Five Ways to Determine Intervention Effectiveness

- Informal assessments
  - Open to subjective, expectancy effects
  - Not point of reference to others
- Raw score improvements
  - Pervasive in education
  - Progress monitoring in RTI/MTSS
  - Distance race analogy
  - Students can be progressing in raw scores while getting farther behind at the same time
- Statistically significant differences between groups
  - Included in intervention studies and commonly reported in abstracts, often without the word "statistical"
  - Experimental and control may both be effective and ineffective
- Effect sizes
  - Virtually required for all instruction/intervention studies
  - Due to pervasiveness in scientific literature it is the most potentially misleading index of effectiveness
  - Technically incapable of telling us if a program is effective
- Standard score point gains on nationally normed tests
  - Stand scores based on nationally stratified norm group can tell us if an intervention or teaching practice is effective
  - Strong inter-correlations among all the major word identification subtests from the leading achievement batteries suggests a fairly stable point of reference to determine improvement
  - Not useful for routine weekly or monthly progress monitoring

Conclusions About Determining Effectiveness

  - The use of nationally normed word identification tests and subtests from the major batteries is our "gold standard" for determining the effectiveness of interventions for word-level reading skills\
  - All other approaches are inherently incapable of determining effectiveness with that same degree of confidence