A Rebuttal to Borchgrevink's Criticism of Our Meta-Analytic Recommendations

Michael Lynn, Cornell University
Brian Mullen, Syracuse University

Borchgrevink presented flawed arguments against our (Lynn & Mullen) previous recommendation that hospitality researchers employ meta-analytic techniques other than those of Hunter and Schmidt. The problems with his criticisms are briefly discussed in this article.

Borchgrevink (1998 [this issue]) takes issue with our recent recommendation (Lynn & Mullen, 1997) that hospitality researchers employ either the Hedges and Olkin (1985) or Rosenthal (1991) techniques of meta-analysis rather than the techniques of Hunter and Schmidt (1990). He presents three arguments supporting use of the Hunter and Schmidt techniques. First, he claims that our evidence against the Hunter and Schmidt approach to meta-analysis (a study by Johnson, Mullen, & Salas, 1995) is fatally flawed. Second, he suggests that the Hunter and Schmidt approach is superior to the other major approaches, because the former approach uses a random-effects model, whereas the other approaches use fixed-effects models. Finally, he implies that the Hunter and Schmidt approach is to be recommended because it emphasizes the use of confidence intervals rather than significance testing. We respond to each of these arguments in the paragraphs below.

Johnson Et Al.'s Study

Our recommendation against using the Hunter and Schmidt (1990) techniques of meta-analysis was based on the results of a study by Johnson et al. (1995). When using the Hedges and Olkin (1985) and Rosenthal (1991) techniques of meta-analysis, Johnson et al. found that the statistical significance of the mean effect size behaved in a reasonable way-for example, consistent with the law of large numbers, it increased with the number of studies in the meta-analytic database and with the average sample size for the included studies. Furthermore, they found that the statistical significance of predictors of effect size also behaved as statistical conventions suggest it should-for example, it

increased with the average sample size of the included studies and with the range of effect sizes in the database. This pattern of results was not observed when Johnson et al. used the Hunter and Schmidt techniques of meta-analysis. Given that the Hunter and Schmidt approach to meta-analysis produced results that failed to conform with statistical convention and with the results of other meta-analytic approaches, it seems prudent to use the other approaches.

Citing an unpublished study (i.e., Schmidt & Hunter, 1997) as evidence, Borchgrevink (1998) argues that the discrepancies in results between the meta-analytic approaches found by Johnson et al. (1995) are attributable to two fatal flaws in their methodology. First, he argues that Johnson et al. employed an incorrect error term when using the Hunter and Schmidt (1990) techniques to test the significance of the mean effect size-that is, they failed to divide the standard deviation of the observed correlations by the square root of the number of studies as Hunter and Schmidt advocate. Second, he argues that Johnson et al. violated an assumption of the fixed-effects model underlying the Hedges and Olkin (1985) and Rosenthal (1991) techniques by applying these meta-analytic techniques to a data set with heterogeneous population effect sizes. We are hesitant to critique an unpublished study that we have not seen, but there are serious problems with the arguments presented by Borchgrevink.

First, only some of the anomalous findings produced by the Hunter and Schmidt (1990) techniques can be explained by Johnson et al.'s (1995) use of the disputed error term. Dividing the error term that Johnson et al. used by the square root of the number of studies in the database would not correct the failure of the Hunter and Schmidt techniques to find that the significance of the mean effect size increases with the average sample size of the studies. Nor would it correct the Hunter and Schmidt technique's tendency to produce anomalous results concerning significance tests of predictors of effect size, because these tests do not involve the disputed error term.

Second, fixed-effects meta-analytic models require homogeneous population effect sizes only to be comparable to random-effects models. This assumption is not necessary to make the use of fixed-effects models appropriate (see Hedges, 1994a, 1994b for a discussion of when fixed- and random-effects models should be used). Thus, Borchgrevink (1998) is correct when he argues that heterogeneous population effects contributed to some of the different results produced by the different meta-analytic techniques in Johnson et al.'s (1995) study. In fact, Johnson et al. explicitly tested and demonstrated this point. However, heterogeneous population effects do not invalidate the results of the Hedges and Olkin (1985) or Rosenthal (1991) techniques as Borchgrevink suggests they do.

Borchgrevink, (1998) suggests that use of the Hunter and Schmidt (1990) techniques of meta-analysis should be encouraged because they use a random-effects model, which has been

recommended by the National Research Council (1992). The National Research Council's recommendation notwithstanding, there are two problems with this argument. First, random-effects models are not necessarily superior, or preferable, to fixed-effects models. Fixed-effects meta-analytic models provide greater statistical power than do random-effects models, and they are perfectly appropriate whenever (a) the tested model explains all the variation in effect size parameters, (b) the unaccounted-for variation in effect size parameters is not regarded as random, and/or (c) the results are statistically generalized only to studies just like those in the meta-analysis (Hedges, 1994b).

Second, the Hunter and Schmidt (1990) approach is not the only, or even the best, meta-analytic technique using a random-effects model. In fact, Erez, Bloom, and Wells (1996) argued that the Hunter and Schmidt approach is really a fixed-effects model because it does not allow true (nonartifactual) between-studies differences to influence parameter estimates. They extended the Rosenthal (1991) meta-analytic approach to include random effects and demonstrated that this model provided more accurate parameter estimates than did the Hunter and Schmidt model. Thus, even should a meta-analyst desire to use a random-effects model, he or she has an alternative that is better than the Hunter and Schmidt techniques.

Statistical Significance Testing

Borchgrevink (1998) points out that Hunter and Schmidt (1990) advocate the use of confidence intervals over the use of significance tests. He seems to suggest (although he does not explicitly state) that this is a reason for using the Hunter and Schmidt techniques. However, the null hypothesis statistical test is not as bad as its critics claim-see Hagan (1997) for a strong defense of this scientific tool. Furthermore, the Hunter and Schmidt approach to meta-analysis is not the only one that provides confidence intervals-the Hedges and Olkin (1985) approach does so as well. Thus, Hunter and Schmidt's emphasis on confidence intervals over significance tests is not a valid reason for using their meta-analytic techniques.

Conclusions

Borchgrevink's (1998) defense of the Hunter and Schmidt (1990) meta-analytic techniques is flawed. We find no reason to change our (Lynn & Mullen, 1997) recommendation that hospitality researchers use the Hedges and Olkin (1985) or Rosenthal (1991) techniques rather than those of Hunter and Schmidt. However, the major purpose of our original article was not to criticize Hunter and

Schmidt but to promote the use of meta-analysis in hospitality research. Thus, we welcome this exchange of ideas and hope that it will encourage scholars in the hospitality field to read and learn more about meta-analysis.

References

Borchgrevink. C.P. (1998). Taking issue with Lynn and Mullen's meta-analytic recommendations. *Journal of Hospitality* & *Tourism Research,* 22(1), 98-101.

Erez, A., Bloom. *M.* C., & Wells, M. T. (1996). Using random rather than fixed *effects* models in meta-analysis: Implications for situational specificity and validity generalization. *Personnel Psychology,* 49, 275-306.

Hagan. R. L. (1997). In praise of the null hypothesis statistical test. *American Psychologist, 52,* 15-24.

Hedges, L. V. (1994a). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 285-300). New York Russell Sage.

Hedges, L.V. (1994b). Statistical considerations. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 29-38). New York: Russell Sage.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for* meta-analysis. Orlando, FL: Academic Press.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA Sage.

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three meta-analytic approaches. *Journal of Applied Psychology,* 80, 94-106.

Lynn, *M.,* & Mullen, B. (1997). The quantitative integration of research: *An* introduction to meta-analysis. *Journal of Hospitality* & *Tourism Research,* 21, 121-139.

National Research Council. (1992). Combining information: Statistical issues and opportunities for research. Washington, DC: National Academy Press.

Rosenthal, R. (1991). *Meta-analytic* procedures *for social research.* Newbury Park, CA Sage.

Schmidt, F. L., & Hunter, J. E. (1997). *Comparison of three meta-analysis methods revisited:* An error *in Johnson, Mullen and Salas (1995).* Unpublished manuscript, College of Business, University of Iowa.