CALCULATING EFFECT SIZES FOR META-ANALYSIS

The goal of modern meta-analysis is to understand general patterns emerging from consideration of results from many separate studies. Different studies vary in the significance of statistical tests and the sample sizes those statistical tests are based on. It should be fairly intuitive that a study based on large samples sizes would contribute more 'weight' than a study with small sample sizes when synthesizing results across many studies. That idea is the essence of modern meta-analysis, which emphasizes the calculation of an effect size to standardize results among studies.

There are many different measures of effect size, and it would be impractical to discuss them all here. For the purposes of this assignment, focus on studies that report statistical results based on a correlation coefficient (r), a t-test (t), an analysis of variance (F), or that report means and standard deviations (or counts) from experimental and control groups. All of these types of results can be converted to a metric of effect size relatively easily.

Calculating effect sizes for studies presenting means, standard deviations, and sample sizes

Some studies will include data on mean values in 'experimental' and 'control' groups—note that the groups do not have to be referred to as experimental or control in the study, but could be any two population mean differences (say a difference in seed mass between endangered and non-endangered plants). When you have data from multiple studies that report population means and standard deviations, you can calculate an effect size metric, Hedge's d. To calculate Hedge's *d*, the following applies: $\overline{X}^E$ is the mean value from one group, and $\overline{X}^C$ is the mean value from the other group, N^E is the sample size of group E, N^C is the sample size of group C, s^E and s^C are standard deviations of group E and C

Hedge's $d = \frac{\overline{X}^E - \overline{X}^C}{S} J$,with S, the pooled standard deviation calculated as:

$$S = \sqrt{\frac{(N^E-1)(s^E)^2+(N^C-1)(s^C)^2}{N^E+N^C-2}}$$

And J calculated as $1 - \frac{3}{4(N^C+N^E-2)-1}$

The variance of Hedge's *d* is calculated as:

$$v_d = \frac{N^C+N^E}{N^C N^E} + \frac{d^2}{2(N^C+N^E)}$$

In other words, with information on population mean values, samples sizes and standard deviations for each population, you can calculate the effect size and variance for each study.

Calculating effect sizes for studies using common statistical tests

The results of the most common univariate statistical tests (*r*, *t*, and *F*) can all be converted to effect sizes using similar transformations.

If studies include results of *t* or *F* tests, they can be converted to the corresponding r value as follows:

$$r = \sqrt{\frac{t^2}{t^2+df}} \qquad r = \sqrt{\frac{F}{F+df}}$$

Values of the correlation coefficient *r* along with transformations of t or *F* to *r* can be converted to the effect size metric Fisher's *z* as:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

with variance $v_z = \frac{1}{n-3}$

Calculating a global mean effect size

Regardless of whether you calculate an effect size using *d* or *z*, the effect sizes for each study and their variance are used to calculate a global mean effect size summed from the effect sizes (*E*) and weights of each study. Here $w_i = 1/v_i$

$$\bar{\bar{E}} = \frac{\sum w_i E_i}{\sum w_i}$$

The variance of $\bar{\bar{E}}$ is:

$$s^2_{\bar{\bar{E}}} = \frac{1}{\sum w_i}$$

From which we can calculate a confidence interval around $\bar{\bar{E}}$ as:

$$\text{CI} = \bar{\bar{E}} \pm t_{\alpha/2[n-1]} * s_{\bar{\bar{E}}}$$

The global mean effect size $\bar{\bar{E}}$ is the overall effect from a group of studies. If $\bar{\bar{E}}$ is significantly different from 0 (i.e., the CI around $\bar{\bar{E}}$ does not intersect 0) then the effect is considered significant.

Note that you cannot combine effect sizes calculated as *d* with *z* for calculation of the global mean effect size $\bar{\bar{E}}$. You will need to survey the literature and determine which effect size metric is best based on the way data are reported for your study.