Why Factor Analyses with Negatively Keyed Items Fail

Kimberly A. Barchard
University of Nevada, Las Vegas

James A. Russell
Boston College

The Problem

Although test designers typically intend negatively keyed items to measure the same dimensions as positively keyed items, negatively keyed items often load on separate factors from positively keyed items or fail to load on any factor.

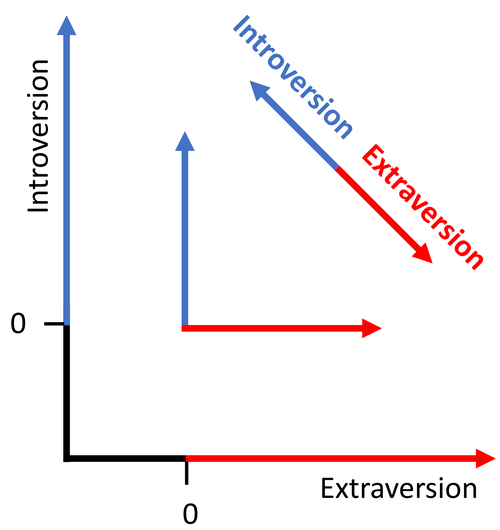
Positively and negatively keyed items form separate factors for two primary reasons. First, the disagreement end of item response scales are often ambiguous. It is clear what different levels of agreement mean. Someone who *strongly agrees* with the statement "I like cats" has greater affinity for felines than someone who *agrees*: Greater agreement indicates greater liking. However, it is not clear what different levels of disagreement mean. Perhaps one person disagrees with this item because they dislike cats. Another because they are allergic. Another because they feel neutral. People who disagree may not have much in common, and people who *strongly disagree* might not be more of any particular thing than people who simply *disagree*.

If two items are designed to measure the same construct and are keyed in opposite directions, agreement with one item should ideally be associated with disagreement on the other. However, people may disagree for a wide range of reasons. Thus, no part of the response scale on one item has a strong relationship with any part of the response scale on the other item, and the correlation is devastated.

Second, if two variables, X and Y, divide a dimension in half, they cannot correlate -1. In the top half of that dimension, X is positive and Y is 0. In the bottom half of that dimension, Y is positive and X is 0. The scatterplot is L-shaped (see Figure 1). If X and Y have a bivariate normal distribution, they will correlate -.467 (Carroll, 2000). Thus, two items that are keyed in the same direction can correlate +1, but two items that are keyed in opposite directions can only correlate -.467, even if both variables have perfect reliability. Because of this, factors tend to be composed of items that are all keyed in the same direction.

When including both positively and negatively keyed items in factor analyses, eliminate gradations of disagreement and use censored data analyses.

Figure 1
Opposites that divide a bipolar dimension in half



Solution Part 1: Eliminate Gradations of Disagreement

To eliminate ambiguity at the disagreement end of response scales, researchers must eliminate gradations of disagreement. There are at least three ways to do this. First, researchers can use interval-level variables such as count data (e.g., How many cats do you own?). Second, test designers can use response scales that provide gradations of agreement/accuracy without providing gradations of disagreement. For example,

Do you like cats at all? Yes or no? If yes, how much?
1 = a little, 2 = moderately, 3 = a lot, 4 = extremely

Third, test designers can specify the meaning of both ends of the response scales.

I like cats a lot	A like cats a little	I neither like nor dislike cats	I dislike cats a little	I dislike cats a lot

Solution Part 2: Censored Data Analysis

If researchers use unipolar items like the first two above, then the items only cover the high ends of the dimensions. Researchers need alternative statistical methods to model the relationships between the full underlying dimensions. An effective method of modeling these relationships is to use censored data analysis. When data points are *censored*, researchers have partial information about the values (Fox, 2016): They know that the value is at least as large as some value (e.g., the person's age is at least 55) or no larger than some value (e.g., the concentration is no more than the lower limit of detection of .001).

Many censored data analysis methods exist (e.g., Allignol & Laouche, 2020; Josse et al., 2020). However, most allow censoring on only one variable, and so cannot estimate correlations when both variables are censored and cannot be the foundation for factor analyses. Fortunately, Holst and Budtz-Jørgensen (2013) developed a method that allows censoring on both X and Y and implemented it in the R package *lava* (Holst, 2020), making it freely available. Recent research (Barchard & Russell, 2020b; Holst et al., 2015) shows the *lava* package provides accurate estimates as long as censoring is only moderate and multivariate normality is reasonable. By using strictly unipolar scales and analyzing the resulting data using censored data methods like the ones in the *lava* package, researchers will be better able to determine the relationships between positively and negatively keyed items, allowing them to measure the full breadth of psychological constructs.

The Details

http://barchard.faculty.unlv.edu/examining-opposites/ provides a presentation with additional figures and a detailed handout.