

Advice on Confidence Intervals

Many educational researchers use analytical conventions and tools that the field of psychology has endorsed. One such tool is the use of confidence intervals (CIs) in graph form. As Cumming and Finch (C&R) recently noted, "the APA [American Psychological Association] Publication Manual states that CIs 'are, in general, the best reporting strategy.'" C&R add that the APA Technical Task Force on Statistical Inference recommends that "in all figures, include graphical representations of interval estimates whenever possible." C&R recently completed a study that should help potential presenters of CIs to avoid several common pitfalls. In that study, C&R make the following points, among others:

1. "For a comparison of two independent means, $p \leq .05$ when the overlap of the 95% CIs is no more than about half the average margin of error, that is, when proportion overlap is about .50 or less...In addition, $p \leq .01$ when the two CIs do not overlap, that is, when proportion overlap is about 0 or there is a positive gap...These relationships are sufficiently accurate when both sample sizes are at least 10, and the margins of error do not differ by more than a factor of 2..." [p.176] [Note: This is a critical point. Under certain conditions, when CIs overlap, the analyst may still be able to infer that the two means differ at the .05 level. So comparing two CIs in a graph should not necessarily replace the actual significance test if the analyst wants to draw a conclusion about statistical significance.]
2. "For paired data, focus on and interpret the mean of the differences and the CI on this mean...Noting whether the CI on the mean of the differences captures 0 is a test of the null hypothesis of no difference between the means. The CIs for the two separate scores (e.g., pretest and posttest) are irrelevant for inferences about the mean of the differences. In general, beware of separate error bars for a repeated-measure IV: They are irrelevant for the inferences likely to be of interest..." [p. 177]
3. "SE bars [SE stands for standard error] are about half the size of 95% CI bars and give approximately a 68% CI, when n is at least 10..." [p.180] [Note: C&R warn us here that SE bars often appear in research papers, and readers should not confuse them with CI bars, especially since the tie to statistical inference differs between SE bars and CI bars.]

C&R provide a treasure of observations about confidence intervals (both the statistics and the graphs), as well as good arguments for their usage. Any researcher who works with studies that cite or provide statistical inference tests could benefit from reading this study. Geoff Cummings and Sue Finch (School of Psychological Science, La Trobe University, Melbourne, Victoria, Australia), present their study in an article ("Inference by Eye: Confidence Intervals and How to Read Pictures of Them") in the journal *American Psychologist* (the February-March 2005 issue or Vol.60, No.2, pp. 170-180). Their article includes seven figures, fifty references, and a handy "Abbreviated Statements of Rules of Eye [vs. Thumb] for Simple Figures Showing Means With Error Bars."

[Abstract done by Willard Hom, Director, Research & Planning Unit, System Office, California Community Colleges, 11/21/05]