Q-STEP R 'HOW TO' GUIDES:

ROBUSTNESS CHECKS WITH SURVEY RESEARCH 1: CRONBACH ALPHA

Creator: Dr James Weinberg

In most social science disciplines, quantitative researchers will work with survey research to develop, build or test theories. In order for such survey research to be as rigorous as possible, it is important that we conduct robustness checks on our results in order to assess the validity of our theoretical claims. Therefore, this guide (the first in a series of 3) focuses on testing the Cronbach alpha (internal reliability) of questionnaire items.

Throughout this guide, you will be given a simple contextual description of Cronbach alpha and when/why it should be used, as well as an example worked through in Rstudio. This guide assumes a basic competency in R from the start - for example, users should already be comfortable with assigning and calling objects.

The example used in this guide is based on a dataset of people's Basic Human Values. Basic values are a personality characteristic that can be measured by psychometric surveys. In this instance, a 20 item questionnaire was administered to 107 people, with two items each tapping one of the ten values in the theory. These ten values can be clustered again into 4 further higher order values. This guide will use Cronbach alpha to assess the internal reliability of the data gathered to test this theory.

Cronbach Alpha

What is it and when to use it?

Cronbach alpha, α (or coefficient alpha), was developed in 1951 by Lee Cronbach. It measures reliability, or internal consistency. "Reliability" is how well a test (or a series of questionnaire items) measures what it should. For example, I have administered a survey on basic values to 107 members of the public. High reliability means that the survey measures the values prescribed in the theory accurately, while low reliability means it measures something else (or possibly nothing at all).

Above all, Cronbach alpha tests to see if multiple-question Likert scale surveys are reliable. These questions measure latent variables - hidden or unobservable variables like a person's basic values (or ideology, neurosis, the list of unobservable variables is endless). These are very difficult to measure in real life. Cronbach's alpha will tell you if the questionnaire items you have designed are accurately measuring the variable of interest.

WARNING: α is sensitive to the number of items in a test. A larger number of items can result in a larger α, and a smaller number of items in a smaller α. This can lead to misleading results with shorter survey measures.

Example:

Start by setting your working directory and reading the data file containing your questionnaire responses. To analyse Cronbach alpha, you will also need to install and load the package "psych".

It is likely that your survey contains a lot more items than you need for this analysis (i.e. socio-demographic data or another item battery). For example, in my dataset I have 165 variables but for the purpose of the current test, I am only interested in the 20 questions related to respondents' basic values. Therefore, you need to isolate these data as new objects in your global environment (top right panel in Rstudio). I am going to organise the items on my questionnaire that are supposed to tap each of the higher order values (Conservation, Openness, Self-Transcendence, and Self-Enhancement) into new data-frames.

Sheffield
Methods
Institute.

```
#install.packages("psych")

#library(psych)

Con <- data.frame(MP$ImpObe, MP$ImpRel, MP$ImpOrg, MP$ImpTra, MP$ImpBeh, MP$I
mpSta)

STran <- data.frame(MP$ImpCar, MP$ImpEqu, MP$ImpSup, MP$ImpPea)

Open <- data.frame(MP$ImpCur, MP$ImpAdv, MP$ImpFun, MP$ImpOri, MP$ImpNew, MP$
ImpEnj)

SEhan <- data.frame(MP$ImpAhe, MP$ImpLea, MP$ImpSuc, MP$ImpCha)
```

Once you have created new objects to group the items together in relevant (theory driven) sequences, then you can use the 'alpha' function in R to calculate their internal reliabilities. If you have any missing values in the dataset, remember to remove these prior to executing the alpha function or use the na.rm = TRUE command to tell R to remove them. Here is the code and output for the four items in my dataset that are supposed to measure Self-Enhancement values.

```
alpha(SEhan, na.rm = TRUE)


## Reliability analysis
## Call: alpha(x = SEhan, na.rm = TRUE)
##
##    raw_alpha std.alpha G6(smc) average_r S/N    ase mean  sd median_r
##        0.87      0.87    0.86      0.63 6.8 0.021  3.6 1.1     0.61
##
##  lower alpha upper      95% confidence boundaries
## 0.83 0.87 0.91
##
##  Reliability if an item is dropped:
##            raw_alpha std.alpha G6(smc) average_r S/N alpha se   var.r med.r
## MP.ImpAhe      0.83      0.83    0.78      0.62 5.0  0.029 0.0042  0.64
## MP.ImpLea      0.85      0.85    0.82      0.66 5.7  0.025 0.0152  0.64
## MP.ImpSuc      0.81      0.82    0.75      0.60 4.5  0.032 0.0052  0.57
## MP.ImpCha      0.84      0.84    0.80      0.64 5.3  0.026 0.0170  0.57
##
##  Item statistics
##             n raw.r std.r r.cor r.drop mean  sd
## MP.ImpAhe 106  0.87  0.85  0.81   0.74  3.7 1.5
## MP.ImpLea 106  0.82  0.82  0.74   0.68  4.0 1.2
## MP.ImpSuc 105  0.89  0.88  0.84   0.78  3.5 1.4
## MP.ImpCha 105  0.83  0.84  0.77   0.71  3.2 1.2
##
## Non missing response frequency for each item
##              1    2    3    4    5    6 miss
## MP.ImpAhe 0.05 0.23 0.20 0.21 0.20 0.12 0.01
## MP.ImpLea 0.01 0.11 0.21 0.30 0.25 0.11 0.01
## MP.ImpSuc 0.05 0.25 0.18 0.25 0.20 0.08 0.02
## MP.ImpCha 0.07 0.26 0.25 0.29 0.12 0.02 0.02
```

Sheffield
Methods
Institute.

What do the summary statistics mean?

• raw_alpha: Cronbach's α (values .7 or .8 indicate good reliability)
• std.alpha: this should be similar to raw_alpha (we only need the raw alpha though)
• G6: Guttman's lambda 6 (calculated from the squared multiple correlation or 'smc')
• average_r: average inter-item correlation (this is used to calculate std.alpha)
• mean: scale mean (the mean of the means of all individuals)
• sd: scale sd

How to interpret 'Reliability if an item is dropped'?

The overall α (raw_alpha) is .87. Each row refers to each item and has a raw alpha associated-this refers to the overall α when that particular item has been dropped/deleted. For example, the first row refers to the question item Qu.ImpAhe, which assesses how ambitious people are to succeed according to social standards. When this item is dropped, the overall α becomes .83. This reflects a worse reliability for the scale, so we know that keeping Qu.ImpAhe creates a more reliable measure of the sample population's Self-Enhancement values.

Essentially this section of the output allows us to check whether any of these raw alpha values are greater than the overall α if .87 if an item is dropped. Where this is the case, then we can assume that the item is redundant and actually hampers the scale.

The other columns of this table refer to how the other statistics will change if that particular item has been dropped/deleted.

How to interpret 'Item statistics'?

• raw.r: correlation between the item and the total score from the scale (i.e., item-total correlations); there is a problem with raw.r, that is, the item itself is included in the total-this means we're correlating the item with itself, so of course it will correlate;

• r.drop: item-total correlation without that item itself (i.e., item-rest correlation or corrected item-total correlation); low item-total correlations indicate that that item doesn't correlate well with the scale overall;

• r.cor: item-total correlation corrected for item overlap and scale reliability;

Sheffield Methods Institute.

- mean and sd: mean and sd of the scale if that item is dropped.

All items should correlate with the total score, so we're looking for items that don't correlate with the overall score from the scale. If r.drop values are less than about .3, it means that particular item doesn't correlate very well with the scale overall.

How to interpret the final frequency table?

This table tells us what percentage of people gave each response to each of the items (i.e., if you have a 6-point scale as in this example, then it tells you how many percent of responses were 1, 2, 3, 4, 5, or 6).

This helps you check the distribution of responses and whether everyone is giving the same responses (which will lead to low reliability).

Sheffield
Methods
Institute.