How to Read a Meta-Analysis of Intervention Studies

What Is a Meta-Analysis?
Meta-analyses primarily seek to summarize past research by synthesizing empirical findings from multiple, separate investigations that address related or identical topics.

Following is key information that readers of meta-analyses should look for to understand the important takeaways.

Defining Study Criteria: MUTOS
Look for the criteria that define study eligibility and consider how they address components of the MUTOS framework:

- What methods are being used (e.g., study design)?
- What are the treatment characteristics (e.g., intervention dosage)?
- What is the setting (e.g., geographic locale)?
- What is the unit of the sample population (e.g., grade level)?
- What are the outcome measures (e.g., outcome measure type)?
- What other requirements did the meta-analyst have (e.g., year of publication)?

Identifying Eligible Studies
The meta-analysis may provide a flowchart similar to the graphic below, explicitly depicting the steps of searching, screening, and identifying eligible studies.

Interpreting Results
Meta-analytic results are expressed as an average of all included studies’ effect sizes—in other words, the average treatment effect for the intervention.

However, effects may differ from one study to the next; this is called “heterogeneity.” Moderator analyses attempt to explain why there is heterogeneity in effect sizes. Look for tables that report the average effect for each moderating variable.

Example: Mathematics Intervention Effects
The following example is from a meta-analysis examining heterogeneity in mathematics intervention effects in Grades PreK–12 (Williams et al., 2021). The overall effect size was moderate and statistically significant ($g = 0.31$, $SE = 0.03$, $p < 0.001$).

In this example, the difference in average effects tells us that studies that used researcher-generated measures yielded an average effect 0.30 standard deviations greater ($SE = 0.08$, $df = 39.80$, $p < .01$) than when standardized achievement measures were used (0.45 vs. 0.15).

<table>
<thead>
<tr>
<th>Moderator: Outcome type</th>
<th>Average effect</th>
<th>Standard error</th>
<th>No. of studies</th>
<th>No. of effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher-generated measure</td>
<td>0.45</td>
<td>0.05</td>
<td>123</td>
<td>639</td>
</tr>
<tr>
<td>Standardized achievement measure</td>
<td>0.15</td>
<td>0.05</td>
<td>107</td>
<td>470</td>
</tr>
</tbody>
</table>

The results of this example may also be expressed in the form of a graph, which serves as a visual representation of the differences in effect sizes between researcher-generated measures (red line) and standardized achievement measures (blue line) across grade levels.