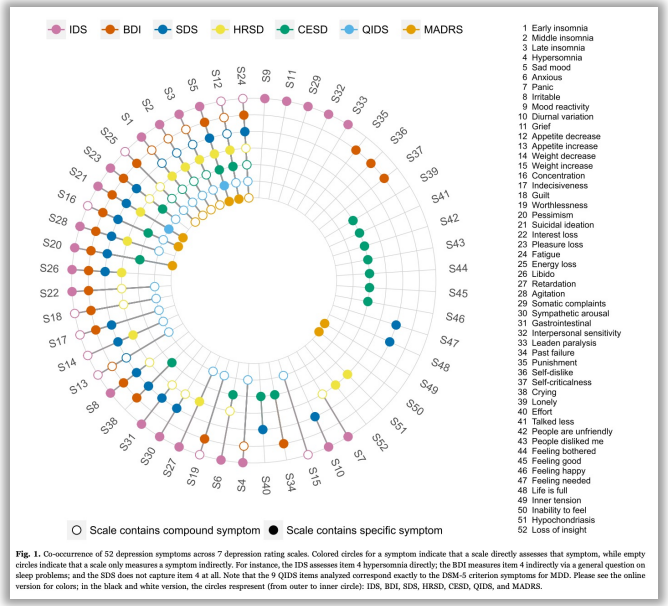
Effect Sizes

- We learned that meta-analysis is a technique which summarizes **quantitative outcomes** from several studies.
- In a primary study, it is usually quite easy to calculate **summary statistics** through which we can describe the data (e.g., the mean $\bar{X}$ or standard deviation S of a continuous outcome)
- This is possible because the outcome has been measured in the same way across all study subjects

→ This is typically **not the case in meta-analyses**, where studies use **different instruments** to measure the outcome



Fried, 2017

Effect Sizes

→ **We must find an effect size measure that can be summarized across all studies**

- Effect sizes are sometimes directly reported in the published article
- More often, we must calculate them from other reported summary information

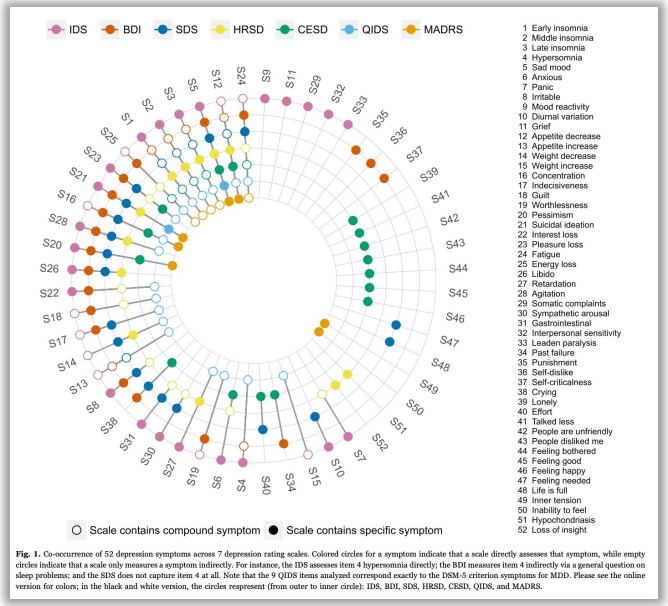


Fig. 1. Co-occurrence of 52 depression symptoms across 7 depression rating scales. Colored circles for a symptom indicate that a scale directly assesses that symptom, while empty circles indicate that a scale only measures a symptom indirectly. For instance, the IDS assesses item 4 hypersomnia directly; the BDI measures item 4 indirectly via a general question on sleep problems; and the SDS does not capture item 4 at all. Note that the 9 QIDS items analyzed correspond exactly to the DSM-5 criterion symptoms for MDD. Please see the online version for colors; in the black and white version, the circles respresent (from outer to inner circle): IDS, BDI, SDS, HRSD, CESD, QIDS, and MADRS.

Fried, 2017

Effect Sizes

Criteria for “good” effect size metrics:

- **Comparable**: Needs to have the same meaning across al studies
- **Computable**: Needs to be calculatable for (almost) all included studies
- **Reliable**: A valid estimator of the sampling variance (standard error) of the effect size must exist
- **Interpretable:**
  - The metric should be appropriate for our research question
  - Optimally, the effect size should have a straightforward clinical interpretation
  - There can be a trade-off between interpretability and other criteria (e.g., odds ratios versus NNTs)

Effect Sizes

What is an effect size?

General definition: ”a metric that quantifies the **relationship** between to entities in terms of **direction and magnitude**”

→ But this is only *one* way to define effect sizes!

→ Some only talk of effect sizes when referring to the results of intervention studies, expressed as differences between the treatment and control group (e.g., Cohen’s *d*)

→ Others disapprove of the term altogether, because the word “effect” in “effect size” suggests that there is a **causal** relationship; but this is not automatically the case