

Organizational Psychology Review

<http://opr.sagepub.com/>

Using accumulated knowledge to calibrate theoretical propositions

Jeffrey R. Edwards and Michael S. Christian

Organizational Psychology Review 2014 4: 279 originally published online 20 May 2014

DOI: 10.1177/2041386614535131

The online version of this article can be found at:

<http://opr.sagepub.com/content/4/3/279>

Published by:



<http://www.sagepublications.com>

On behalf of:



e a w o p

[European Association of Work and Organizational Psychology](#)

Additional services and information for *Organizational Psychology Review* can be found at:

Email Alerts: <http://opr.sagepub.com/cgi/alerts>

Subscriptions: <http://opr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://opr.sagepub.com/content/4/3/279.refs.html>

>> [Version of Record](#) - Jul 14, 2014

[OnlineFirst Version of Record](#) - May 20, 2014

[What is This?](#)

Using accumulated knowledge to calibrate theoretical propositions

Organizational Psychology Review

2014, Vol. 4(3) 279–291

© The Author(s) 2014

Reprints and permission:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/2041386614535131

opr.sagepub.com



Jeffrey R. Edwards and Michael S. Christian

University of North Carolina at Chapel Hill, USA

Abstract

In organizational psychology research, most theories put forth directional predictions, such as stating that an increase in one construct will result in an increase or decrease in another construct. Such predictions are imprecise, given that they can be confirmed by a wide range of values, and theories that rely on such predictions bear little risk of falsification. In this article, we present an approach for increasing theoretical precision by using results from meta-analyses to calibrate the predictions embedded in a theory. Our approach provides point values for theoretical predictions along with credibility intervals that capture the likely range of the predicted effects. We illustrate this approach by drawing from research on work engagement and calibrate the predictions represented by two conceptual models. Contributions and limitations of the proposed approach are discussed.

Keywords

Creativity & innovation, fit, statistics/methods

Advancements in organizational psychology research depend heavily on the development and application of strong theory. Much attention has been devoted to guidelines for developing theory (Campbell, 1990; Dubin, 1976; Ferris, Hochwater, & Buckley, 2012; Sutton & Staw, 1995; van Knippenberg, 2011; Webster & Starbuck, 1988; Weick, 1989; Whetten, 1989), and the field

has established journals dedicated to publishing conceptual articles that advance theory, including the *Academy of Management Review*, *Psychological Review*, and *Organizational Psychology Review*. Moreover, developing theory is regarded as evidence of significant scholarly impact (Miner, 2003), characteristic of those who achieve great stature in the field (Smith & Hitt,

Paper received 19 April 2014; revised version accepted 19 April 2014.

Corresponding author:

Jeffrey R. Edwards, The Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3490, USA.

Email: jredwards@unc.edu

2005) and earn recognition for distinguished scientific contributions by associations such as the Academy of Management and the Society for Industrial and Organizational Psychology.

Guidelines for developing strong theory typically emphasize several key criteria. In particular, the constructs that constitute theories should be clearly defined and comprehensively cover the conceptual domain of the theory while maintaining reasonable parsimony. The relationships among theoretical constructs should be fully described, thereby establishing the causal structure of the theory. Boundary conditions should be established to delineate what the theory purports to predict and explain. The theory should be useful, contributing to the understanding of relevant and important phenomena. Perhaps most important, strong theory should explain the conceptual logic underlying the processes represented by the theory, such that the rationale of the theory is explicit and justified.

Although the foregoing criteria provide sound guidance for developing and evaluating theory, they do not address an important but overlooked feature of strong theory, which concerns the *magnitude* of the relationships among theoretical constructs. Most theories in organizational psychology and related fields contain propositions that merely state the expected direction of the relationship between constructs, such as whether an increase in one construct will cause an increase or decrease in another construct (Edwards & Berry, 2010; Gigerenzer, 1998; Meehl, 1990). Directional propositions provide a weak basis for theory testing, given that observing a nonzero effect in empirical research is virtually inevitable (Cohen, 1994; Murphy, 1990), and such effects are increasingly likely to register as statistically significant as studies implement stronger methods, such as larger sample sizes and more reliable measures (Meehl, 1967). More fundamentally, theories that stipulate directional propositions run afoul of the principle that, in order to be viable, theories must be falsifiable (Popper, 1959). As asserted by Rushton (cited

in Platt, 1964, p. 349), “a theory that cannot be mortally endangered cannot be alive.” Although a directional prediction can be falsified by obtaining an effect in the opposite direction, the range of acceptance is wide, because any value that deviates from zero in the predicted direction qualifies as supporting evidence (Edwards & Berry, 2010; Meehl, 1967).

Theories that are limited to directional propositions represent a significant barrier to progress in organizational psychology research. Because such theories yield predictions that tend to evade falsification, they are rarely eliminated from consideration (Edwards & Berry, 2010; Meehl, 1990; Pfeffer, 1993). Moreover, the accumulation of knowledge generated by such theories represents little more than tallying the frequency with which an effect falls in the expected direction, as opposed to homing in on the magnitude of an effect. From an applied standpoint, findings from research based on theories that make directional predictions have limited value, because practitioners and policy makers can derive little value from statements along the lines that increasing X will increase or decrease Y to some unspecified degree. As observed by Tukey (1969, p. 86):

The physical sciences have learned much by storing up amounts, not just directions. If, for example, elasticity had been confined to “When you pull on it, it gets longer!” Hooke’s law, the elastic limit, plasticity, and many other important topics could not have appeared.

Directional predictions are partners in crime with null hypothesis significance tests, which together create a mutually reinforcing cycle of imprecision. Although null hypothesis significance testing has been criticized for decades (e.g., Bakan, 1966; Carver, 1978; Cohen, 1994; Krueger, 2001; Nickerson, 2000; Rozeboom, 1960), its survival is partly due to the fact that most theories in organizational psychology simply predict deviations from the null in a positive or negative direction, which is

what null hypothesis tests are designed to detect. By drawing from theories that make such feeble predictions, there is little wonder that we continue to rely on statistical tests that repeatedly flog the null hypothesis, even when it can be declared dead on arrival (Cohen, 1994; Gigerenzer, 1998; Murphy, 1990).

One way to extricate ourselves from this vicious cycle is to develop theories that offer predictions that are more precise than directional statements (Edwards & Berry, 2010). The approach we examine in this article involves the development of propositions that articulate the expected range of the magnitude of each relationship among the variables in a theory, in the spirit of the “good-enough belt” advocated by Serlin and Lapsley (1985). Of the various ways to develop range predictions, one that holds particular promise is to draw from accumulated empirical evidence about relationships that are conceptually consistent with those in a theory under development. This evidence can be found in meta-analyses, which are pervasive in the organizational psychology literature. Although meta-analyses are typically used as little more than sources for summarizing an area of research or documenting the prevailing direction of an effect (Carlson & Ji, 2011), most contain information that can be used to calibrate the expected magnitude of a relationship indicated by a theoretical proposition. This information constitutes a vast untapped resource that can be leveraged to increase the precision of theoretical predictions. In this article, we discuss how results from meta-analyses can be used to calibrate theoretical predictions and provide an example that draws from published meta-analyses.

Using results from meta-analyses to calibrate theoretical predictions

Meta-analysis has become firmly established in organizational psychology research as the

primary method for synthesizing quantitative findings from primary studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; Cooper, Hedges, & Valentine, 2009; Raju, Burke, Normand, & Langlois, 1991; Schmidt & Hunter, 2014). In general, the goal of meta-analysis is to derive summary estimates of effect sizes along with intervals that describe the uncertainty of these estimates and the potential ranges the effects might display. These estimates can be corrected for various artifacts, such as sampling variability, measurement error, and range restriction, which tends to strengthen the correspondence between meta-analytic estimates and the population parameters they are intended to represent.

Various meta-analytic approaches have been developed to summarize effect sizes and derive their associated intervals. Typically, effect sizes are summarized by computing a weighted average of the effects found in primary studies, where the weights are the sample sizes of each study (Hedges & Olkin, 1985; Raju et al., 1991; Schmidt & Hunter, 2014) or within-study variance, which is itself an inverse function of sample size (Borenstein et al., 2009). Summary effect size estimates are usually accompanied by a confidence interval, which is derived from the standard error of the effect and thus captures the precision of the estimate (Borenstein et al., 2009; Raju et al., 1991; Schmidt & Hunter, 2014). All else being equal, the confidence interval around an effect size estimate decreases as the combined sample sizes of the primary studies increase. Confidence intervals should be distinguished from credibility intervals, which are derived from the standard deviation of the corrected effects and thus capture the variation of population effect size estimates after sampling variability has been removed (Whitener, 1990). Prediction intervals integrate the information contained in confidence and credibility intervals to indicate the expected range of an effect for the next study to be conducted, taking into account sampling variability as well as the range of population effect

sizes (Borenstein et al., 2009; Higgins, Thompson, & Spiegelhalter, 2009).

Of the various statistics generated by meta-analysis, the most relevant for calibrating theoretical predictions are the summary effect and the credibility interval. The summary effect provides an unbiased estimate of the effect in the population, which can serve as a point value for a theoretical prediction. The credibility interval can be used to specify the range of a theoretical prediction around the point value and set its upper and lower bounds. Because these bounds are corrected for sampling variability, they are appropriate for calibrating predictions at the theoretical level, which refer to effects in the population designated by the theory independent of a particular sample drawn from that population. The summary effect and credibility interval may be corrected for artifacts that can bias estimates in primary studies, such as measurement error and range restriction, given that these artifacts affect empirical research and, in principle, do not exist at the theoretical level.

In some meta-analyses, the credibility interval for the summary effect is zero, meaning there is no variation in the effect. This situation characterizes fixed-effects meta-analysis, which is based on the assumption that a single population effect underlies the effects found in primary studies, which differ from one another only due to sampling variability (Borenstein et al., 2009; Hedges & Vevea, 1998; Schmidt & Hunter, 2014). When a credibility interval is zero, the summary effect can be used to specify a theoretical prediction represented as a single point value. Doing so does *not* imply that a hypothesis derived from the theory to be tested empirically should be stated as a point value, given that virtually any effect estimated in sample data will deviate from a pre-specified point, even if by a very small amount (Cohen, 1994; Murphy, 1990). This deviation is accommodated by placing a prediction interval around the summary effect which, as noted earlier, specifies the expected

range of an effect in sample randomly drawn from the population (Borenstein et al., 2009; Higgins et al., 2009). A prediction interval around a hypothesized point value is consistent with the good-enough belt advocated by Serlin and Lapsley (1985), which specifies the range within which effects are expected to fall in empirical research. When a credibility interval is nonzero, its corresponding prediction interval is wider because it takes sampling variability into account (Borenstein et al., 2009). Because our goal is to *calibrate* theoretical predictions, we focus on credibility intervals while noting that prediction intervals should be used in studies designed to *test* the predictions of a theory.

An illustrative example

To demonstrate the use of summary effects and credibility intervals to calibrate theoretical predictions, we consider research on work engagement, which refers to the investment of personal energy in the experience or performance of work (Christian, Garza, & Slaughter, 2011; Kahn, 1990; Macey & Schneider, 2008; Rich, LePine, & Crawford, 2010; Schaufeli, Salanova, Gonzalez-Roma, & Bakker, 2002). This research suggests that work engagement is influenced by various individual and situational factors and, in turn, affects outcomes such as job-related attitudes and performance. For this illustration, we focus on five constructs: task significance, transformational leadership, conscientiousness, work engagement, and task performance. We specify the relationships among these constructs using two conceptual models, one in which the effects of task significance, transformational leadership, and conscientiousness on task performance are fully mediated by work engagement, and another in which these effects are partially mediated by work engagement (see Figure 1).

To calibrate the predictions in these theoretical models, we draw from a meta-analysis

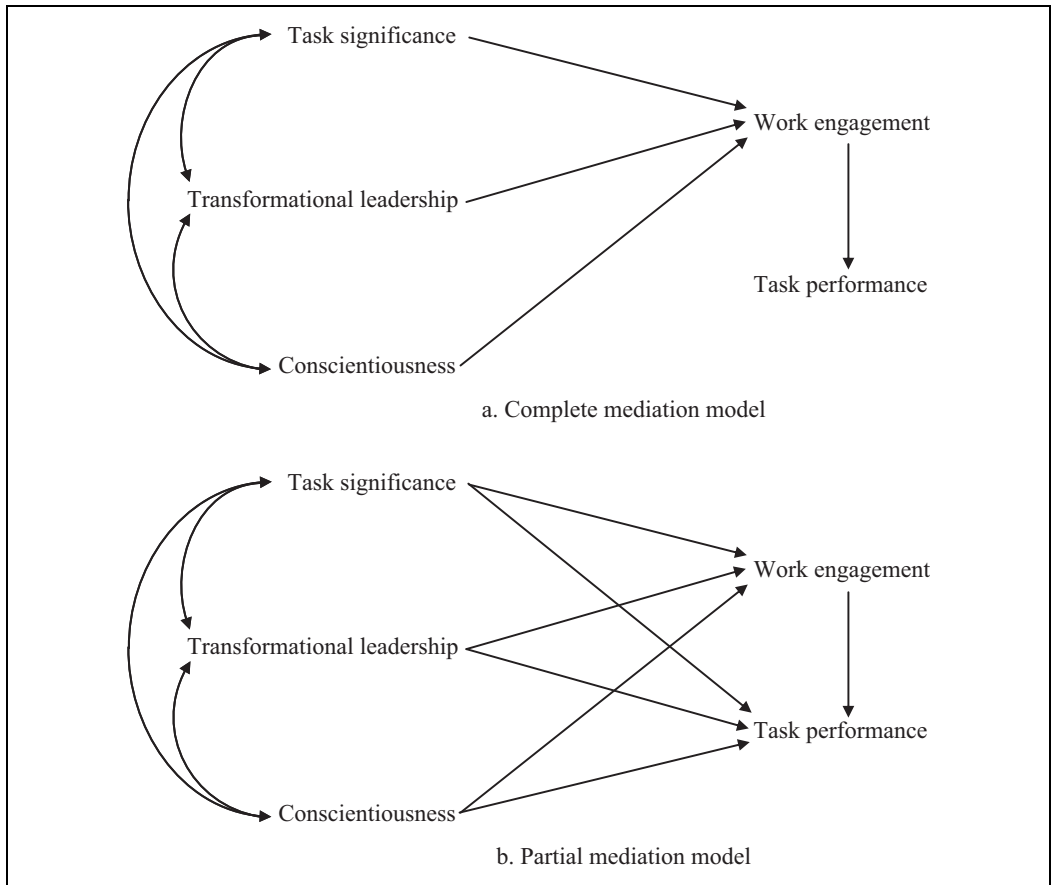


Figure 1. Complete and partial mediation models relating work engagement to task significance, transformational leadership, conscientiousness, and task performance.

conducted by Christian et al. (2011) using the approach developed by Raju et al. (1991). We used information collected for this meta-analysis to derive population estimates of the correlations among the five constructs of interest, computed as the sample-size weighted means of the correlations from the primary studies (see Table 1). We corrected the correlations from the primary studies for measurement error individually for each study, which takes into account the sampling variability of the reliability estimates used (Burke & Landis, 2003). We also computed 80% credibility intervals, as shown in Table 1. We used intervals of

this width in accordance with conventions in meta-analysis, although intervals that are narrower or wider can be used, which would respectively increase or decrease the precision of the associated theoretical propositions. All credibility intervals were nonzero, ranging in width from 0.16 to 0.42. The total sample sizes for the mean correlations and associated credibility intervals ranged from 777 to 12,893 with a harmonic mean of 2,701.

If the predictions specified by a theory were limited to simple bivariate associations, then the point values and intervals in Table 1 could be used for calibration purposes. However,

Table 1. Correlations and bounds for meta-analytic correlations.

	1	2	3	4	5
1. Task significance	1.00				
2. Transformational leadership	0.29 (.21, .37)	1.00			
3. Conscientiousness	0.15 (.06, .24)	0.07 (-.02, .16)	1.00		
4. Work engagement	0.51 (.43, .59)	0.27 (.19, .35)	0.42 (.30, .54)	1.00	
5. Task performance	0.23 (.02, .44)	0.20 (.12, .28)	0.23 (.10, .36)	0.39 (.27, .51)	1.00

Note. $H = 2701$ (H refers to the harmonic mean). All correlations are corrected for measurement error and range restriction. The numbers in parentheses below each correlation are the lower and upper bounds of the 80% credibility interval.

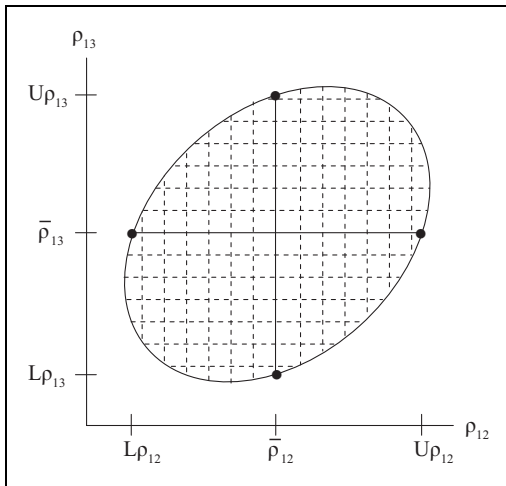


Figure 2. Hypothetical credibility region for two summary effects.

most theories specify causal relationships among multiple constructs, such as those depicted by the models in Figure 1. For such theories, predictions should be calibrated not based on bivariate associations, but instead using partialled relationships that are consistent with the structure of the theory. For instance, according to the models in Figure 1, the effect of task significance on work engagement should be calibrated not using the correlation of .51 and credibility interval of (.43, .59) in Table 1, but instead using information that takes into account the correlation of task significance with transformational leadership and conscientiousness

along with the effects of these two constructs on work engagement. To obtain this information, the mean correlations and credibility intervals in Table 1 were used as input into structural equation models specified according to the two models in Figure 1. We used this approach not to empirically test the models, but instead to compute the coefficients used to calibrate the paths of the models.

Point values for calibrating the theoretical relationships were generated by estimating models that used the mean correlations in Table 1 as input. Credibility intervals for the relationships were derived by sequentially replacing each mean correlation with the lower and upper bounds of its credibility interval, leaving the remaining mean correlations at their point values. This approach was based on the principle that the credibility interval for each mean correlation is at its widest when the other correlations are at their means. This principle is illustrated by the graph in Figure 2. The axes labeled ρ_{12} and ρ_{13} represent two population correlations, the average values on each axis are the summary effects, and the lower and upper bounds of the credibility intervals of the correlations are labeled $L\rho_{12}$ and $U\rho_{12}$ for ρ_{12} and $L\rho_{13}$ and $U\rho_{13}$ for ρ_{13} . The ellipse signifies the credibility region of ρ_{12} and ρ_{13} considered jointly and is slanted to represent the general case in which ρ_{12} and ρ_{13} are not independent. The solid lines crossing the ellipse correspond to the credibility intervals of ρ_{12} and ρ_{13} , and the dashed lines run parallel to the solid lines. When

Table 2. Point estimates and bounds for complete and partial mediation models.

		Point value	Lower bound	Upper bound
Complete mediation model				
Task significance	Work engagement	0.42	0.33	0.51
Transformational leadership	Work engagement	0.12	0.04	0.21
Conscientiousness	Work engagement	0.35	0.23	0.47
Work engagement	Task performance	0.39	0.27	0.51
Partial mediation model				
Task significance	Work engagement	0.42	0.33	0.51
Transformational leadership	Work engagement	0.12	0.04	0.21
Conscientiousness	Work engagement	0.35	0.23	0.47
Task significance	Task performance	0.03	-0.27	0.32
Transformational leadership	Task performance	0.10	0.01	0.19
Conscientiousness	Task performance	0.09	-0.07	0.25
Work engagement	Task performance	0.31	0.11	0.51

Note. Table entries are standardized paths for the models shown in Figure 1.

projected onto the axes, the dashed lines show that, for each correlation, the width of each credibility interval depends on the level of the other correlation. For instance, the horizontal lines show that the credibility interval for ρ_{12} becomes narrower as ρ_{13} deviates from its mean, given that each dashed line is shorter than the solid line. This principle also holds for the credibility interval of ρ_{13} , as can be seen by comparing the lengths of the solid and dashed vertical lines. The mathematical basis for this principle rests on the fact that, for any set of parallel chords crossing an ellipse, the longest chord is the diameter, which passes through the center of the ellipse.¹ This principle generalizes to higher dimensions, such as the 10-dimensional space corresponding to the 10 mean correlations in Table 1, and it holds regardless of whether the mean correlations are dependent or independent. Replacing the point value of each mean correlation with the lower and upper bounds of its credibility interval yielded 20 correlation matrices, which were used as input to compute paths for the models in Figure 1. From these results, we identified the minimum and maximum values of each path and used them as the lower and upper bounds, respectively, of the relationships in the two models.

The calibrated relationships for the models in Figure 1 are provided in Table 2. For the complete mediation model, the point values of the effects of task significance, transformational leadership, and conscientiousness on work engagement were 0.42, 0.12, and 0.35, respectively. These values are lower than the corresponding mean correlations in Table 1, which can be attributed to the fact that the results in Table 2 are partialled rather than bivariate relationships. The sizes of the credibility intervals for the effects were 0.18, 0.17, and 0.24, respectively, which are similar to those for the mean correlations in Table 1. The point value and credibility interval for the effect of work engagement on task performance is the same as those for the mean correlation in Table 1, given that the complete mediation model depicts this relationship as a bivariate association from which no other effects are partialled.

For the partial mediation model, the point values and credibility intervals for the effects of task significance, transformational leadership, and conscientiousness on work engagement are the same as those for the complete mediation model, given that both models specify these effects in the same manner (i.e., these three constructs are the causes of work engagement).

However, with the partial mediation model, work engagement is now accompanied by task significance, transformational leadership, and conscientiousness as causes of task performance. The point value for the effect of work engagement on task performance decreased slightly from 0.39 to 0.31, and the width of the credibility interval became increased from 0.24 to 0.40. The point values of the effects of task significance, transformational leadership, and conscientiousness on task performance were small, ranging from 0.03 to 0.10, and their credibility intervals ranged in width from 0.18 to 0.59. Across both models, the credibility intervals for all but two paths excluded zero, and both of these paths were less than 0.10 in magnitude.

Discussion

Developing theories with propositions that are more precise than mere directional statements promises to enhance theoretical progress in organizational psychology research. In this article, we have presented an approach to increasing the precision of theoretical propositions by drawing from quantitative summaries of effect sizes and credibility intervals available in meta-analyses, which provide a vast untapped resource for calibrating theoretical predictions. We used this approach to calibrate predictions for two conceptual models that involve the causes and outcomes of work engagement. Our approach incorporated mean correlations and credibility intervals with the causal structures of the models to specify the magnitudes of the predictions embedded in these models and the lower and upper bounds of these predictions. The approach we presented can be used to calibrate theoretical predictions in other conceptual domains that involve different constructs and pose alternative causal models.

Although the approach presented here bears some similarity to meta-analytic structural equation modeling (Cheung & Chan, 2005;

Viswesvaran & Ones, 1995), our approach is not intended to test models, but instead is designed to calibrate theoretical predictions to be tested in subsequent empirical research. As noted earlier, our application of structural equation modeling was merely a convenient technique to compute point values and credibility intervals for the relationships specified by a theory. Because this type of application is a computational shortcut rather than an empirical test, we are not concerned with model fit, tests of parameter estimates, and other issues that merit attention in conventional applications of structural equation modeling.

Although our approach is designed to calibrate relationships in theoretical models, it can be modified to specify point values and ranges for predictions to be tested in empirical studies. This can be accomplished by replacing the credibility intervals used in our approach with prediction intervals which, as noted earlier, provide the expected range of effect sizes for the next empirical study, taking into account the variability of population effect sizes and variance due to sampling error (Borenstein et al., 2009; Higgins et al., 2009). Because prediction intervals incorporate both of these sources of variation, they are wider than either confidence intervals or credibility intervals. For our illustrative examples, the 80% prediction intervals for the mean correlations in Table 1 were only slightly wider than the 80% credibility intervals, with an average increase in width of 0.02. The bounds of prediction intervals can be used to derive the expected range of empirical relationships using the procedure we demonstrated to obtain bounds of the theoretical relationships for the models in our example. The obtained prediction intervals can be compared to empirical results using procedures for testing range hypotheses (Nickerson, 2000; Serlin & Lapsley, 1985)

Admittedly, the approach presented here has several limitations. First, it requires meta-analytic evidence that is relevant to the theory under development. As such, researchers should

carefully evaluate the correspondence between the sampling, method, and design of the primary studies and the substance of the theoretical model under development (Cooper, 2010). For theories that strive for novelty, it might be difficult to locate meta-analyses that provide results needed to calibrate theoretical predictions. In such cases, researchers might resort to meta-analyses that involve constructs that are conceptually similar to those in the focal theory, acknowledging that the calibrated predictions likely represent rough approximations. Second, our illustration represented population effect sizes as correlations. When used as input for structural equation models, correlation matrices can implicitly modify the model being analyzed, a problem that is avoided by using covariance matrices (Cudeck, 1989). Nonetheless, correlations provide a convenient metric for quantifying effect sizes when the studies involved measure variables on different scales, as is generally the case in meta-analysis. On balance, we believe that correlations provide a reasonable basis for calibrating theoretical predictions, although we recognize that other metrics for quantifying effect sizes can be used (Babakus & Ferguson, 1988; Cortina & Nouri, 2000; Grissom & Kim, 2005; Liebetrau, 1983). Third, the correlations derived from meta-analyses involve samples drawn from different studies, and the resulting correlation matrices might not be positive definite, which can yield aberrant results such as R^2 values that exceed unity. Moreover, the primary studies might involve samples that differ in substantive respects, which would undermine the treatment of the correlations as a unified set. This limitation applies to meta-analytic structural equation modeling in general, as opposed to our approach in particular, but it nonetheless merits attention.

We should add that theoretical propositions calibrated using our approach are only as valid as the primary studies used for meta-analyses and the methods used to conduct these analyses. Although meta-analysis can address some shortcomings of primary studies, such

as measurement error and range restriction, many types of methodological flaws are not remedied when primary studies are combined using meta-analysis (Aguinis, Pierce, Bosco, Dalton, & Dalton, 2011; Eysenck, 1978; Sharpe, 1997). In addition, meta-analysis relies on statistical assumptions, such as whether effects are fixed or random (Hedges & Vevea, 1998) and the notion that random effects are normally distributed (Higgins et al., 2009; Kraemer & Andrews, 1982), which can be violated in practice. Moreover, meta-analysis requires numerous judgment calls (Wanous, Sullivan, & Malinak, 1989), although the effects of these decisions on effect size estimates are apparently less than generally believed (Aguinis, Dalton, Bosco, Pierce, & Dalton, 2011). These and other factors that impact the results of meta-analyses should be carefully evaluated as a precursor to using these results to calibrate theoretical propositions.

We acknowledge that our approach might engender skepticism among organizational psychology researchers. The premise of calibrating theoretical predictions could be viewed as assuming a level of precision that cannot be achieved in psychology, given the nature of the phenomena under study. However, researchers implicitly calibrate an effect size when they designate it as small, medium, or large (Cohen, 1992) or evaluate the practical significance of an effect (Kirk, 1996). Moreover, the primary purpose of meta-analysis is to calibrate effect sizes and gauge their uncertainty and potential ranges, and there is no shortage of meta-analyses in the organizational psychology literature. With regard to calibrating theoretical predictions, the primary question is not whether we have the required information, but whether we will choose to use that information to make our theories more precise.

Finally, the approach outlined here should be considered one of several ways to increase the precision of theoretical propositions (Edwards & Berry, 2010). Other approaches rely on the conceptual reasoning used to explain the relationships

in a theory. For instance, when a proposition is supported by several distinct theoretical arguments that supplement one another, it is reasonable to predict that associated effect will be stronger than an effect that is based on a single theoretical rationale. In addition, when a proposition refers to a relationship between constructs that are causally proximal rather than distal, then the relationship is likely to be relatively strong, because distal relationships effectively operate through mediating mechanisms that dampen the magnitude of the relationship. This reasoning is manifested in a conceptual model of workplace safety presented by Christian, Bradley, Wallace, and Burke (2009), in which distal person and situation factors (e.g., personality, climate) influenced proximal person factors (e.g., safety motivation and knowledge), which in turn affected safety performance. Their meta-analysis supported the prediction that proximal factors have larger effects than factors that are conceptually more distal from safety performance. Theoretical propositions can also be informed by practical considerations, which can be used to stipulate the range of effect sizes that would be considered meaningful and useful (Fowler, 1985; Kirk, 1996). Thus, we encourage researchers to draw from empirical, conceptual, and practical approaches to increase the precision of theoretical propositions. Doing so would help us move beyond directional predictions that say little more than whether we expect to fall off a log to the left or the right. Taking these steps to increase theoretical precision can greatly enhance the accumulation of knowledge in organizational psychology research.

Acknowledgements

The authors thank Michael Burke, Marcus Butts, Barry Cipra, Daan van Knippenberg, Huy Le, In-Sue Oh, Hannah Rothstein, and Frank Schmidt for their comments and suggestions during the development of this article. Data and syntax files used for the illustrative example can be obtained from the first author.

Note

1. We are indebted to Barry Cipra for explaining the rationale for this principle to us.

References

- Aguinis, H., Dalton, D. A., Bosco, F. A., Pierce, C. A., & Dalton, C. M. (2011). Meta-analytic choices and judgment calls: Implications for theory building and testing, obtained effect sizes, and scholarly impact. *Journal of Management, 37*, 5–38.
- Aguinis, H., Pierce, C. A., Bosco, F. A., Dalton, D. R., & Dalton, C. M. (2011). Debunking myths and urban legends about meta-analysis. *Organizational Research Methods, 14*, 306–331.
- Babakus, E., & Ferguson, C. E., Jr. (1988). On choosing the appropriate measure of association when analyzing rating scale data. *Journal of the Academy of Marketing Science, 16*, 95–102.
- Bakan, D. (1966). The test of significance in psychological research. *Psychological Bulletin, 66*, 423–437.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley.
- Burke, M. J., & Landis, R. S. (2003). Methodological and conceptual challenges in conducting and interpreting meta-analyses. In K. Murphy (Ed.), *Validity generalization: A critical review* (pp. 287–309). Mahwah, NJ: Erlbaum.
- Campbell, J. P. (1990). The role of theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 39–73). Palo Alto, CA: Consulting Psychologists Press.
- Carlson, K. D., & Ji, F. X. (2011). Citing and building on meta-analytic findings: A review and recommendations. *Organizational Research Methods, 14*, 696–717.
- Carver, R. P. (1978). The case against statistical significance testing. *Harvard Education Review, 48*, 378–399.
- Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-

- stage approach. *Psychological Methods*, 10, 40–64.
- Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology*, 94, 1103–1127.
- Christian, M. S., Garza, A. S., & Slaughter, J. E. (2011). Work engagement: A quantitative review and test of its relations with task and contextual performance. *Personnel Psychology*, 64, 89–136.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997–1003.
- Cooper, H. M. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Cooper, H. M., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.
- Cortina, J. M., & Nouri, H. (2000). *Computing effect sizes from complex ANOVA designs*. Thousand Oaks, CA: Sage.
- Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, 105, 317–327.
- Dubin, R. (1976). Theory building in applied areas. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 17–39). Chicago, IL: Rand McNally.
- Edwards, J. R., & Berry, J. W. (2010). The presence of something or the absence of nothing: Increasing theoretical precision in management research. *Organizational Research Methods*, 13, 668–689.
- Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517.
- Ferris, G. R., Hochwater, W. A., & Buckley, M. R. (2012). Theory in the organizational sciences: How will we know it when we see it? *Organizational Psychology Review*, 2, 94–106.
- Fowler, R. L. (1985). Testing for substantive significance in applied research by specifying nonzero effect null hypotheses. *Journal of Applied Psychology*, 70, 215–218.
- Gigerenzer, G. (1998). Surrogates for theories. *Theory and Psychology*, 8, 195–204.
- Grissom, R., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486–504.
- Higgins, J. P. T., Thompson, S. G., & Spiegelhalter, D. J. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society*, 172, 137–159.
- Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, 33, 692–724.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56, 746–759.
- Kraemer, H. C., & Andrews, G. (1982). A nonparametric technique for meta-analysis effect size calculation. *Psychological Bulletin*, 91, 404–412.
- Krueger, J. (2001). Null hypothesis significance testing: On the survival of a flawed method. *American Psychologist*, 56, 16–26.
- Liebetrau, A. M. (1983). *Measures of association*. Newbury Park, CA: Sage.
- Macey, W., & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology*, 1, 3–30.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34, 103–115.
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1, 108–141.
- Miner, J. B. (2003). The rated importance, scientific validity, and practical usefulness of organizational behavior theories: A quantitative review. *Academy of Management Learning & Education*, 2, 250–268.
- Murphy, K. R. (1990). If the null hypothesis is impossible, why test it? *American Psychologist*, 45, 403–404.

- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*, 241–301.
- Pfeffer, J. (1993). Barriers to the advance of organizational science: Paradigm development as a dependent variable. *Academy of Management Review, 18*, 599–620.
- Platt, J. R. (1964). Strong inference. *Science, 146*, 347–353.
- Popper, K. R. (1959). *The logic of scientific discovery*. New York, NY: Basic Books.
- Raju, N. S., Burke, M. J., Normand, J., & Langlois, G. M. (1991). A new meta-analytic approach. *Journal of Applied Psychology, 76*, 432–446.
- Rich, B. L., LePine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal, 53*, 617–635.
- Rozeboom, W. W. (1960). The fallacy of the null hypothesis significance test. *Psychological Bulletin, 57*, 416–428.
- Schaufeli, W. B., Salanova, M., Gonzalez-Roma, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies, 3*, 71–92.
- Schmidt, F., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.
- Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good-enough principle. *American Psychologist, 40*, 73–83.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*, 881–901.
- Smith, K. G., & Hitt, M. A. (2005). *Great minds in management: The process of theory development*. New York, NY: Oxford University Press.
- Sutton, R. I., & Staw, B. M. (1995). What theory is not. *Administrative Science Quarterly, 40*, 371–384.
- Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist, 24*, 83–91.
- Van Knippenberg, D. (2011). Advancing theory in organizational psychology. *Organizational Psychology Review, 1*, 3–8.
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology, 48*, 865–886.
- Wanous, J. P., Sullivan, S. E., & Malinak, J. (1989). The role of judgment calls in meta-analysis. *Journal of Applied Psychology, 74*, 259–264.
- Webster, J., & Starbuck, W. H. (1988). Theory building in industrial and organizational psychology. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 93–138). Chichester, UK: Wiley.
- Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review, 14*, 516–531.
- Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*, 490–495.
- Whitener, E. M. (1990). Confusion of confidence intervals and credibility intervals in meta-analysis. *Journal of Applied Psychology, 75*, 315–321.

Author biographies

Jeffrey R. Edwards is the Belk Distinguished Professor of Organizational Behavior at the Kenan-Flagler Business School, University of North Carolina. He is past Editor of *Organizational Behavior and Human Decision Processes*, past Chair of the Research Methods Division of the Academy of Management, and a fellow of the Academy of Management, the American Psychological Association, and the Society for Industrial and Organizational Psychology. His research addresses person–environment fit, stress and coping, work–family issues, and methodological topics that include difference scores, polynomial regression, structural equation modeling, construct validity, and the development and evaluation of theory.

Michael S. Christian is an Assistant Professor of organizational behavior at the Kenan-Flagler

Business School at the University of North Carolina at Chapel Hill. His current research is focused on human energy and self-regulation in the workplace, and the psychological drivers of the ebbs and flows in withdrawal, engagement, and unethical behavior. *Administrative Science Quarterly*, *The Academy of Management Journal*, *The*

Journal of Applied Psychology, *Personnel Psychology and Organizational Behavior*, and *Human Decision Processes* have published his research. His work has been covered by media outlets such as *The Wall Street Journal*, *The Financial Times*, *The Washington Post*, and CNN.