Running Head: Meta-analysis

Application of Meta-analysis in Sport and Exercise Science

Order of Authorship and Contact Information

Soyeon Ahn
School of Education and Human Development
University of Miami
1507 Levante Avenue
Max Orovitz Building, Rm. 310B
Coral Gables, FL, 33124-2040, USA
Phone: 1.305.284.5389
E-mail: s.ahn@miami.edu

Min Lu
School of Education and Human Development
University of Miami
1507 Levante Avenue, Max Orovitz Building, Rm. 307.02
Coral Gables, FL, 33146, USA
Phone: 1.305.284.3001
E-mail: m.lu6@umiami.edu

G. Tyler Lefevor
School of Education and Human Development
University of Miami
1507 Levante Avenue, Max Orovitz Building, Rm. 307.02
Coral Gables, FL, 33146, USA
Phone: 1.305.284.3001
E-mail: g.lefevor@umiami.edu

Alicia L. Fedewa
Department of Educational, School, and Counseling Psychology
University of Kentucky
238 Dickey Hall (Mail)
Lexington, KY 40506-0017
Phone: 1.859.257.9338
E-mail: alicia.fedewa@uky.edu

Seniz Celimli
School of Education and Human Development
University of Miami
1507 Levante Avenue, Max Orovitz Building, Rm.
Coral Gables, FL, 33146, USA
Phone: 1.305.284.3001
E-mail: s.celimli@miami.edu

Running Head: Meta-analysis

Application of Meta-analysis in Sport and Exercise Science

Abstract

Meta-analysis is a tool that provides a cohesive picture of a phenomenon of interest by statistically integrating study findings on the “same” or a “similar” phenomenon across a large collection of studies. Due to a number of strengths, since being introduced, meta-analysis has been widely used in the area of Sport and Exercise Science, and we believe that will continue to be the case. The current chapter aims to (1) provide a brief overview of meta-analysis and discuss how to choose the most appropriate meta-analytic method and then (2) offer a step-by-step tutorial in R that illustrates how to perform both univariate and multivariate meta-analysis using the dataset to answer essentially the same research questions. This chapter is designed to introduce researchers to meta-analysis as a data analytic approach and familiarize them with the best-practices within meta-analysis to address research questions in the area of Sport and Exercise Science.

Keywords: Meta-analysis, Univariate Meta-analysis, Multivariate Meta-analysis, R

General Introduction

Meta-analysis,[1] or "analysis of analyses," was coined by Gene V. Glass (1976) in his presidential talk at the annual meeting of the American Educational Research Association. Meta-analysis[1] is a tool that provides a cohesive picture of a phenomenon of interest by statistically integrating study findings on the "*same*" or "*similar*" phenomenon across a large collection of studies. It boasts a number of strengths, such as an increase in statistical power by including more samples from selected studies, which have made meta-analysis popular in many disciplines, including sport and exercise science (Becker & Ahn, 2012).

Stages of Meta-analysis

The general procedure for conducting a meta-analysis is almost identical to that for primary research. Cooper (1984) conceptualized meta-analysis in five stages, which he further expanded to seven stages in 2010 (Cooper, 2010). Stage 1, *problem formulation*, involves formulating research questions and hypotheses. Stage 2, *data collection,* involves searching the literature and coding information from relevant studies. Stage 3, *data evaluation,* involves evaluating the quality of study and dealing with differences in the quality of studies. Stage 4, *data analysis*, involves analyzing and integrating effect sizes across studies. Stage 5, *presentation,* involves interpreting the cumulative evidence obtained through the analysis.

In each of the five stages, Cooper outlined questions to be addressed that would "enhance or undermine the trustworthiness of the conclusion" (2010, p. 11). These questions allow a researcher to better ensure the validity of the subsequent meta-analysis. Though important, space prevents a detailed discussion of each stage in the present chapter. The interested reader is referred to several excellent sources that give a more detailed description of questions for consideration in each stage of meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2008;

Cooper, 1984, 2010; Hedges & Olkin, 1985; Wanous, Sullivan, & Malinak, 1989). In this chapter, our focus is primarily on stage 4 as our main goal is to illustrate how to apply meta-analytic methods in sport and exercise science.

Key Elements of Meta-analysis

Figure 11.1 is a forest plot for a meta-analysis of 18 studies on the treatment effect of physical activity in reducing children's depression, which were included in a published meta-analysis conducted by Ahn and Fedewa (2011). The forest plot in Figure 11.1 presents four key elements for conducting a meta-analysis: (1) *individual studies*, (2) the *effect size* of each study, (3) the *variance* (or *standard error*) of each effect size as a measure of *precision*, and (4) *study weight* (Borenstein et al., 2009).

Insert Figure 11.1 about here

The first column of this plot shows the study ID that is assigned to each of the eighteen ***individual studies*** that examine treatment effect with physical activity by comparing the level of children's depression between a treatment and control group. These 18 studies examine the "*same*" or a "*similar*" phenomenon – the treatment effect of physical activity on children's depression. Therefore, the findings from those studies are comparable and can be synthesized in a meta-analysis.

From each of the individual studies included in a meta-analysis, a standardized mean difference (*d*) between the treatment and control groups on the level of children's depression is computed and serves as the ***effect size*** of each study. The effect sizes are shown in the second column of this plot. The standardized mean difference (*d*) in children's depression score (*X*) between treatment group (*t*) with physical activity and control group (*c*) is computed via $d =$

$(\overline{X_t} - \overline{X_c})/S_{pooled}$, with $S_{pooled} = \sqrt{[(n_t - 1)SD_t^2 + (n_c - 1)SD_c^2]/(n_t + n_c - 2)}$, where *n* is sample size and *SD* is a standard deviation.

This effect size represents difference in children's depression between the treatment and control groups on the standard metric and thus enables us to compare the effects across studies. A couple of other commonly used effect sizes are the correlation coefficient (*r*) and the odds ratio, both of which represent the strength and direction of the relationship of interest. In this example, a standardized mean difference of 0 (the center) represents no difference on children's depression between treatment and control groups: A negative value (falling left of the center) indicates a lower mean on children's depression in treatment group and thus a beneficial effect of treatment; and a positive value (falling right of the center) indicates a higher mean on children's depression in treatment. As indicated in Figure 11.1, the strongest treatment effect (farthest from the center to the left) is found in the study #7 with *d* = -1.82.

The next element of the meta-analysis is the ***variance*** of the effect size (or the ***standard error***, the square root of the variance). Because the sampling distribution of the effect size is known as the [within-study] variance of the effect size is known, meta-analysis is also referred to as "level-1 variance-known or V-known application" (Raudenbush & Bryk, 2002, p. 207) in the Multilevel Modeling (Hierarchical Linear Modeling) literature. In this plot, the variance of standardized mean difference (shown in the third column of this plot) is computed via $v(d) = [(n_t + n_c)/n_t n_c] + d^2/[2(n_t + n_c)]$. The variance can be used to construct the confidence interval (CI) around the effect size (shown in the last column of this plot). The variance is a measure of ***precision*** of the estimated effect size. The bigger the variance of an effect size is, the wider the confidence interval will be, which reflects a less precise effect size. In this example, study #3 has the widest confidence interval around the effect size with the largest variance ($v(d)$

= 0.60), while study #1, study #12, and study #13 have the narrowest confidence intervals around the effect size with the smallest variances ($v(d) = 0.05$).

The fourth element is ***study weight*** (w), which is displayed in the fifth column of this plot. Study weight varies in its size across the studies and is the inverse of the variance of the effect size. In meta-analysis, bigger weights are assigned to effect sizes that are estimated more precisely with smaller variance. In this example, the biggest study weights are assigned to study #1, study #12, and study #13 ($w = 7.62\%$), while the smallest study weight is assigned to study #3 ($w = 1.60\%$). Because the variance of the effect size is primarily dependent upon the sample size,[2] study effects are often known as being weighted by the sample size (Borenstein et al. 2009).

Goals of Meta-analysis

With the four key elements in mind, the researcher conducts meta-analysis to reach two main goals. These include (1) to compute the overall effect size, and (2) to explore the potential variation in effect sizes. The first goal is to compute the overall effect size of study findings on the question of interest. This is referred as the ***overall analysis***. On the forest plot shown in Figure 11.1, the overall effect of the 18 studies is shown on the bottom of the figure. The estimated overall effect is -0.49, with a 95% confidence interval of -0.69 and -0.28, indicating a significant treatment effect of physical activity in lowering the level of children's depression. An overall effect is simply a weighted mean ($\bar{d}_{..}$) of the individual effects, computed via

$\bar{d}_{..} = \sum_{i=1}^{k} w_i d_i / \sum_{i=1}^{k} w_i$, where d_i is the effect size for the i^{th} study for $i = 1, \ldots, k$, k is the number of studies, and w_i is the inverse-variance weight for the i^{th} study.

Depending on whether the participants in the included studies were assumed to be sampled from a single population, a fixed- or random-effects model is chosen to estimate the

overall effect (for a more extensive discussion see Borenstein et al., 2009). Under the fixed-effects model, it is assumed that samples from selected studies come from the same population and they vary only due to the sampling error. Thus, the common effect is computed based solely on the effects inversely weighted by their sampling variance (v_i), $w_i = 1/v_i$. Under the random-effects model, it is assumed that samples from selected studies do not come from the same population and so the true effect varies across studies. Thus, an average effect is estimated based on the effect sizes inversely weighted by both the between-study difference[3] (s_q^2) and their sampling variance (v_i), $w^*_i = 1/v^*_i = 1/(s_q^2 + v_i)$.

The second goal of meta-analysis is to explore the potential sources of between-study variation in the effect sizes using exploratory variable(s), which is referred to as ***moderator analysis***. If the true effect size is consistent (and a fixed-effects model is adopted in estimating the overall analysis), the focus is solely on estimating an overall effect. However, when the true effect varies considerably across studies (and a random-effects model is adopted in estimating the overall analysis), the important task for meta-analyst is to identify the source of variation. Figure 11.2 shows a forest plot with two subgroups, which differ by whether a Randomized Control Trial (RCT) is used or not (non-RCT). In this example, the first fourteen studies were based on RCTs and yielded an overall effect of -0.44 with a 95% CI of -0.73 and -0.14, while the last four studies were based on non-RCTs and yielded the overall effect of -0.51 with a 95% CI of -0.77 and -0.24. This result indicates that no significant difference exists in the treatment effect between the RCT studies and the non-RCT studies as the CIs for two groups are overlapping.

Insert Figure 11.2 about here

For moderator analysis, either a fixed- or a mixed-effects model with a moderator is available. In particular, when an exploratory variable(s) is assumed to explain all the between-study difference in the effect sizes, a fixed-effects model with a moderator(s) is deemed appropriate. However, when an exploratory variable(s) is assumed to explain some but not all the between-study difference in the effect sizes, a mixed-effects model with a moderator analysis is more appropriate.

Utility of the Methodology in Sport and Exercise Science

Based on the computerized searches of MEDLINE and PsycINFO in 2009, Becker and Ahn (2012) found that the change in the number of the published peer-reviewed meta-analyses in the area of Sport and Exercise Science between 1975 and 2009 follows an exponential growth line, with an R^2 value of .99. This trend has continued in the last five years, with a total number of 108,227 peer-reviewed journal articles published between 2010 and 2015 in our informal searches of MEDLINE and PsycINFO conducted in February, 2015 using the key terms “meta-analysis”, “sports science”, “kinesiology”, and “exercise”.

In his review of meta-analyses in sport and exercise research, Hagger (2006) pointed out that many researchers in the area employ the fixed-effects model, with only a few discussing difference between the fixed- and random-effects model. Hagger also noted that, when the effect (e.g., the effectiveness of physical activity) is examined on multiple outcomes/constructs (e.g., depression, self-concept, and externalizing problems), researchers in the area either summarize the effects by each outcome separately or conduct the moderator analysis comparing the effects by type of outcome. Likewise, when more than one effect sizes on several outcomes/constructs are extracted from the same study, yielding a multivariate data, the current practice in the area relies on the univariate meta-analysis. However, multivariate meta-analysis may be more

appropriate for such data. Hagger recommended researchers adopt the best meta-analytic practice that is appropriate for the nature of the data and its underlying assumption.

Insert Figure 11.3 about here

Figure 11.3 presents a flowchart that can be used to guide meta-analytic researchers to choose the appropriate meta-analytic model. As shown in Figure 11.3, multivariate meta-analysis might be an appropriate option for multivariate data when (1) the number of effect sizes is sufficiently large and (2) statistical information needed to compute the dependent effect sizes and their variance-covariance is available. Otherwise, univariate meta-analysis is likely more appropriate. Regardless of the meta-analytic approach used, the decision of whether to use a fixed- or a random-effects model for the overall analysis and whether to use a fixed- or a mixed-effects model for the moderator analysis should be made based on both theoretical and statistical assumptions about the nature of the data (one true effect size vs. multiple true effect size). More discussion on the choice of models can be found in Borenstein et al. (2009), Cooper and Hedges (1994), Cooper, Hedges, and Valentine (2014) and Hedges and Vevea (1998).

The Substantive Example

Keeping the general guidelines on choosing either a univariate or a multivariate meta-analysis in mind, we illustrate a step-by-step example to demonstrate both methods using the free software R. Our illustration follows four steps: Step 1, *data management*; step 2, *test of the homogeneity assumption in effects*; Step 3, *overall analysis*; and step 4, *moderator analysis*.

In this example, we use the data from a published meta-analysis (Craft, Magyar, Becker, & Feltz, 2003). Using correlation coefficients (*r*) as a primary effect size measure, the main purposes of their meta-analysis were to examine the interrelationships between three subscales: cognitive anxiety, somatic anxiety, and self-concept measured by the Competitive State Anxiety

Inventory (CSAI-2) and athletic performance. We further explore the potential moderators that might explain the variation observed across the effect sizes.

Craft et al. (2003) extracted 246 correlations from 29 independent studies (20 published and 9 unpublished studies) that report at least one correlation coefficient describing the relationship among cognitive anxiety, somatic anxiety, self-concept, and sport performance in athletes. Across the 29 included studies, the number of reported correlations from a single study varied from one to six, with a mean of 3.6 out of 6 possible correlations (i.e., $[4 \times 3] / 2 = 6$). For the purpose of demonstration, we use a subset of the data that provides all six correlation coefficients among cognitive anxiety, somatic anxiety, self-concept, and sport performance in athletes.

Insert Table 11.1 about here

Table 11.1 presents the eighteen studies that provide all six correlations among the four variables of interest. These include (1) the correlation between cognitive anxiety and somatic anxiety (C1), (2) the correlation between cognitive anxiety and self-confidence (C2), (3) the correlation between somatic anxiety and self-confidence (C3), (4) the correlation between cognitive anxiety and athletic performance (C4), (5) the correlation between somatic anxiety and athletic performance (C5), and (6) the correlation between self-confidence and athletic performance (C6). Sample size (*n*), gender (*Gender*: Male, Female, and Mixed-gender group), and % of male participants (*% Male*) are also presented along with the six reported correlations.

The Synergy

Univariate Meta-analysis

A univariate meta-analysis is first presented to investigate two distinct research questions: 1) what is the overall effect across studies? and 2) is the variation in study effects

explained by hypothesized moderating variables such as gender or race? In other words, overall and moderator analyses are conducted using the data provided. The overall relationships between each of the three subscales (i.e., cognitive anxiety, somatic anxiety, and self-concept) of the CSAI-2 and athletic performance (C4, C5, and C6 in Table 11.1) are first estimated. Then, the follow-up moderator analyses are conducted to examine the effect of gender as a potential moderator on the effect sizes. All analyses were conducted using the *metafor* package (Viechtbauer, 2010) written in the free statistical software R (R Development Core Team, 2014).

Step 1: Data management. The first step involves two main tasks for creating a dataset for univariate meta-analysis in the R statistical software. First, effect sizes extracted from the included studies should be stacked in a column. For instance, as shown in the dataset (*craft.txt*), the *r* between somatic anxiety and performance are stacked and saved in a column called C5. For missing correlations, "NA" should be entered in the dataset. This can be done in Excel and should be saved as a "tab delimited" file.

Once the "tab delimited" dataset is imported to the R statistical software, the next step is to compute the effect size (*r*) and its variance. A transformation is recommended to stabilize the raw correlation, *r* (Becker, 2000). Therefore, in this example the *r* was transformed to Fisher's *z* via $z_r = .5\log(1+r)/(1-r)$ and then its associated variance ($v(z_r)$) was computed via $v(z_r) = 1/(n-3)$. These transformations can all be easily done using the "*escalc*" function available in the *metafor* package written in the R statistical software (see Appendix 11.1 for R code).

Step 2: Test of the homogeneity assumption in effects. In step 2, we seek to determine whether to use a fixed- or random-effects model for an overall analysis. Theoretically, the relationship between each of the three subscales of the CSAI-2 and athletic performance would

differ depending on a number of factors such as gender, ethnicity, muscle mass, and experience. Therefore, a random-effects model, which assumes that the true relationships between these variables vary from study to study, appears appropriate. In addition to our theoretical justification, we can test the assumption of homogeneity in effects regarding whether effects are from the same population.

The test of the homogeneity of variance of the effect sizes is based on the Q_{total} statistic (Cooper et al., 2009), which quantifies a total observed variance in effect sizes across the k studies computed via $Q_{total} = \sum_{i=1}^{k} w_i(r_i - r_.)^2$ where $w_i = 1/v_i$. Under the null hypothesis that all the studies come from the same population, the Q_{total} will follows a central chi-squared distribution with degrees of freedom equal to $k - 1$. If the Q_{total} is found to be statistically significant at a preset alpha level, it is assumed that study effects are not from the same population, suggesting that a random-effects model should be adopted for the overall analysis. Otherwise, a fixed-effects model should be adopted for the overall analysis.

Insert Table 11.2 about here

In our example, Q_{total} was computed using "*rma*" function available in the *metafor* package written in the R statistical software (see Appendix 11.1 for R code). The Q_{total} statistics for C4, C5, and C6 are shown in the third column of Table 11.2. The tests of the homogeneity assumption in effect sizes indicate that statistically significant between-study variation exists in the relationship (C4) between cognitive anxiety and athlete performance (Q_{total} (17) = 28.57, p = .04), the relationship (C5) between somatic anxiety and athlete performance (Q_{total} (17) = 29.84, p = .03), and the relationship (C6) between self-confidence and athlete performance (Q_{total} (17) = 75.19, $p < .01$). These suggest that the overall relationships for C4, C5, and C6 should be

estimated under the random-effects model, in which effects are assumed to come from different populations.

Step 3: Overall analysis. In step 3, the overall effect size is estimated under the chosen model using *rma* function available in the *metafor* package written in the R statistical software (see Appendix 11.1 for R code). For our example, overall effects were estimated under the random-effects model, where study effects were weighted by the sum of the between-study variance (s_q^2) as well as the within-study variance (v_i). Under the random-effects model, the estimated overall correlations for C4, C5, and C6 were -.06 (95% CI: -.16, .05), -.04 (95% CI: -.14, .06), and .06 (95% CI: -.07, .19), respectively, and they were all found to be not statistically significant. These results suggest that none of the three subscales (i.e., cognitive anxiety, somatic anxiety, and self-concept) of CSAI-2 is related to athletic performance. Figures 11.4 – 11.6 and Table 11.2 summarize the results based on a random-effects model on C4, C5, and C6.

Insert Figures 11.4 – 11.6 about here

Step 4: Moderator analysis. In step 4, when the significant amount of between-study variation in effect sizes exists (when the Q_{total} is found to be significant in the preceding step), a series of moderator analysis can be performed to find the source of variation among the effects. For this example, we will use gender with three subgroups (Female, Male, and Mixed-gender group) as a categorical moderator and the percentage of males in the sample (*p_male*) as a continuous moderator to see whether each of the relationships (C4, C5, and C6) significantly differs by gender (either *gender* or *p_male*).

The first step in a moderator analysis is to determine whether to use the fixed- vs. mixed-effects model on theoretical and statistical bases. In particular, two pieces of the statistical evidence that helps researchers choose the model are the significance tests of (1) Q_{model}, which

quantifies the amount of variation in effects explained by a moderator; and (2) Q_{error}, which quantifies the amount of variation in effects unexplained by a moderator. When a moderator is assumed to explain all between-study variation in effects (i.e., Q_{model} is significant, but Q_{error} is not significant), a fixed-effects model with a moderator is deemed appropriate. When a moderator is assumed to explain some but not all between-study variation in effects (i.e., both Q_{model} and Q_{error} are significant), a mixed-effects model with a moderator is deemed appropriate. When Q_{model} is not found to be statistically significant, another moderator should be sought out to explain the between-study variation in effects.

Insert Table 11.3 about here

For our example, as shown in Table 11.3, none of moderators were found to significantly explain the between-study variation in the relationships (C4) between cognitive anxiety and athletic performance (Q_{model} (1) = .02, p = .88 for % of males in the sample, Q_{model} (2) = 2.43, p = .30 for gender) or the relationships (C5) between somatic anxiety and athletic performance (Q_{model} (1) = 2.34, p = .13 for % of males in the sample, Q_{model} (1) = 5.58, p = .06 for gender). Both moderators were found to significantly explain some but not all between-study variation in the relationship (C6) between self-confidence and athletic performance (Q_{model} (1) = 7.21, p < .01 and Q_{error} (16) = 67.98, p < .01 for % of males in the sample; Q_{model} (2) = 31.55, p < 01 and Q_{error} (15) = 43.65, p < .01 for gender). Therefore, a mixed-effects model was applied to perform a moderator analysis with % of males in the sample or gender.

Under the mixed-effects model with % of males in the sample as a continuous moderator, neither the intercept nor the slope was found to be statistically significant (Q_{model} (2) = .0004, p = .98). In particular, an insignificant intercept of .07 (95% CI: -.36 and .47) suggests that the relationship between self-confidence and athletic performance was not significantly different

from zero when females are included in the sample (% of males in the sample equals to zero). And, an insignificant slope related to % of males of -.0001 (95% CI: -.005 and .005) suggests that the change in the relationship between self-confidence and athletic performance is not significant for an additional 1% increase in males in the sample.

Under the mixed-effects model with a gender as a categorical moderator, the estimated relationship (C6) between self-confidence and athletic performance was not statistically significant for female-only ($\bar{r}$ = -.14, 95% CI: -.51 and .27) and male-only samples ($\bar{r}$ = .01, 95% CI: -.13 and .15), while it was statistically significant for mixed-gender sample ($\bar{r}$ = .32, 95% CI: .06 and .53). However, the estimated correlation between self-confidence and athletic performance was not significantly different among those three gender groups (Q_{model} (2) = 6.44, p = .09). Typically, when the Q_{model} value is not significant for a moderator, separate group analyses are not conducted. However, they are conducted in this example for the purpose of illustration.

Multivariate Meta-analysis

A multivariate meta-analysis is performed on the same dataset to investigate research questions similar to those we examined in the univariate meta-analysis: 1) what is the overall effect across studies after accounting for interrelationships among multiple effects? and 2) is the variation in study effects explained by hypothesized moderating variables such as gender or race after accounting for interrelationships among multiple effects? In other words, overall and moderator analyses are conducted using the data, but the effects are estimated more precisely by incorporating non-zero correlations among dependent effect sizes (Becker, 2009; Kim & Becker, 2010; Nam, Mengerson, & Garthwaite, 2003).

Presently, the overall relationships between each of the three subscales (i.e., cognitive anxiety, somatic anxiety, and self-concept) of the CSAI-2 and athletic performance (C4, C5, and C6 in Table 11.1) after accounting for interrelationships among three subscales (i.e., cognitive anxiety, somatic anxiety, and self-concept: C1, C2, and C3 in Table 11.1) are first estimated. Then, follow-up moderator analyses are conducted to examine if each relationship varies by a potential moderator (i.e., gender) after accounting for interrelationships among three subscales (i.e., cognitive anxiety, somatic anxiety, and self-concept: C1, C2, and C3 in Table 11). A number of different data analytic techniques are available for multivariate meta-analysis; our presentation in this section is based on the generalized least squares (GLS) method (Becker, 1992, 2000). All analyses were conducted using the *mvmeta* package (Gasparrini, Armstrong, & Kenward, 2012) written in the free R statistical software (R Development Core Team, 2014**).**

Step1: Data management. There are two main tasks for creating a dataset for multivariate meta-analysis. First, all the observed correlation values across the 18 studies are stacked in a column. As shown in the dataset (*craft.txt*), each of six correlations (C1 to C6) among four measures extracted from the 18 studies are saved in each column and saved as a "tab delimited" file.

Once the "tab delimited" dataset is imported to the R software, the next step is to transform the six reported raw correlation values (C1 to C6) to Fisher's *z* values via $z_{st} = .5\log(1+r_{st})/(1-r_{st})$ where r_{st} is observed correlation between *s*th and *t*th variables and four variable index labels *s*, *t*, *u*, and *v* range from 1 to the number of variables in correlation matrix. Then, the associated within-study variance-covariance matrix for the transformed Fisher's *z* values for each study should be computed using equations proposed by Olkin and Siotain (1976): $\text{var}(z_{st}) = 1/(n-3)$ and $\text{cov}(z_{st}, z_{uv}) = SD_{st,uv}/\left[(1-r_{st}^2)(1-r_{uv}^2)\right]$. Appendix 11.2 provides R syntax

code written by authors for computing within-study variance-covariance matrix for the transformed Fisher's *z*.

Step 2: Test of the homogeneity assumption in effect-size vector. In step 2, the test of the homogeneity assumption in an effect-size vector is performed in order to see whether the vector of effect-size measures is from the same population. For our example, a homogeneity test of the effect-size vector indicates that the set of six observed correlations do not come from the same (common) population (Q_{total} (102) = 349.86, $p < .01$), suggesting that a vector of overall correlations among four measures should be estimated under the random-effects model rather than the fixed-effects model. A multivariate homogeneity test was performed using "*mvmeta*" function available in the *mvmeta* package (see Appendix 11.2 for R code).

Step 3: Overall analysis. In step 3, a set of effect sizes on multiple measures is estimated using either a fixed- or random-effects model depending on the results of a multivariate homogeneity test described above. In our example, as suggested in step 2, overall relationships among four variables were estimated under the random-effects model.

Insert Tables 11.4 – 11.5 about here

Under the random-effects model, the estimated correlations among the four variables are shown in Tables 11.4 – 11.5. Of the six estimated correlations, three pairs of variables – cognitive anxiety and somatic anxiety ($\bar{r}$ = .47, z = 7.90, $p < .01$, 95% CI: .36 and .56), cognitive anxiety and self-confidence ($\bar{r}$ = -.26, z = -6.34, $p < .01$, 95% CI: -.34 and -.18), and somatic anxiety and self-confidence ($\bar{r}$ = -.34, z = -4.34, $p < .01$, 95% CI: -.47 and -.19) – were found to be statistically significant. However, none of the three subscales of CSAI-2 (i.e., cognitive anxiety, somatic anxiety, and self-concept) were related to athletic performance: -.02 (95% CI: -.13, .08) for C4, -.004 (95% CI: -.10, .10) for C5, and .03 (95% CI: -.10, .16) for C6. These were

all estimated using *mvmeta* function available in the *mvmeta* package written in the R statistical software (see Appendix 11.2 for R code).

Step 4: Moderator analysis. In step 4, a fixed-effects model with predictors or a mixed-effects model can be performed to explore the sources of between-study variations in a set of effect sizes on multiple measures. In our example, we used the percentage of male participants as a potential moderator and ran a mixed-effects model to see whether the effect sizes differed by the percentage of males in the sample (*p_male*). These were all estimated using "*mvmeta*" function available in the *mvmeta* package (see Appendix 11.2 for R code).

Insert Table 11.6 about here

A significant Q_{model} indicates that the percentage of male athletes significantly explained the between-study variation in a set of correlations among four variables (Q_{model} (6) = 43.88 [Q_{total} of 349.86 estimated in step 2 – Q_{error} of 305.98 estimated in step 4], $p < .01$). As shown in Table 11.6, the relationship between cognitive anxiety and somatic anxiety was found to be statistically significant and positive for female athletes ($\bar{r}$ = .57, z = 3.11, p = .002, 95% CI: .23 and .78). However, the relationship between cognitive anxiety and self-confidence was found to be negative and significant for female athletes ($\bar{r}$ = -.31, z = -2.07, p = .03, 95% CI: -.55 and -.02). Results suggest that cognitive anxiety is positively related to somatic anxiety but is negatively related to self-confidence among female athletes.

Summary

Since being introduced, meta-analysis has been widely used in sport and exercise science, and we believe that this will continue to be the case. The increased number of available computer programs for meta-analysis makes meta-analytic research more accessible than ever before. However, the ease of use of meta-analytic software might

result in poorly conducted analyses unless researchers have a thorough understanding of the underlying principles of meta-analysis and its application appropriate for the nature of data as well as the underlying assumptions. Our hope is to guide researchers to adopt the best available meta-analytic practice. This chapter provided a brief overview of meta-analysis and discussed how to choose the most appropriate meta-analytic methods. We offered a step-by-step tutorial in R that illustrates the process of performing both univariate and multivariate meta-analysis using a real dataset in the area of Sport and Exercise Science.

Univariate meta-analysis is an appropriate option for synthesizing the study effect of a single outcome/construct. However, researchers in practice are often interested in examining study effects on multiple outcomes/constructs rather than on a single one. In such cases, more than one effect size may be extracted from a study and consequently these effect sizes are dependent on one another, which leads to a multivariate scenario. Other circumstances that produce a multivariate data in meta-analysis are when multiple comparison groups are used within the study and their pairs are contrasted to address treatment effects (Gleser & Olkin, 1994; 2009; Ryan, Blakeslee, & Furst, 1986), and(or) when the same participants are measured on multiple outcomes over several time periods (Wei & Higgins, 2013a; 2013b).

Different univariate meta-analytic approaches are available for dealing with a multivariate data, including ignoring dependence of effect sizes, averaging dependent effect sizes, and separating effect sizes into independent subgroups and analyzing them separately by subgroup (Borenstein et al., 2009). However, when multivariate meta-analysis, with non-zero within-study correlations among dependent effect sizes (Becker, 2009; Kim & Becker, 2010;

Nam et al., 2003) is used for data analysis, it provides more precise estimates and control for Type I error rate. Although multivariate meta-analysis is known to hold many advantages (Becker, 1992, 2009; Becker & Schram, 1994; Gleser & Olkin, 2009; Kim & Becker, 2010; Nam et al., 2003; van Houwelingen, Arends, & Stijnen, 2002; Wei & Higgins, 2013b), a number of challenges have been identified in its application to practice.

Some challenges in applying multivariate meta-analysis include dealing with missing information in estimating effect sizes and their associated within-study variance-covariance matrix of multiple correlated effect sizes (Becker, 1992, 2009; Furlow & Beretvas, 2005; Ishak, Platt, Joseph, & Hanley, 2008; Riley, 2009; Riley, Thompson, & Abrams, 2008; Wei & Higgins, 2013a), an estimation of between-study variance-covariance matrix in performing a random-effects model or mixed-effects model (White, 2009), the number of primary studies exploring outcomes/constructs of researcher's interest that is sufficient enough for multivariate meta-analysis, and potential differences in the psychometrics properties of outcome(s) representing the underlying construct. Likewise, multivariate meta-analysis might not be applicable for all the circumstances. Therefore, researchers are encouraged to evaluate whether to use univariate or multivariate method in advance, and choose the most appropriate method for their dataset.

In this chapter, we demonstrated to the steps for perform both univariate and multivariate meta-analysis using an existing dataset for answering essentially the same research questions. We do so in hopes to equip readers with the knowledge of available tools, and provides information that helps to choose the most appropriate one for their own research. We believe that paying more attentions to the recent advances in meta-analysis and the capacities of available computer software available to conduct meta-analysis will lead to more accurate results. In addition, we encourage researchers to adhere the guidelines for reporting meta-

analytic results. These include (1) Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA), which can be found in http://www.prisma-statement.org/, and (2) Meta-Analysis Reporting Standards (MARS), which can be found in https://www.apa.org/pubs/authors/jars.pdf.

References

Ahn, S., & Fedewa, A. L. (2011). A meta-analysis of the relationship between children's physical activity and mental health. *Journal of pediatric psychology, 36*, 385-397. doi: 10.1093/jpepsy/jsq107

Becker, B. J. (1992). Using results from replicated studies to estimate linear models. *Journal of Educational and Behavioral Statistics, 17*, 341-362. doi: 10.2307/1165128

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling.* (pp. 499-525). San Diego, CA, US: Academic Press.

Becker, B. J. (2009). Model-based meta-analysis. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis (2nd ed.).* (pp. 377-395). New York, NY, US: Russell Sage Foundation.

Becker, B. J., & Ahn, S. (2012). Synthesizing measurement outcomes through meta-analysis. In G. Tenenbaum, R. C. Eklund & A. Kamata (Eds.), *Measurement in sport and exercise psychology.* (pp. 153-168). Champaign, IL, US: Human Kinetics.

Becker, B. J., & Schram, C. M. (1994). Examining explanatory models through research synthesis. In H. Cooper and L. V. Hedges (Eds.). *The handbook of research synthesis* (pp. 357-381). New York: Russell Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, England: John Wiley & Songs, Ltd.

Cooper, H. M. (1984). *The integrative research review: A systematic approach*. Beverly Hills, Calif: Sage Publications.

Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach (4th ed.)*.

Thousand Oaks, CA, US: Sage Publications, Inc.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York, NY, US: Russell Sage Foundation.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis (2nd ed.)*. New York, NY, US: Russell Sage Foundation.

Craft, L. L., Magyar, T. M., Becker, B. J., & Feltz, D. L. (2003). The Relationship Between the Competitive State Anxiety Inventory-2 and Sport Performance: A Meta-Analysis. *Journal of Sport & Exercise Psychology, 25*(1), 44-65.

Furlow, C. F., & Beretvas, S. N. (2005). Meta-analytic methods of pooling correlation matrices for structural equation modeling under different patterns of missing data. *Psychological methods, 10*, 227-254. doi: 10.1037/1082-989X.10.2.227

Gasparrini A., Armstrong, B., & Kenward M. G. (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*. *31,* 3821-3839. doi: 10.1002/sim.5471

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis.* (pp. 339-355). New York, NY, US: Russell Sage Foundation.

Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis (2nd ed.).* (pp. 357-376). New York, NY, US: Russell Sage Foundation.

Hagger, M. S. (2006). Meta-analysis in sport and exercise research: Review, recent

developments, and recommendations. European Journal of Sport Sciences, 6, 103-115. doi: 10.1080/17461390500528527

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

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. doi: 10.1037/1082-989X.3.4.486

Ishak, K. J., Platt, R. W., Joseph, L., & Hanley, J. A. (2008). Impact of approximating or ignoring within-study covariances in multivariate meta-analyses. *Statistics in Medicine, 27*, 670-686. doi: 10.1002/sim.2913

Kim, R.-S., & Becker, B. J. (2010). The degree of dependence between multiple-treatment effect sizes. *Multivariate Behavioral Research, 45*, 213-238. doi: 10.1037/a0032734

Nam, I. S., Mengersen, K., & Garthwaite, P. (2003). Multivariate meta‐analysis. *Statistics in medicine, 22*, 2309-2333. doi: 10.1002/sim.1410

Olkin, I., & Siotain, M. (1976). Asymptotic distribution of function of a correlation matrix. In I. S. Ikeda (Ed.), *Essays in probability and statistics* (pp. 235-251). Tokyo, Japan: Shinko Tsusho Co., Ltd.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and data analysis methods* (*2nd ed*). Thousand Oaks, CA, U.S.: Sage Publications.

R Development Core Team (2014). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

Riley, R. D. (2009). Multivariate meta-analysis: the effect of ignoring within-study correlation. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 172*, 789-811. doi: 10.1111/j.1467-985X.2008.00593.x

Riley, R. D., Thompson, J. R., & Abrams, K. R. (2008). An alternative model for bivariate random-effects meta-analysis when the within-study correlations are unknown. *Biostatistics, 9*, 172-186. doi: 10.1093/biostatistics/kxm023

Ryan, E. D., Blakeslee, T., & Furst, D. M. (1986). Mental practice and motor skill learning: An indirect test of the neuromuscular feedback hypothesis. *International Journal of Sport and Exercise Psychology, 17*(1), 60-70.

Van Houwelingen, H. C., Arends, L. R., & Stijnen, T. (2002). Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in medicine, 21*, 589-624. doi: 10.1002/sim.1040

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational And Behavioral Statistics*, *30*, 261-293. doi: 10.3102/10769986030003261

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. URL: http://www.jstatsoft.org/v36/i03/.

Wanous, J. P., Sullivan, S. E., & Malinak, J. (1989). The role of judgment calls in meta-analysis. *Journal Of Applied Psychology*, *74*, 259-264. doi:10.1037/0021-9010.74.2.259

Wei, Y., & Higgins, J. P. T. (2013a). Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics In Medicine, 32*, 1191-1205. doi: 10.1002/sim.5679

Wei, Y., & Higgins, J. P. T. (2013b). Bayesian multivariate meta-analysis with multiple outcomes. *Statistics In Medicine, 32*, 2911-2934. doi: 10.1002/sim.5745

White, I. R. (2009). Multivariate random-effects meta-analysis. *Stata Journal, 9*(1), 40.

Footnotes

[1] Other terms that have been interchangeably used in the literature include *research synthesis* and *quantitative review*.

[2] Other factors affecting the precision of the effect size are study design (Borenstein, Hedges, Higgins, & Rothstein, 2009) and study quality (Cooper & Hedges, 1994).

[3] A number of methods are available for estimating the between-study variation (s_q^2). These include an estimator proposed by Hunter and Schmidt (1990), an estimator proposed by Hedges (1983), and estimator by Dersimonian and Laird (1986), and two estimators based on Maximum Likelihood (ML) estimation, and Restricted Maximum Likelihood (REML) estimation. Of these estimators, Viechtbauer (2005) showed that REML estimation provides the most unbiased and efficient estimator of between-study variance.

Table 11.1

Relationships between Three Subscales of CSAI and Athletic Performance in 18 Studies

ID	*n*	Gender	% Male	Correlation (*r*)					
				C1	C2	C3	C4	C5	C6
Caruso1990a	24	Male	100	0.240	-0.400	-0.340	-0.170	0.080	0.210
Caruso1990b	27	Male	100	0.420	-0.420	-0.480	-0.080	0.170	-0.120
Caruso1990c	30	Male	100	0.330	-0.190	-0.590	-0.140	-0.110	0.090
Barnes1986	14	Male	100	0.212	-0.544	-0.431	-0.391	-0.166	0.191
Edwards1996	45	Female	0	0.470	-0.370	-0.500	0.100	0.310	-0.170
Maynard1995a	24	Male	100	0.670	-0.360	-0.720	-0.150	-0.400	0.350
Maynard1995b	24	Male	100	0.670	-0.410	-0.720	-0.240	-0.240	0.360
Maynard1995c	24	Male	100	0.130	-0.040	-0.500	-0.060	-0.160	0.220
Rodrigo1990	51	Male	100	0.570	-0.180	-0.260	-0.520	-0.430	0.160
McAuley	7	Female	0	0.171	-0.311	-0.177	0.106	-0.008	-0.007
Abouzekri2010a	61	Male	100	0.425	-0.128	-0.158	0.084	0.086	-0.218
Abouzekri2010b	61	Male	100	0.145	-0.049	-0.219	0.066	0.061	-0.187
Abouzekri2010c	61	Male	100	0.456	-0.136	-0.124	0.120	-0.008	-0.322
Kais2005a	24	Male	100	0.500	-0.520	-0.500	0.290	-0.150	0.000
Kais2005b	24	Male	100	0.500	-0.520	-0.500	0.290	-0.150	0.000
Nicholls2010	307	Mixed	82	0.490	-0.200	-0.380	-0.080	-0.030	0.130
Otten2009	243	Mixed	63	0.790	-0.330	-0.230	-0.150	-0.110	0.490
Sanchez2010	19	Male	100	0.109	-0.668	0.690	0.130	0.542	-0.158

Note. C1: correlation (*r*) between cognitive anxiety and somatic anxiety; C2: *r* between cognitive anxiety and self-confidence; C3: *r* between somatic anxiety and self-confidence; C4; *r* between cognitive anxiety and athletic performance; C5: *r* between somatic anxiety and athletic performance; C6: *r* between self-confidence and athletic performance.

Table 11.2

Estimated Correlations among Each subscale of CSAI from the Univariate Meta-analysis

					95% CI	
Interrelationship	*k*	Q_{total}	*df*	$\bar{r}$	*LL*	*UL*
Cognitive anxiety and Athletic performance (C4)	18	28.57*	17	-.06	-.16	.05
Somatic anxiety and Athletic performance (C5)	18	29.84*	17	-.04	-.14	.06
Self-confidence and Athletic performance (C6)	18	75.19**	17	.06	-.07	.19

Note. k: Number of effects; * p <.05; ** p < .01

Table 11.3

Results from Univariate Moderator Analysis

Moderators	*k*	Q_{model}	df_{model}	Q_{error}	df_{error}	$\bar{r}$	95% CI *LL*	95% CI *UL*
Cognitive anxiety and Athletic performance (C4)								
% of males in the sample	18	.02	1	28.55*	16	-	-	-
Gender	18	2.43	2	26.13*	15	-	-	-
Somatic anxiety and Athletic performance (C5)								
% of males in the sample	18	2.34	1	27.50*	16	-	-	-
Gender	18	5.58	2	24.25	15	-	-	-
Self-confidence and Athletic performance (C6)								
% of males in the sample	18	7.21**/.0004†	1	67.98**	16			
Intercept						.07	-.36	.47
Slope related to % of males in the sample						-.0001	-.005	.005
Gender	18	31.55**/6.44†	2	43.64**	15			
Female	14					-.14	-.51	.27
Male	2					.01	-.13	.15
Mixed-gender group	2					.32*	.06	.53

Note. k: Number of effects; * p <.05; ** p < .01; † Q_{model} was estimated under the mixed-effects model.

Table 11.4

Estimated Correlations between Three Subscales of CSAI and Athletic Performance using Multivariate Meta-analysis

Interrelationship	*k*	$\bar{r}$	95% CI	
			LL	*UL*
Cognitive anxiety and Somatic anxiety (C1)	18	.47**	.36	.56
Cognitive anxiety and Self-confidence (C2)	18	-.26**	-.34	-.18
Somatic anxiety and Self-confidence (C3)	18	-.34**	-.47	-.19
Cognitive anxiety and Athletic performance (C4)	18	-.02	-.13	.08
Somatic anxiety and Athletic performance (C5)	18	-.004	-.10	.10
Self-confidence and Athletic performance (C6)	18	.03	-.10	.16

Note. k: Number of effects; ** $p < .01$

Table 4.4

Estimated Correlation Matrix among Three subscales of CSAI and Athletic Performance using Multivariate Meta-analysis

	Cog	Som	SC	P
Cognitive anxiety (Cog)				
Somatic anxiety (Som)	.47**			
Self-confidence (Sc)	-.26**	-.34**		
Performance (P)	-.02	-.004	.03	

Note. k: Number of effects; ** $p < .01$

Table 11.6

Results from Multivariate Moderator Analysis using % of Males in the Samples

		95% CI	
Interrelationships	$\bar{r}$	*LL*	*UL*
Cognitive anxiety and Somatic anxiety (C1)			
Intercept	.57**	.23	.78
% of males in the sample	-.002	-.006	.003
Cognitive anxiety and Self-confidence (C2)			
Intercept	-.31*	-.55	-.02
% of males in the sample	.0005	-.003	.004
Somatic anxiety and Self-confidence (C3)			
Intercept	-.37	-.73	.16
% of males in the sample	.0002	-.006	.006
Cognitive anxiety and Athletic performance (C4)			
Intercept	.08	-.28	.41
% of males in the sample	-.001	-.005	.003
Somatic anxiety and Athletic performance (C5)			
Intercept	.30	-.04	.57
% of males in the sample	-.004	-.008	.0002
Self-confidence and Athletic performance (C6)			
Intercept	.03	-.39	.44
% of males in the sample	.0001	-.005	.005

Note. * $p < .05$; ** $p < .01$

Figure Captions

Figure 11.1 Forest plot of 18 treatment effects of physical activity on children's depression included in Ahn and Fedewa (2011).

Figure 11.2 Forest plot of 18 treatment effects of physical activity on children's depression by RCT vs. non-RCT studies

Figure 11.3 Flowchart for choosing the appropriate meta-analytic model

Figure 11.4 Forest plot of correlation between cognitive anxiety and athlete performance

Figure 11.5 Forest plot of correlation between somatic anxiety and athlete performance

Figure 11.6 Forest plot of correlation between self-confidence and athlete performance