A Comparison of Meta-Analytic Approaches on the Consequences of Role Stressors

by

Edward Rickamer Hoover

(Under the direction of Cheolwoo Park)

Abstract

The use of conducting meta-analytic work is on the rise, yet analysts tend to overlook differences between approaches and simply select one that is common to the field of study. This study contrasts two popular meta-analytic approaches based on work by Hedges and Olkin (1985) and Hunter, Schmidt, and Jackson (1982) and compares them both in theory and in application. Broadly, while Hedges and Olkin (1985) first corrects for statistical biases, Hunter et al. (1982) uses the biased estimates but corrects for statistical artifacts prior to integration. Conceptual and statistical differences between these approaches lead to numerous disparities between estimates. Ultimately, it is difficult to provide guidelines for which meta-analytic approach is the best as it differs by scenario and the information available. This decision must be made holistically by examining the assumptions behind the nature of the constructs in question, the characteristics of the data set, and possible types of inaccuracies that must most be guarded against.

Index words: Effect size, Fixed effects, Job satisfaction, Meta-analysis, Random effects, Role ambiguity

A Comparison of Meta-Analytic Approaches on the Consequences of Role Stressors

by

Edward Rickamer Hoover

B.S., University of Georgia, 2001

M.A., University of South Florida, 2007

Ph.D., University of South Florida, 2009

A Thesis Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

Master of Science

Athens, Georgia

2013

A Comparison of Meta-Analytic Approaches
on the Consequences of Role Stressors

by

Edward Rickamer Hoover

Approved: June 7th, 2013

Professor: Cheolwoo Park

Committee: Paul Schiekelman
Jaxk Reeves

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2013

A Comparison of Meta-Analytic Approaches on the Consequences of Role Stressors

Edward Rickamer Hoover

June 11, 2013

Acknowledgements

I am grateful for the guidance of my major professor and friend Cheolwoo Park. His continual support and patience made this project successful.

Contents

Chapter 1

Introduction

A meta-analysis is "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating findings" (Glass, 1976). It is a quantified summary of existing research that augments more classical and traditional qualitative narrative summaries and reviews. The rise in popularity of such methods is in response to the overwhelming number of individual studies being produced in many disciplines. The benefit of a high quality meta-analysis is that it imparts a great deal of information concisely.

The goal of this study is to provide a general overview of the meta-analysis field with exploration into various techniques, and to do so at an approachable level. To accomplish this goal, this study explores and contrasts two of the most popular approaches developed by Hedges and Olkin (1985) and Hunter, Schmidt, and Jackson (1982). Broadly, Hedges and Olkin (1985) first transforms individual effect sizes to unbiased estimators prior to integration. In contrast, the Hunter et al. (1982) technique uses biased estimates but attempts to correct them for artifacts, such as measurement error and unreliability, prior to integration. This study begins with an investigation into the statistical underpinnings behind each approach, including preliminary issues such as transforming individual studies for preparation for integration, calculating effect size, testing significance, and examining homogeneity of effect sizes. Next, these topics are compared in application. In order to motivate the discussion, 10 studies examining the influence of role ambiguity on job satisfaction were analyzed and contrasted. Finally, in an exploratory vein, the magnitude of the studies' effect sizes, the number of participants in individual samples, and the number of studies composing the meta-analysis will were artificially manipulated in order to gain a deeper understanding of the influence of these factors on meta-analytic results.

The unique contribution of this study is that it provides both a statistical discussion and concrete exploration of the differences between two prominent meta-analytic techniques. In alignment with this contribution, Hedges and Olkin (1985) and Hunter et al. (1982) meta-analysis approaches and relevant issues are discussed, the data set employed provided, and the relevant syntax is appended. It is desired to provide a readily accessible resources to those less familiar with meta-analysis.

Chapter 2

Literature Review

This chapter begins by introducing the topic of meta-analysis. Next, the strengths and weaknesses inherent in the meta-analysis methodology (Section 2.1.1) are discussed. In Section 2.1.3 two popular approaches are introduced, that of Hedges and Olkin (1985) (Section 2.1.4) and Hunter et al. (1982) (Section 2.1.5). For each, topics such as calculating weighted average effect size, level of significance, and homogeneity are discussed. Section 2.1.6 provides some initial comparisons between these two approaches. The chapter ends with a brief introduction to role ambiguity and job satisfaction which constitute the motivating example (Section 2.2).

2.1 Meta-Analysis

A meta-analysis begins with a systematic review of all relevant studies in order to develop a balanced and impartial pool of individual studies. The effects of the various studies are then integrated, depending on technique and underlying assumptions. By combining the results of similar studies, the precision of estimates and power to detect treatment effects are strengthened. In addition, variability may be examined to identify potential unidentified moderators. This allows the analyst to identify patterns and other interesting relationships.

2.1.1 Strength and Weakness of Meta-Analytic Technique

The benefit of a well-conducted meta-analysis is that it provides a great deal of information concisely. Specifically, it provides a weighted estimate that more accurately reflects the true effect size influencing the variables under consideration than the individual studies from which it is composed. Furthermore, as the "true" effect is being better estimated, this value is more generalizable in application.

A primary tenet of the meta-analysis strategy is that the included studies are a representative sample of the population of studies. While missing at random is acceptable, non-random missing data may bias estimates. However, there are many types of biases that may diminish or even corrupt meta-analytic results. Some commonly mentioned biases include publication bias, search bias, and selection or agenda-driven bias. Publication bias (also referred to more commonly as the file-drawer problem) reflects that studies that have found "interesting" effects are more likely to be disseminated. Most often this entails being published in scholarly journal and books, as well as being presented at conferences. As a meta-analysis is synthesis of available studies, an analyst must make conscious efforts to obtain studies that may contain null results. Another common bias is search bias. Even if publication bias were not present, an analyst must be cognizant of the employed search criteria. The inclusiveness of search criteria or exclusionary criteria directly influences the nature of the data set. This not only includes relevant terms and criteria, but, in the present day, the search engine employed to locate studies. A final common bias is that of a selection or agenda-driven bias. A subjective part of the selection process is the decision of which studies to include in one's analysis. While it may or may not be purposeful, a non-equitable selection criteria will most likely bias results. The potential of this bias is often mitigated by including multiple individuals in the decision process for which studies to include.

One method to assess the impact of a non-random sample is to calculate the failsafe N (Corwin, 1983). Simply put, a fail-safe N is a calculation of the number of studies necessary to draw a meaningful effect size to non-meaningful effect size. The reasoning is that if it would take a relatively large number of studies to make the effect size negligible, then the researcher can be confident certain types of non-randomness, such as the often cited file-drawer problem, has had minimal influence such that

$$k_{fs} = \frac{k_{obt}(d_{obt} - d_c)}{d_c - \overline{d}_{fs}}$$

where k_{fs} is the number of fail safe studies, k_{obt} is studies included in the meta-analysis, d_{obt} is summary effect size, d_c is desired lower bound, and $\overline{d}_{fs}$ is the studies with this size needed to lower the effect size. Note, while not necessary, it is common to set $\overline{d}_{fs}$ to zero to indicate no effect. Obviously, such a statistic can not be used as an indicator of other forms of bias, such as selection or agenda-driven bias.

2.1.2 Comparison of Fixed vs. Random Effect Models

While an "effect" is difficult to definitely define, conceptually, it is the underlying influence that is the cause of the difference between two dissimilar or the relationship between two similar entities or concepts. Correspondingly, an effect size refers to the magnitude of the effect. There are two possibilities concerning the conceptualization of a population (i.e. true) effect sizes. Fixed effect models assume that sampling error is the sole source of variation amongst the effect sizes of the included studies and that they share a common effect size. This assumption is plausible when the studies are close replications of one another, use the same procedures, measures, or have other similar characteristics. Thus, the observed effect sizes for the i^{th} study, T_i, will be distributed about the common effect size μ by the random sampling error ε_i, such that

$$T_i = \mu + \varepsilon_i; \varepsilon_i \sim N(0, v_i^2)$$

where v_i^2 is the sampling variance of the effect size. Because it is assumed that the included studies are estimating the same effect size, weights are based on the random error within the studies.

While fixed effect models are based on the assumption that the true effect is the same across the included studies, in application, this is a problematic assumption. While the included studies have enough in common to justify their combination, there is generally no reason to assume that the true effect size is exactly the same across studies. Thus, random effect models assume that the true effect size varies between studies. The studies included in the meta-analysis are assumed to be a random sample of the relevant distribution of effect sizes, and the combined effect size estimates the mean effect size in this distribution. Thus the observed effect size, T_i, differs from the underlying population mean, μ, due to both sampling study variance, ξ_i, and underlying population error, ε_i, such that

$$T_i = \mu + \xi_i + \varepsilon_i; \xi_i \sim N(0, \tau^2), \varepsilon_i \sim N(0, v_i^2)$$

where v_i^2 is the sampling variance of the effect size and τ^2 is the underlying population variance. Thus, there are two levels of sampling and two levels of error. First, each study is used to estimate the true effect in a specific population. Second, all of the true effects are used to estimate the mean of the true effects. It follows that τ^2 will increase as either the variance within-studies decreases and/or the observed variance increases. Therefore, the ability to estimate the combined effect size precisely will depend on both the number of subjects within studies, as well as the total number of studies. While studies based on large samples may yield more precise estimates than studies based on smaller samples, each study is assumed to be estimating a different effect size. In comparison with the fixed effect model, the weights assigned under random effects are more balanced, with studies based on large samples being less likely to dominate the analysis (Hedges and Olkin, 1985).

Hunter and Schmidt (1990) provided sharp criticism to Hedges and Olkin (1985) distinction between fixed and random effect models. Simply put, Hedges and Olkin (1985) assumed that consistency among effect sizes is equivalent to the individual studies being a fixed effect model and that a lack of homogeneity is indicative of an unmeasured random effect component. Hunter and Schmidt (1990) contend that opposed to a fixed effect model, Hedges and Olkin (1985) are in fact referencing a two-factor model in which the random effect factor is nested under the fixed effect factor. To elaborate, one may consider a mean difference effect size. Hedges and Olkin's (1985) implicit assumption is that the difference between conditions (i.e. classes) represents all possible classes. Thus, given this assumption is true, the random effects component (i.e. studies) would then be nested under this fixed effect factor.

2.1.3 Popular Techniques

Two of the more popular meta-analysis techniques were developed by Hedges and Olkin (1985) and Hunter et al. (1982). While both methods have evolved somewhat over time these classic publications are often cited as the guiding methodology behind other's meta-analyses (Hedges, 1983; Hedges and Olkin, 1985; Hedges and Vevea, 1998; Schmidt and Hunter, 1977; Hunter and Schmidt, 1990, Hunter and Schmidt, 2004). Both methods provide an estimate of the overall mean effect size and an estimate of the variability of infinite-sample effect sizes. A sharp distinction is that Hedges and Olkin (1985) first transforms the estimator to an unbiased version prior to integration, while Hunter et al. (1982) employs the biased estimator but makes corrections for various artifacts. It is of note that Hedges and Olkin (1985) equally focus on methods for both mean difference and relational effect sizes. Here, mean difference effect size represents the effect size existing between two disparate constructs or entities, while relational effect size represents the the similarity between two similar constructs or entities. As their technique focuses on characteristics such as unbiased estimators and standardization, one may readily move from effect sizes arising from mean differences to relational effect sizes through conversions. In contrast, Schmidt and Hunter (1990) focus on relational effect sizes. This is likely a reflection of necessity as both Schmidt and Hunter are researchers in the social sciences. As a majority of studies in the social sciences depend on cross-sectional survey designs that prevent inferences of causality. Furthermore, internal validity is difficult to maintain without loss to external validity. In other words, the complexity of human behavior is unlikely to behave naturally in the presence of unnatural situational constructs. Thus for the social sciences, due to these and other factors, the common metric base for a relational effect size between two variables (x and y) is the correlation coefficient. Arguably, the most common correlation measure is the Pearson product moment correlation coefficient;

$$r_{xy} = \frac{\Sigma z_{x_i} z_{y_i}}{n}$$

where n is the number of observational pairings and z_{x_i} and z_{y_i} are the standardized scores of x_i and y_i for case i. As the motivational application concerns the social construct of role theory, the focus of this thesis will concern relational effect sizes.

2.1.4 Hedges and Olkin (1985)

Hedges and Olkin (1985) provide various techniques for analyzing both mean difference and relational effect sizes. The common measure of an individual relational effect size is the Pearson product-moment correlation (r), which is a sample estimator of the population correlation (ρ). It can be shown that while r is the maximum likelihood estimator of ρ it is biased such that

$$Bias(r) \cong -\frac{\rho(1-\rho^2)}{2n}$$

with the sample correlation coefficient tending to underestimate ρ (if $\rho > 0$). It is of note that the sampling variance of r is obtained through an infinite series, but the approximate variance of r (to order $\frac{1}{n}$) is $Var(r) \cong \frac{(1-\rho^2)^2}{n}$. Thus, when the coefficient is large in magnitude (or assessed from a large sample size), the bias of the sample is minimal.

Hedges and Olkin (1985) referenced three methods to correct the bias in r. One potential transformation is to convert r into its unbiased estimate

$$\widetilde{G}_{(r)} \cong r + \frac{r(1-r^2)}{2(n-3)}.$$

$\widetilde{G}_{(r)}$ has been shown to be accurate to within .01 if $n \geq 8$ and to within .001 when $n \geq 18$. Also, $\widetilde{G}_{(r)}$ has been shown to have the same asymptotic distribution as r (Hedges and Olkin, 1985).

The sampling distribution of r is also skewed, especially when the correlation coefficients are large. Therefore, a second potential transformation is to correct for this by converting the correlation to z scores using Fisher's (1921) r-to-z transformation

$$z = \frac{1}{2} \ln \left[\frac{1+r}{1-r} \right]. \tag{2.1}$$

This method is the most commonly employed as it provides multiple benefits. For instance, Fisher's transformation normalizes the variance of r, as well as makes the variance independent of ρ. In addition, simulation studies have shown that linear combinations integrating $\widetilde{G}_{(r)}$ differ only minimally than those integrating z. Furthermore, as will be shown, the sampling variance and corresponding inverse weight simplify to a great extent.

Finally, Hedges and Olkin (1985) briefly referenced a t-transformation proposed by Kramer (1974, 1975) such that $t = \frac{\sqrt{n-2}(r-\rho)}{\sqrt{(1-r^2)(1-\rho^2)}}$ which has an approximately Students t-distribution with $n-2$ degrees of freedom and works well with very small sample sizes. An important note of their work is that these are just three possible conversions to eliminate/minimize biases. Others are possible and may provide superior results depending on scenario.

Effect size

One of the primary considerations when integrating multiple effect sizes into a single estimate concerns the weights placed on individual studies. Hedges and Olkin (1985) approach uses the sample variance $\sigma_r^2 = \frac{(1-\rho^2)^2}{(n-1)}$ in order to establish the fixed effect inverse weights of $w = \frac{(n-1)}{(1-\rho^2)^2}$. Note, as ρ is contained within σ_r^2, larger correlations will receive greater weight and thus may bias the estimate. As previously mentioned two of the main reasons for the popularity of Fisher's r-to-z transformation is that the sample variance converts to

$$\sigma_{z_i}^2 = \frac{1}{n_i - 3}$$

eliminating the bias and the inverse weight simplify to

$$w_i = n_i - 3.$$

for the i^{th} study. Along these lines, employing Fishers r-to-z transformation results in the combined effect size (i.e. weighted average) of

$$\overline{z} = \Sigma_{i=1}^{k} \frac{(n_i - 3)z_i}{\Sigma_{i=1}^{k}(n_i - 3)}.$$

where k is the number of studies included in the meta-analysis.

Significance of effect size

Statistical significance concerns the likelihood of the observed results being obtained from a distribution described in the null hypothesis. The significance may either be described as values such as the standard normal deviates, z, the p-values or in terms of confidence intervals. It can be shown that when $\rho_1 = \rho_2 = \ldots = \rho_k = \rho$ and $n_1, n_2, \ldots, n_k$ increase at approximately the same rate, $\overline{z}$ is approximately normally distributed with mean $\zeta = z(\rho)$ and variance $\frac{1}{\Sigma_{i=1}^{k}(n_i-3)}$. The large sample approximation to the distribution of $\overline{z}$ can be used to test hypotheses concerning ρ by transforming to ζ. For instance, $\rho = \rho_0$ corresponds to the hypotheses $\zeta = \zeta_0 = z(\rho_0)$. This hypothesis would be compared at significance level α using the test statistic

$$z_s ig = (\overline{z} - \zeta_0)\sqrt{\Sigma_{i=1}^{k}(n_i - 3)} \tag{2.2}$$

compared to the $100 \times \alpha$ percent two-tail critical value of the standard normal distribution.

With the weighted average effect size and sample variance it is possible to construct a confidence interval of the population value such that

$$\overline{z} \pm z * \sqrt{\frac{1}{\Sigma_{i=1}^{k}(n_i - 3)}}$$

where $z*$ is the critical value from the standard normal distribution such that the area between $-z*$ and $z*$ is equal to the desired confidence interval. Once the confidence interval has been established for $\overline{z}$ it is possible to revert to the original r through Fisher's corresponding z-to-r transformation

$$r = \frac{e^{2\overline{z}-1}}{e^{2\overline{z}+1}}.$$

Homogeneity of effect size

It is possible to test whether the observed variance is consistent with the hypothesis that there is only a single underlying value of the effect size. The most popular test of the homogeneity of underlying effect sizes is based on work by Cochran (Hedges and Olkin, 1985; Lipsey and Wilson, 2001). Given that $r_1, r_2, ...r_k$ are a series of independent sample correlations and, correspondingly, result in a series of independent z-transformed values, then a test of homogeneity of the population correlation can be constructed from the test statistic

$$Q = \Sigma_{i=1}^{k}(n_i - 3)(z_i - \overline{z})^2, \tag{2.3}$$

where the null hypothesis is rejected if the p-value is less than α, by convention, from a chi-square distribution with $k - 1$ degrees of freedom. Here, Q is a weighted sum of squared deviations from the mean.

Random Error Variance Component

If the Q statistic is found to be significantly different from the null, then there is evidence that effect size is best explained by a random effect model as opposed to a fixed effect model. In other words, the lack of homogeneity in distribution of effect sizes is likely a result in differences between studies. The impact of the random effect component ($\sigma_\rho^2 = REVC$) is expressed as

$$REVC = \frac{Q - (k-1)}{\Sigma w_i - (\frac{\Sigma w_i^2}{\Sigma w_i})}. \tag{2.4}$$

REVC is then used to modify the inverse variance weight such that

$$w_i* = \frac{1}{\frac{1}{w_i} + REVC}. \tag{2.5}$$

Hedges and Olkin (1985) approach is re-conducted utilizing the modified inverse variance weights in order to provide the random effects model estimates.

2.1.5 Hunter et al. (1982)

The method developed by Hunter et al. (1982) does not attempt to correct the biases in effect size prior to integration. Instead the authors' approach attempts to correct the effect size for potential sources of error. It is important to note that the influence of a specific type of artifact is independent across studies. Because of this, it is possible to base meta-analysis on artifact distributions. Hunter and Schmidt (1990) listed some of the prominent artifacts including sampling error, error of measurement in the dependent and independent variables, dichotomization of continuous dependent and independent variables, range variation in the independent variable, attrition artifacts, deviations from perfect construct validity in the dependent and independent variables, reporting or transcriptional error, and variance due to extraneous factors. While corrections for many different artifacts are possible, the information necessary to correct the influence of artifacts are only sporadically reported. Therefore, the artifacts of sampling error, attenuation, and range restriction are the ones most often corrected.

For instance, a reality in the social sciences is that constructs are rarely, if ever, perfectly measured. A commonly cited rule of thumb is that it is acceptable to use scales with reliability as low as .7 (Nunnally, 1978). In this context, internal consistency (or reliability) is the correlation between a measure's scale and true scores. Individual correlation coefficients may be corrected prior to integration for unreliability in x and/or y through the correction

$$r_c = \frac{r_{xy}}{\sqrt{r_{xx}}\sqrt{r_{yy}}},$$

where r_{xx} and r_{yy} are the internal consistencies of x and y, respectively.

Effect Size

After correcting the individual estimates for such artifacts as unreliability, Hunter et al. (1982) estimate ρ as the simple weighted average of sample correlations such that

$$\overline{r} = \frac{\Sigma_{i=1}^{k} n_i r_i}{\Sigma_{i=1}^{k} n_i}$$

with a corresponding variance of

$$\sigma_r^2 = \frac{\Sigma_{i=1}^{k}[n_i(r_i - \overline{r})^2]}{\Sigma_{i=1}^{k} n_i}$$

where n_i is the number of observations and r_i the correlation between x and y for study i.

Significance of Effect Size

For significance testing Hunter et al. (1982) calculates the test statistic as the standard normal *z-score* such that

$$Z = \frac{\overline{r}}{\sigma_{(\overline{r})}} \tag{2.6}$$

where

$$\sigma_{(\bar{r})} = \frac{\Sigma_{i=1}^{k} n_i (r_i - \bar{r})^2}{\Sigma_{i=1}^{k} n_i}.$$

Homogeneity

Hunter et al. (1982) provide a method to test for systematic variation in effect sizes employing a χ^2 test

$$\chi^2_{k-1} = \frac{\Sigma_{i=1}^{k} n_i}{(1 - \bar{r}^2)^2} \sigma_r^2. \tag{2.7}$$

Significant Chi-square values provide evidence that potential moderators may be lurking. As the Chi-square is a summative statistic it is sensitive to changes in sample size and number of studies included in the meta-analysis.

2.1.6 Comparison of Meta-Analytic Strategies

Hunter et al. (1982) differs from Hedges and Olkin (1985) techniques as it does not attempt to correct the biases in effect size indexes before deriving mean effect sizes or before applying moderators to these indexes. Rather, this approach attempts to correct effect size indexes for potential sources of error, such as sampling error, attenuation, and reliability, before meta-analytically integrating the effect size across studies. However, this may be less beneficial as few studies report sources of error.

If the r_i's have a nonlinear bias then Hedges and Olkin (1985) approach, which corrects for biases, should produce a weighted average effect size that differs from Hunter et al. (1982). Specifically, Hunter et al. (1982) should underestimate r. In previous comparisons of the two approaches, the Hunter et al. (1982) approach has generally provided more accurate results than has the Hedges and Olkin (1995) approach (Field, 2001; Hall and Brannick, 2002; Schulze, 2004). Such a result is unexpected because Hedges and Olkin (1985) showed that the maximum likelihood estimator of the mean in the random-effects case depends upon both sampling variance of the individual studies and the variance of infinite-sample effect sizes (random effect variance component or REVC), but the Hunter et al. (1982) procedure uses sample size weights, which do not incorporate the REVC. Thus, the Hunter et al. (1982) weights can be shown to be suboptimal. However, both effect size and the REVC, are subject to sampling error, and thus in practice, they may not provide more accurate estimates. This becomes particularly true if the individual study sample sizes are small (Brannick, Yang, and Cafri, 2008).

2.2 Role Stressors and Consequences

Role stressors are a psychological reaction that arises from ones subjective evaluation of a significant others expectations (Khan et al., 1964). When expectations are unclear, conflicting, and/or overwhelming, both personal and organizationally relevant outcomes may be negatively influenced (Örtqvist and Wincent, 2006). Two of the primary forms of role stressors are role ambiguity and role conflict. With consideration for this motivational example only, role ambiguity will be investigated with it defined as an inadequacy, being either unclear or inconsistent, of a message to convey the necessary information to predict the outcome of the focal person's behavior. Role ambiguity may result from a failure to know what the role expectations are or of actions necessary to conform to those expectations (O'Driscoll, Ilgen, and Hildreth, 1992). Relevant to the motivational application, job satisfaction is considered an attitude to work-related conditions or other aspects of organizational life, including co-workers, pay, and so forth (Wiener, 1982). Örtqvist and Wincent (2006) found that role ambiguity was shown to have a significant negative linear relationship with job satisfaction ($\bar{r} = -.39, k = 39, \Sigma_i^k n_i = 9780$). Although the authors failed to specify their meta-analytic technique it seems likely from the material present that they employed Hunter et al. (1982) technique without corrections.

Chapter 3

Data Analysis

The organization of this chapter is as follows: Section 3.1 introduces and provides the data set under investigation. Section 3.2 presents the results of the various meta-analytic approaches, with discussion into some of the pertinent differences between methods. Section 3.3 introduces the moderating conditions under investigation; these include the strength of the correlation coefficient, sample size, and number of studies composing the meta-analysis. Specific details and discussion concerning the moderating conditions effect on the magnitude of the weighted average effect size, level of significance, homogeneity, and REVC are presented in Sections 3.3.1, 3.3.2, 3.3.3, and 3.3.4, respectively.

3.1 Description of Data

To motivate the investigation of the differences between meta-analytic approaches ten studies ($k = 10$) were selected from the research literature. Specifically, the first ten studies retrieved that successfully reported the correlation between role ambiguity

(x) and job satisfaction (y), their corresponding internal consistencies (r_{xx} and r_{yy}), and sample size (n) were retained for the purpose of this study (Table 3.1).

Table 3.1: Data set

Citation	Year	Sample Size	r_{xx}	r_{yy}	r_{xy}	r_{Corr}
Abdel-Halim (1982)	1982	89	0.670	0.780	-0.570	-0.788
Babin & Boles (1998)	1998	328	0.840	0.920	-0.370	-0.421
Behrman & Perreault, Jr. (1984)	1984	196	0.830	0.960	-0.547	-0.613
Dobreva-Martinova et al. (2002)	2002	1068	0.800	0.920	-0.390	-0.455
Fry & Hellriegel (1987)	1987	187	0.820	0.910	-0.310	-0.359
Glisson & Durick (1988)	1988	319	0.810	0.860	-0.680	-0.815
Johnston et al. (1990)	1990	102	0.807	0.850	-0.355	-0.429
Keenan & McBain (1979)	1979	90	0.840	0.730	-0.480	-0.613
Kemery, Mossholder, & Bedeian (1987)	1987	370	0.810	0.930	-0.360	-0.415
ODriscoll & Beehr (1994)	1994	236	0.860	0.860	-0.600	-0.698

Notes: $\Sigma_i^k = 2985$. r_{xx} is the reliability of role ambiguity. r_{yy} is the reliability of job satisfaction.

3.2 Results

Table 3.1 was analyzed employing two meta-analytic approaches. One approach was that endorsed by Hedges and Olkin (1985) which provides estimates for both the fixed effect (HO FEM) and random effect (HO REM) models (Appendix A). The second approach was that endorsed by Hunter et al. (1982). This approach was conducted without correcting for artifacts (Hunter and Schmidt (1990) refer to this as a bare-bones analysis (HSJ BB)) along with a variant correcting for the artifact of attenuation (HSJ Corr)(Appendix B). Thus, four different meta-analyses were conducted for comparative purposes being that of HO FEM, HO REM, HS JBB, and HJS Corr (Table 3.2).

Table 3.2: Meta-analytic estimates by approach

Approach	$\overline{z}$	$\overline{r}$	s.e.	LCB	UCB	$\lvert Z_{sig}\rvert$	Q	χ^2	REVC
HO FEM	-0.488	-0.453	0.018	-0.481	-0.423	26.528	72.387		
HO REM	-0.516	-0.474	0.056	-0.556	-0.385	9.161			0.026
HSJ BB		-0.444	0.036	-0.515	-0.373	34.142		59.953	
HSJ Corr		-0.522	0.045	-0.611	-0.433	59.953		115.900	

Note: LCB = 95% lower confidence bound. UCB = 95% upper confidence bound.

Several insights may be taken from Table 3.2. Results showed only minimal differences between HO FEM ($\overline{z} = -.488$) and HSJ BB ($\overline{r} = -.444$). This is as expected as by Hunter and Schmidt (1990) admission their technique parallels Hedges and Olkin (1985) fixed effect model. However, within the individual approaches large differences occur. For instance, results showed a substantial difference in the weighted average effect size between HO FEM ($\overline{z} = -.488$) and HO REM ($\overline{z} = -.516$). This is a result of the incorporation of the REVC into the revised weight in (2.5). In addition, results demonstrate a substantial difference in the weighted average effect size between HSJ BB ($\overline{r} = -.444$) and HSJ Corr ($\overline{r} = -.528$). This is a direct result of the correction of $\frac{r_{xy}}{\sqrt{r_{xx}}\sqrt{r_{yy}}}$ imposed on HSJ Corr. In cases of perfect reliability in which both r_{xx} and r_{yy} equal 1, $\frac{r_{xy}}{\sqrt{r_{xx}}\sqrt{r_{yy}}}$ equals to r_{xy}. However, if unreliability is present in x and/or y then correcting an unreliability artifact results in an increase to the magnitude of the weighted average effect size.

There was also large variations in the standard error between models, and correspondingly their confidence intervals and magnitude of the $z - value$ (i.e.$|z_{sig}|$). A major disparity was the difference between HO FEM ($s.e.$ = .018) and HO REM ($s.e.$ = .056) which is a direct result of the REVC being incorporated into the weight of the HO REM standard error (2.5). Another disparity arises between HSJ BB

($s.e.$ = .036) and HSJ Corr ($s.e.$ = .045). This is a result of utilizing the corrected weighted average effect size, opposed to the simple weighted average effect size in 2.1.5. Thus, the greater artifact correction imposed the higher the standard error.

Homogeneity is assessed through χ^2, which in general is measuring the weighted sum of squares of the effect size about the weighted mean effect size. Table 3.2 showed that there are large difference in the χ^2 values between HO FEM ($\chi^2 = 72.387$), HSJ BB ($\chi^2 = 59.953$), and HSJ Corr ($\chi^2 = 115.900$). These differences are partially a result of the variances used to impose the weight on the χ^2. Specifically, while HO FEM employs the reciprocal of the inverse weight of $(n_i - 3)$ in (2.3) the Hunter et al. (1982) approach uses $\frac{1}{(1-\bar{r}^2)^2}$ in (2.7).

3.3 Potential Moderators

The characteristics of Table 3.1 were manipulated to investigate the effect of the magnitude of correlation coefficient (r), size of the sample (n), and the number of studies included in the meta-analysis (k) on various characteristics of the meta-analyses. One manipulated condition was the effect of the magnitude of the correlation prior to integration. To investigate this influence each meta-analytic technique was conducted using $r + .15$, r, and $r - .15$. Note that $\pm.15$ was selected only with considerations of not violating the natural range of correlation coefficients or reversing the direction of the correlation coefficient. Furthermore, the correlation between role ambiguity and job satisfaction is negative, thus $r - .15$ results in a stronger and $r + .15$ a weaker strength correlation. A second manipulated condition was to investigate the effect of sample size on meta-analytic estimates. To do so, the sample size of each study was either doubled ($n \times 2$) or tripled ($n \times 3$) depending on condition. The final manipulated condition was to investigate the effect of the number of studies on meta-analytic estimates. This was done by doubling ($k \times 2$) or tripling ($k \times 3$) the data set (Table 3.1) depending on condition. One caveat is that to examine the influence of the number of studies included in the meta-analysis on REVC, k was manipulated in (2.4) by multiplying the value by 1, 2, or 3 depending on condition. Results for HO FEM, HO REM, HSJ BB, and HSJ Corr are located in Tables 3.3, 3.4, 3.5, 3.6, respectively.

Table 3.3: Hedges and Olkin (1985) Fixed Effect Model

Database	$\bar{r}$	s.e.	LCB	UCB	$\lvert Z_{sig} \rvert$	Q
Table 3.1	-0.453	0.018	-0.481	-0.426	26.528	72.387
r - .15	-0.610	0.018	-0.632	-0.587	38.535	132.590
r + .15	-0.299	0.018	-0.331	-0.266	16.753	51.454
n x 2	-0.453	0.013	-0.473	-0.432	37.623	145.587
n x 3	-0.453	0.011	-0.469	-0.436	46.122	218.787
k x 2	-0.453	0.013	-0.473	-0.432	37.516	144.775
k x 3	-0.453	0.011	-0.469	-0.436	45.948	217.162

Table 3.4: Hedges and Olkin (1985) Random Effect Model

Database	$\overline{r}$	s.e.	LCB	UCB	$\|Z_{sig}\|$	REVC
Table 3.1	-0.474	0.056	-0.556	-0.385	9.161	0.026
r - .15	-0.632	0.076	-0.713	-0.536	9.887	0.051
r + .15	-0.320	0.048	-0.402	-0.234	6.932	0.018
n x 2	-0.475	0.030	-0.520	-0.428	17.184	0.028
n x 3	-0.475	0.030	-0.520	-0.429	17.294	0.029
k x 2	-0.474	0.0380	-0.530	-0.415	13.596	0.023
k x 3	-0.474	0.031	-0.520	-0.427	16.897	0.023

Table 3.5: Hunter, Schmidt, and Johnson (1982) Bare-bones Approach

Database	$\overline{r}$	s.e.	LCB	UCB	$\|Z_{sig}\|$	χ^2
Table 3.1	-0.444	0.036	-0.515	-0.373	34.142	59.953
$r - .15$	-0.594	0.036	-0.665	-0.523	45.680	92.252
$r + .15$	-0.294	0.036	-0.365	-0.223	22.604	46.310
$n \times 2$	-0.444	0.036	-0.515	-0.373	34.142	120.137
$n \times 3$	-0.444	0.036	-0.515	-0.373	34.142	180.321
$k \times 2$	-0.444	0.025	-0.494	-0.394	34.142	119.905
$k \times 3$	-0.444	0.021	-0.485	-0.403	34.142	179.858

Table 3.6: Hunter, Schmidt, and Johnson (1982) Corrected for Attenuation

Database	$\overline{r}$	s.e.	LCB	UCB	|Zsig|	χ^2
Table 3.1	-0.522	0.045	-0.611	-0.433	25.280	115.895
$r - .15$	-0.672	0.045	-0.761	-0.583	32.545	203.910
$r + .15$	-0.372	0.045	-0.461	-0.283	18.014	82.637
$n \times 2$	-0.522	0.045	-0.611	-0.433	25.280	232.292
$n \times 3$	-0.522	0.045	-0.611	-0.433	25.280	348.689
$k \times 2$	-0.522	0.032	-0.585	-0.459	25.280	231.790
$k \times 3$	-0.522	0.026	-0.573	-0.470	25.280	347.686

3.3.1 Effect Size

Figure 3.1 illustrates the effect of the strength of individual correlation coefficients on the weighted average effect size. Conditions were investigated by artificially manipulating the strength of the coefficients by $\pm.15$ for each meta-analysis. As expected, the weighted average effect size increased or decreased based on the condition (-.15 and +.15, respectively).

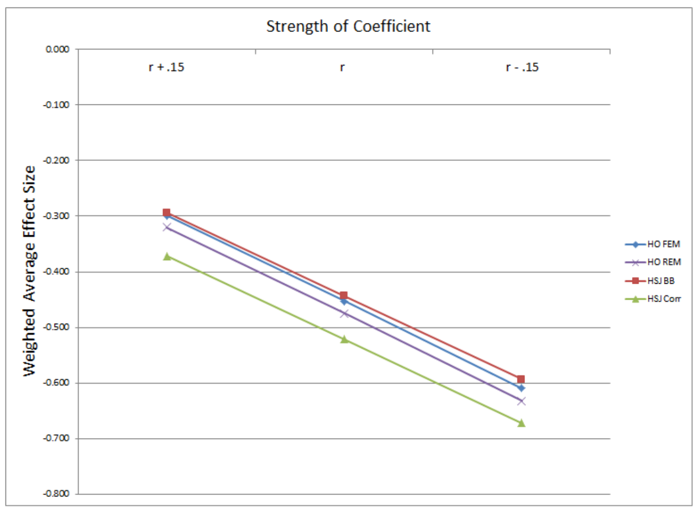

Figure 3.1: The effect of the magnitude of r on the weighted average effect size

While sample size effects weights of individual effect sizes prior to integration a simple proportional increase across all studies does not. Figure 3.2 demonstrates that the weighted average effect size is invariant to proportion changes in sample size. Similarly, Figure 3.3 demonstrates that if all else remains constant, a simple proportional increase in the number of studies does not influence the weighted average effect size.

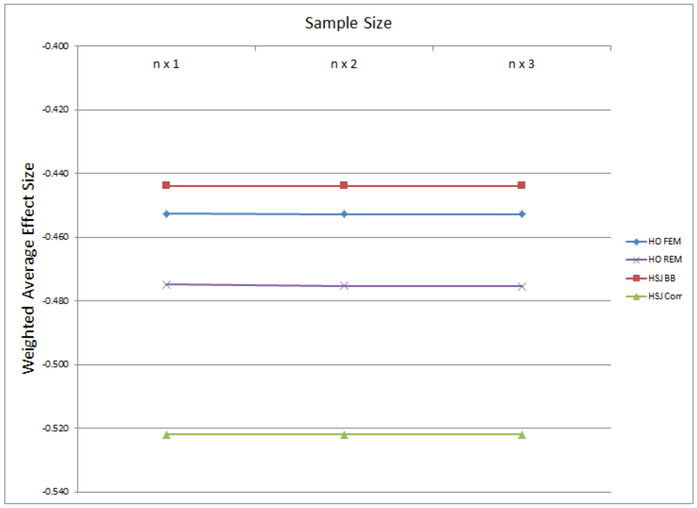

Figure 3.2: The effect of sample size on the weighted average effect size

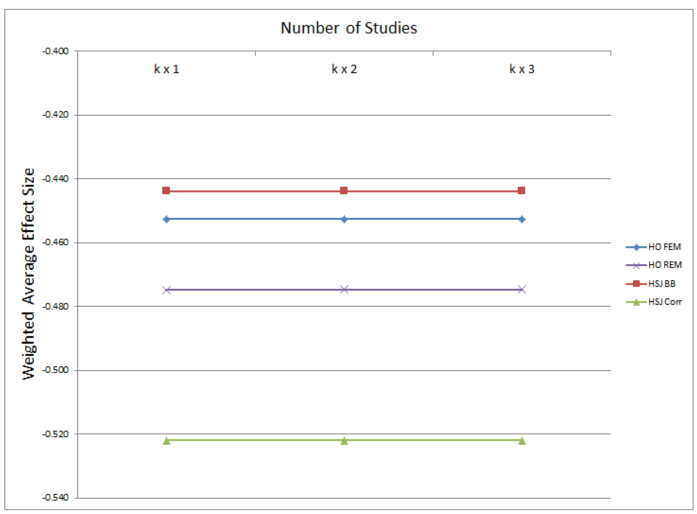

Figure 3.3: The effect of the number of studies on the weighted average effect size

3.3.2 Significance of Effect Size

Figure 3.4 shows that as the strength of the correlation coefficients increase so does the magnitude of the standard normal z-score. An increase in the magnitude of weighted average effect size results in high magnitude z-score as the weighted average effect size is placed in the Z_{sig} numerator for both Hedges and Olkin (1985) in (2.2) and Hunter et al. (1982) in (2.6) approaches. However, HO REM increases at a slower rate due to the incorporation of the REVC into random effects inverse weight in (2.4).

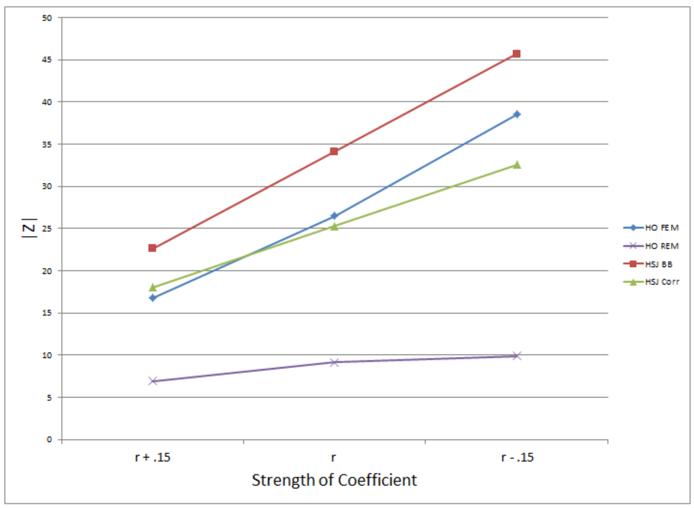

Figure 3.4: The effect of the magnitude of r on $|Z_{sig}|$

Figures 3.5 and 3.6 show that the Hunter et al. (1982) approach is invariant under proportional increases in sample size and number of studies in the meta-analysis. In contrast, the Hedges and Olkin (1985) approach is influenced, with HO FEM to a greater extent than HO REM. This is a result of differences in the standard error terms in the fixed effect (2.3) and random effect (2.5) models.

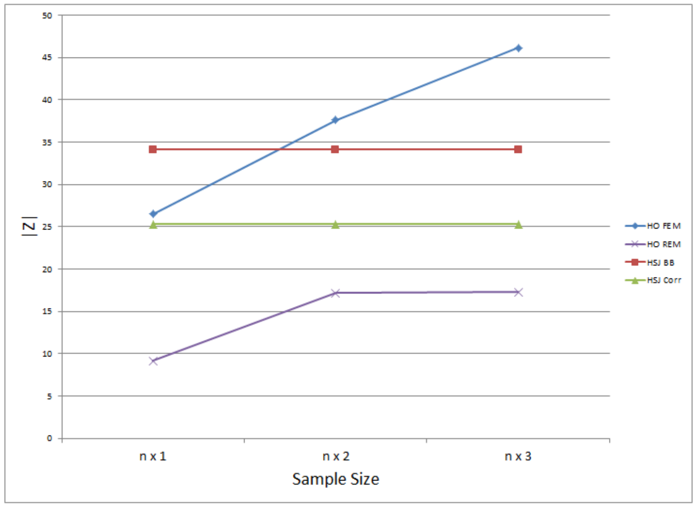

Figure 3.5: The effect of sample size on $|Z_{sig}|$

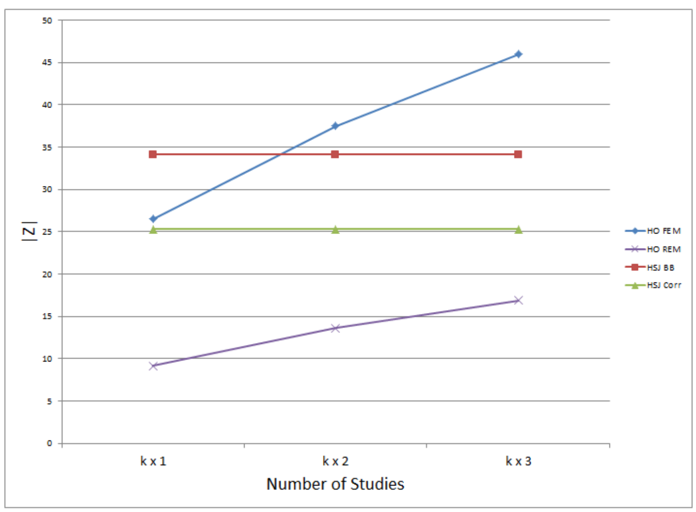

Figure 3.6: The effect of the number of studies on $|Z_{sig}|$

3.3.3 Homogeneity of Effect Size

Hedges and Olkin's (1985) Q statistic is a χ^2 test that assesses for consistency of effect sizes. When the χ^2 is significant, a random effect model is statistically justified. While not inherently clear in the original formula (2.3), the computational formula

$$Q = \Sigma(w_i z_i^2) - \frac{(\Sigma w_i z_i)^2}{\Sigma w_i} \tag{3.1}$$

conveys that the $\Sigma(w_i z_i^2)$ term increases at at faster rate than the $(\Sigma w_i z_i)^2$ term for stronger correlation coefficients given a constant weight. Thus an increase in the strength of the correlation coefficient results in larger χ^2 values.

Correspondingly, Hunter et al. (1982) provide a method to test for systematic variation in effect sizes employing a χ^2 test in (2.7). Figure 3.7 illustrates that homogeneity decreases as the strength of the correlation coefficients increases. This is a direct result of the $(1 - r^2)^2$ term in (2.7). Thus, an increase in the magnitude of the weighted average effect size results in larger χ^2 values.

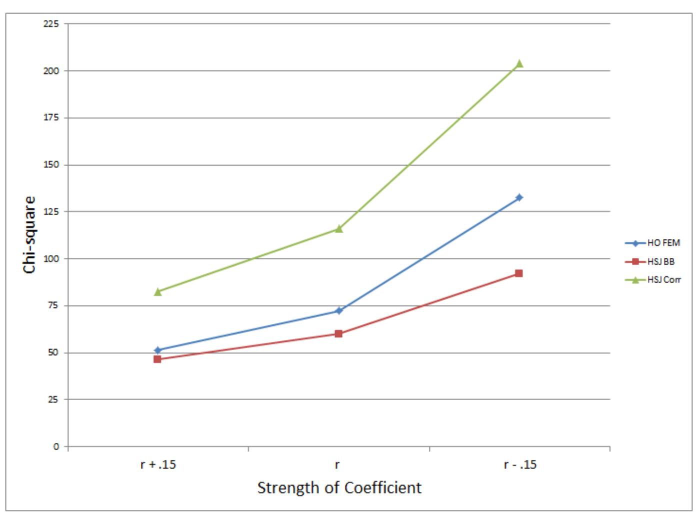

Figure 3.7: The effect of the magnitude of r on χ^2 value

Figures 3.8 and 3.9 both emphasize that the χ^2 is a summative statistic. Thus, increases in sample sizes or the number of studies decreases homogeneity of the measure.

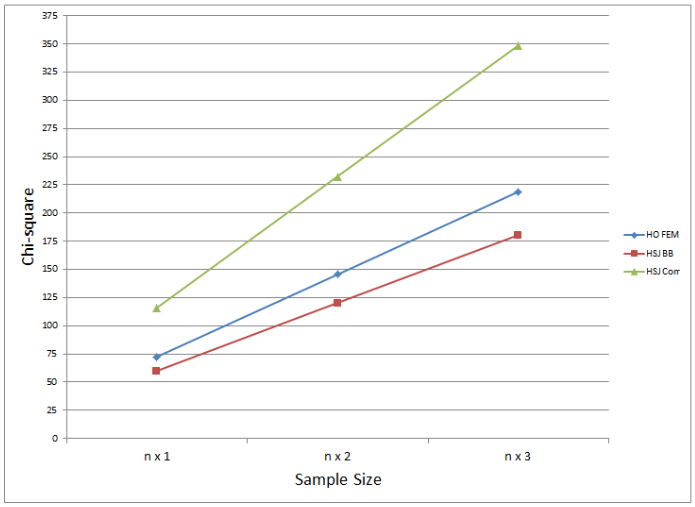

Figure 3.8: The effect of sample size on χ^2 value

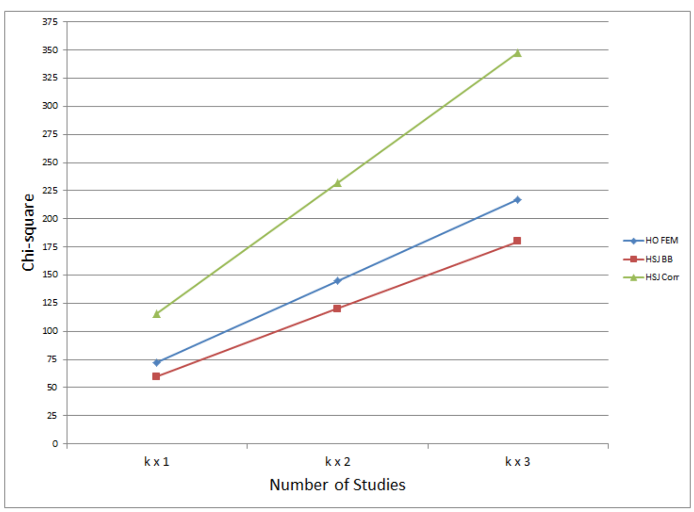

Figure 3.9: The effect of the number of studies on χ^2 value

3.3.4 Random Error Variance Component

Table 3.4 shows that the REVC is sensitive to manipulations to the strength of the correlation coefficient. An increase in the strength of the coefficients was shown to result in a substantial increase in the REVC (value for $r+.15$ was .018, r was .026, and $r-.15$ was .051). This result holds true for Fisher's r-to-z transformation as higher magnitude values are increased at a greater rate under a standard normal distribution than r. However, this finding may not hold for other potential transformations such as those proposed by Hedges and Olkin (1985). Proportional increases in sample size were shown to only negligibly increase the REVC (value for $n \times 1$ was .026, $n \times 2$ was .028, and $n \times 3$ was .029). In contrast, results showed a negative relationship between the number of studies included in the meta-analysis and REVC (value for $k \times 1$ was .026, $k \times 2$ was .022, $k \times 3$ was .018). This result conforms to expectations given in (2.4).

Chapter 4

Discussion

The purpose of this study was to investigate and compare two popular meta-analytic approaches. In general, the results of this study provide two insights to practitioners. First is that the selection of meta-analytic technique does have a direct impact on one's findings. For instance, the differences between HO FEM, HO REM, HSJ BB, and HSJ corr were -.453, -.475, -.444, and -.522, respectively. Unfortunately the second realization is that there is no "correct" methodology. It is up to the analyst to investigate the nature of the constructs and measures under consideration and make a subjective judgement which technique is best suited to the circumstance. For instance, if one is dealing with scales believed to be unduly influenced by artifacts, then a Hunter et al. (1982) approach imposing corrections may be the most appropriate. However, if information concerning those artifacts is unavailable, then another technique may be more appropriate. Even within a single approach, such as Hedges and Olkin (1985), there are multiple transformations possible to correct for biasness depending on the characteristics on the data set. Thus, both across and within different meta-analytic techniques multiple decisions must be made.

A primary goal of this study was to investigate how manipulating the characteristics of a meta-analysis data set effects estimates on a "real" data set. While the realism of the data set adds value, it comes at a cost, as the data set was artificially manipulated in order to investigate topics of interest. For instance, to investigate the effects on an increased data set (i.e. $k \times 1$, $k \times 2$, $k \times 3$) the initial data set was simply doubled or tripled. While this did increase k technically the homogeneity of the effects becomes convoluted due to violations concerning the independence of observations. Yet, other methods have their own drawbacks with no single approach free from their own inherent weaknesses.

Chapter 5

Bibliography

[1] *Abdel-Halim, A. A. (1982). Social support and managerial affective responses to job stress. *Journal of Occupational Behavior, 3,* 281-295.

[2] *Babin, B. J., & Boles, J. S. (1998). Employee behavior in a service environment: A model and test of potential differences between men and women. *Journal of Marketing, 62,* 77-91.

[3] * Behrman, D. N., & Perreault, W. D. Jr. (1984). A role stress model of the performance and satisfaction of industrial salespersons. *Journal of Marketing,* 48, 9 -21.

[4] Brannick, M. T., Yang, L., & Cafri, G. (2008). Comparison of weights in meta-analysis under realistic conditions. Poster presented at the 23rd annual conference of the Society for Industrial and Organizational Psychology, San Francisco, CA, April 2008.

[5] Brannick, M. T. (June, 2013). *Hedges SAS* Retrieved from http://luna.cas.usf.edu/ mbrannic/files/meta/HedgesP88.sas

[6] Brannick, M. T. (June, 2013). *Schmidt and Hunter bare bones approach.* Retrieved from http://luna.cas.usf.edu/ mbrannic/files/meta/shbarbonesr.sas

[7] Corwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics, 8,* 157-159.

[8] * Dobreva-Martinova, T., Villeneuve, M., Strickland, L., & Matheson, K. (2002). Occupational role stress in the Canadian forces: Its association with individual and organizational well-being. *Canadian Journal of Behavioural Science, 34,* 111-121.

[9] Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed-and random-effects methods. *Psychological Methods, 6,* 161-180.

[10] Fisher, R.A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron, 1,* 332.

[11] * Fry, L. W., & Hellriegel, D. (1987). The role and expectancy participation model: An empirical assessment and extension. *Journal of Occupational Behavior, 8,* 295-309.

[12] Glass, G. V. (1976). Primary, secondary, and meta-analysis. *Educational Researcher, 5,* 3-8.

[13] * Glisson, C., & Durick, M. (1988). Predictors of job satisfaction and organizational commitment in human service organizations. *Administrative Science Quarterly,* 33, 61- 81.

[14] Hall, S. M., & Brannick, M. T. (2002). Comparison of two random-effects methods of meta-analysis. *Journal of Applied Psychology, 87,* 377-389.

[15] Hedges, L. V. (1983). A random effects model for effect sizes. *Psychological Bulletin, 93,* 388-395.

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

[17] Hedges, L. V. & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486-504.

[18] Hunter, J.E., Schmidt, F.L., & Jackson, G.B. (1982). Meta-analysis: Cumulating findings across research. Beverly Hills, CA: Sage.

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

[20] Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings (2nd edition).* Newbury Park, CA: Sage. Hillsdale, NJ

[21] Johnson, B.T, Mullen, B., & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology, 80(1),* 94-106.

[22] * Johnston, M. W., Parasuraman, A., Futrell, C. M., & Black, W. C. (1990). A longitudinal assessment of the impact of selected organizational influences on salespeople's organizational commitment during early employment. *Journal of Marketing Research, 27,* 333-344.

[23] Kahn, R. L., Wolfe, D., Quinn, A., Snoek, J. D., & Rosenthal, R. (1964). *Organizational Stress: Studies in Role conflict and Role Ambiguity.* New York: Wiley and Sons.

[24] * Keenan, A., & McBain, G. D. M. (1979). Effects of type A behavior, intolerance of ambiguity, and locus of control on the relationship between role stress and work-related outcomes. *Journal of Occupational Psychology, 52,* 277-285.

[25] * Kemery, E. R., Mossholder, K. W., & Bedeian, A. G. (1987). Role stress, physical symptom- atology, and turnover intentions: A causal analysis of three alternative specifications. *Journal of Occupational Behavior, 8,* 11-23.

[26] Kramer, H. C. (1974). Improved approximation to the non-null distribution of the correlation coefficient. *Journal of American Statistical Association, 68,* 1004-1008.

[27] Kramer, H. C. (1975). On estimation and hypothesis testing problems for correlation coefficients. *Psychometrika, 40,* 473-485.

[28] Nunnally, J. C. (1978). *Psychometric theory (2nd ed.).* New York: McGraw-Hill.

[29] * O'Driscoll, M. P., & Beehr, T. A. (1994). Supervisor behaviors, role stressors and uncertainty as predictors of personal outcomes for subordinates. *Journal of Organizational Behavior, 15,* 141-155.

[30] Örtqvist, D. & Wincent, J. (2006). Prominent consequences of role stress: A meta-analytic review. *International Journal of Stress Management, 13(4),* 399-422.

[31] Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62,* 529-540.

[32] Schulze, R. (2004). *Meta-analysis: A comparison of approaches.* Gottingen, Germany: Hogrefe & Huber.

[33] Wiener, Y. (1982). Commitment in organizations. A normative view. *Academy of Management Review, 7(3)*, 418-428.

Appendix A

Hedges and Olkin (1985) SAS Macro (Brannick, 2013)

```
%MACRO Hedges(es,w,dsn=_last_,print=raw) ;
*****************************************************************
* When using this macro, the default is the
* effect size d (or g) which are not transformed
* for the analysis.  The results are printed by default
* in whatever metric you input.Therefore, if you input
* d, you are good to go. But if you input *log transformed*
* data and you want results in the original metric, put
* Print = EXP in the call statement, e.g., %Hedges(es,w, print=EXP).
* If you input *Fisher z transformed* data and you want
* original metric output, put Print=BKR to get it back to r.
*****************************************************************
*Note that the variance and other statistics
*except the means and confidence intervals
*reported by the program will still be in the transformed metric,
*even though the confidence intervals will be in the original metric.
```

```
*You cannot convert tau or tau squared into the original metric, only
*the interval resulting from the application of tau.
*********************************************************************;
proc iml;
use &dsn ;
read all var{&es} into es   where(&es^=. & &w^=.) ;
read all var{&w}  into w    where(&es^=. & &w^=.) ;
k = nrow(es)  ;
df = k - 1   ;
***********************************
*  Fixed effects
***********************************;
*mean effect size;
mes = sum(es#w)/sum(w);
*standard error of the mean Fixed;
sem = sqrt(1/sum(w));
*lower bound for the mean   Fixed;
les = mes - 1.95996*sem;
*upper bound for the mean   Fixed;
ues = mes + 1.95996*sem;
* z for mean                Fixed;
z   = mes/sem ;
* p value for mean          Fixed;
pz  = (1 - (.5+erf(abs(z)/sqrt(2))/2))*2 ;
* maximum value of the effect sizes;
maxes = max(es) ;
* minimum value of the effect sizes;
mines = min(es) ;
*********************************
* random effects
```

```
*********************************;
*weighted sum of squares, Q;
q = sum((es#es)#w) - sum(es#w)#sum(es#w)/sum(w);
wsd   = sqrt(q*w[+,]**-1) ;
* p value for Q;
pq   = 1-probchi(q,df) ;
*T-squared (estimated tau-squared;
Tsquare    = (q - df)/(w[+,]-sum(w#w)/w[+,]) ;
* set to zero if less than zero;
if Tsquare<0 then ; do ; Tsquare = 0 ; end ;
* compute random effects weights;
wre    = 1/(1/w + Tsquare) ;
* Mean ES Random;
mesre = sum(es#wre)/sum(wre) ;
* Standard error of the mean Random;
semre = sqrt(1/sum(wre)) ;
* Variance of the mean Random;
VmRE = semre#semre;
* Lower bound for the mean Random;
lesre = mesre -    1.95996*semre ;
* Upper bound for the mean Random;
uesre = mesre +    1.95996*semre ;
* z test for the mean Random;
zre   = mesre/semre ;
*p value for the test of the H0: mean is zero, Random;
pzre  = (1 - (.5+erf(abs(zre)/sqrt(2))/2))*2 ;
* I squared;
Isquare = ((q-df)/q)#100;
* T or estimated tau;
tau = sqrt(Tsquare);
```

```
* sums of weights;
sw1 = sum(w);
sw2 = sum(w#w);
sw3 = sum(w#w#w);
* scaling factor;
C = sw1-(sw2/sw1);
A1 = df+2#(sw1-sw2/sw1)#Tsquare;
A2  = (sw2-2#(sw3/sw1)+(sw2#sw2)/(sw1#sw1))#(Tsquare#Tsquare);
* for computing CI for REVC;
A = A1+A2;
B = .5#(log(q)-log(df))/(sqrt(2#q)-sqrt(2#df-1));
b1=1-(1/(3#(df-1)#(df-1))); b2 = 2#(df-1)#b1; b3 = sqrt(1/b2);
* for computing CI for REVC;
if q <= (df+1) then B = b3;
L = exp(.5#log(q/df)-1.96#B);
U = exp(.5#log(q/df)+1.96#B);
* lower bound for tau squared;
LLtsq = (df#(L#L-1))/C;
if LLtsq < 0 then LLtsq =0;
* upper bound for tau squared;
ULtsq = (df#(U#U-1))/C;
if ULtsq < 0 then Ultsq =0;
* lower bound for tau;
LLtau = sqrt(LLtsq);
* upper bound for tau;
ULtau = sqrt(ULtsq);
*******************************
* prediction interval
*******************************;
dist = tinv(.975,(df-1));
```

```
LLpred = mesre - dist#sqrt(Tsquare+VmRE);
ULpred = mesre + dist#sqrt(Tsquare+VmRe);
%if %upcase(&print) = EXP %then
%do;
mes = exp(mes) ;
les = exp(les) ;
ues = exp(ues) ;
sem = . ;
mes_re = exp(mes_re) ;
les_re = exp(les_re) ;
ues_re = exp(ues_re) ;
semre = . ;
print 'means and CIs ONLY converted from log to original';
%end;
%if %upcase(&print) = BKR %then
%do;
mes = (exp(2#mes)-1)/(exp(2#mes)+1);
les = (exp(2#les)-1)/(exp(2#les)+1) ;
ues = (exp(2#ues)-1)/(exp(2#ues)+1) ;
mesre = (exp(2#mesre)-1)/(exp(2#mesre)+1);
lesre = (exp(2#lesre)-1)/(exp(2#lesre)+1) ;
uesre = (exp(2#uesre)-1)/(exp(2#uesre)+1) ;
LLpred = (exp(2#LLpred)-1)/(exp(2#LLpred)+1) ;
ULpred = (exp(2#ULpred)-1)/(exp(2#ULpred)+1) ;
print 'means and CIs ONLY converted from z to r';
%end;
print '-------------- Distribution Description --------------';
mattrib         k     label={"No. of obs."}     maxes label={"Max Obs."}
mines label={"Min Obs."} wsd   label={"Weighted SD"};
print k mines maxes wsd [format=12.5];
```

```
print " ";
print '-------------- Homogeneity Analysis --------------';
mattrib df label={"df"} pq label={"p"} ;
print q [format=12.5] df pq [format=12.5];
print Isquare [rowname="Proportion of observed variance due to random effects = "];
print Tsquare [rowname="Random effects var. component = "];
print 'Lower and Upper bounds for tausquared';
print LLtsq ULtsq;
print tau [rowname="Standard deviation of random effects = "];
print'Lower and Upper bounds for tau';
print LLtau ULtau;
print 'Lower and Upper Bounds for 95 percent Credibility (Prediction) Interval';
print LLpred ULpred;
print " ";
print '-------------- Fixed & Random Effects Model --------------';
fixed = mes || sem || les || ues || z || pz ;
random = mesre || semre || lesre || uesre || zre || pzre ;
model = fixed // random ;
mattrib model rowname=({'Fixed', 'Random'})
colname=({'Mean' 'SE' '-95%CI' '+95%CI' 'z' 'p'})
label={"Model"};
print model [format=10.5];
quit;
%MEND Hedges;
run;
%Hedges(es,w,print=BKR);
run;
```

Appendix B

Hunter, Schmidt, and Jackson (1982) SAS Syntax (Brannick, 2013)

```
data d1;
input r n;
cards;
.25 100
.30 250
.20 200
.40 150
proc print;
proc iml;
*Schmidt and Hunter Bare Bones;
**************************************************;
use d1;
read all into x;
```

```
obsr = x[,1];          *observed correlations;
n = x[,2];             *sample size N;
k = nrow(X);
sumn=n[+];             *sum of N;
aven = sumn/k;
*print x;
nr= obsr`*n;           *sum weighted r;
aver=nr/sumn;          *weighted mean;
varr1= obsr - aver;    *deviation from weighted mean;
varr2=n`* varr1##2;    *sum weighted squared deviations;
varr=varr2/sumn;       *weighted variance of obs r (s-squared sub r);
samperr = (1-aver**2)**2/((sumn/k)-1); *sampling error variance;
resr=varr-samperr;     *residual variance (variance of rho);
if resr < 0 then resr = 0;  *keep boundary on residual variance;
sdrho=resr**.5;         *print sdrho;
CI95L = aver-1.96#sqrt(varr/k);
CI95U = aver+1.96#sqrt(varr/k);
CR95L = aver-1.96#sqrt(resr);
CR95U = aver+1.96#sqrt(resr);
********************************************;
Print 'Number of studies is' k;
Print 'Average sample size is' aven;
Print 'Estimated population mean is' aver;
Print 'Observed Variance is' varr;
Print 'Sampling Error Variance is' samperr;
Print 'SDrho is' sdrho;
Print '95 percent confidence interval for mean is' CI95L CI95U;
Print '95 percent credibility interval is' CR95L CR95U;
quit;
run;
```