An excess of barely-reliable Cronbach's values:

Preregistration of analysis of I/O dataset

Ian Hussey, Taym Alsalti, Malte Elson, Frank Bosco, & Ruben Arslan

# Background

Dataset 1 (psychology literature/APA journals) was analysed in a more exploratory fashion: we compared the fit of kernel smoothing and Beta regression, as well as different smoothing and bin width settings within the kernel smoothing. We decided against intercept-only Beta regression, because it demonstrated poor fit (much unmodelled variance around .85 to .95), and it was implausible that the data were generated by a single Beta distribution.

Dataset 2 (Bosco’s I/O literature dataset) will be explored using a more confirmatory approach, by preregistering the analysis code and parameters (e.g., kernel bandwidth = 0.01) used for dataset 1 and applying it to dataset 2.

Note that the datasets come from different literatures, which were possibly generated by communities with different research cultures and norms. As such, the analysis of dataset 2 is not a replication of the analysis of dataset 1 because we cannot assume that these datasets were sampled from the same population (i.e., communities with identical distributions of research practices that generated the alpha estimates being analysed). We therefore consider these tests to be independent tests of different literatures rather than a replication of the first analysis.

# Method

## Research question and hypotheses

We examine whether an excess of Cronbach’s estimates at rule-of-thumb cut-off values (i.e., .70, .80, .90) relative to non-cut-off values were reported in the literature. Based on the relative popularity of these rules-of-thumb, we had two related hypotheses: (1) that there would be an excess of .70 values; and (2) that there would be an excess of .70, .80, and .90 values.

## Code implementing the analyses

We provide a verbal description of the planned analysis below. We also preregister the R code that will be used to implement these analyses: see [osf.io/pe3t7](http://osf.io/pe3t7/), specifically “analysis/analysis\_IO.Rmd”. Should any discrepancies exist between the verbal description and the code, the code represents the authoritative preregistration. We describe the method in the past tense, with placeholders for results in square brackets highlighted in yellow, given that this text is likely to be used in the manuscript.

## Kernel smoothing

*[Relevant chunks in analysis.Rmd file: Kernel smoothing>Model and plot]*

We applied kernel smoothing to the extracted estimates in order to estimate their distribution and quantify the excess of values at the cut-off thresholds. The extracted estimates were rounded to two decimal places and converted to counts. We chose a data driven method with few assumptions to model their distribution: kernel density smoothing. Density was estimated at 99 equally spaced bins in the interval (i.e., from 0.01 to 0.99). We opted for the default options in R’s ‘density’ function: gaussian distributed kernels with a smoothing bandwidth set using Silverman’s rule of thumb (Silverman, 1986; i.e., the settings kernel = "gaussian" and bw = 0.01). As expected with large sample sizes (Sheather, 2004), the choice of kernel did not have a noticeable impact on the resulting density distribution. The bandwidth was chosen based on Silverman’s rule-of-thumb, which seemed to provide the best fit to the data as it yielded a relatively narrow bandwidth, which is appropriate for large sample sizes (Trosset, 2009, p. 172). These analytic choices were made in the first dataset (APA psychology journals) and then preregistered for the remaining analyses of the other datasets prior to having access to them. The observed counts and fitted smoothed curve can be seen in Figure XX (upper panel).

## Permutation tests on residuals

*[Relevant chunks in analysis.Rmd file: Kernel smoothing>Permutation tests and Kernel smoothing>Quantify excess]*

Residuals between the fitted model and the observed data were calculated for each of the values (i.e., from 0.01 to 0.99). This allowed us to calculate the excess or deficit of the observed count of each bin relative to its predicted value according to the smoothed curve.

Independence permutation tests were used to assess whether residuals at the cut-off values were larger than non-cut-off values at an above-chance level. These were implemented using the R package coin (REF). Two tests were run. The first test compared the .70 bin against all other bins. [A significant excess / No significant excess] of values of .70 relative to other values was found, *p* = XXXX. A second permutation test compared all three cut-off values against the non-cut-off values. [A significant excess / No significant excess] of values of .70, .80 and .90 relative to non-cutoff values was observed, *p* = XXXX. XX% more values of .70, XX% more values of .80, and XX% more values of .90 were found than predicted by the smoothed curve of observed values (see Figure XX, lower panel).