



Positively Packed Rating Scales

The traditional Likert-scale (1932) is a summated agreement response scale, balanced between disagreement and agreement. It uses a neutral mid-point and provides two options each for agreement and disagreement.

Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22, 5–55.

This is problematic in many cases because assuming people are not neutral there are only 2 choices to express the strength of their either positive or negative position--this reduces variance in the variable and makes it harder to estimate true degree of endorsement or rejection.

Furthermore, the neutral point is confounded with various alternative explanations such as 'I don't know', 'I don't care', 'I don't want to say', or even 'I don't understand'. Hence, the use of a mid-point is highly problematic.

Hence, positively-packed rating scales have been developed as a way of allowing people a way to show larger variation in their fundamentally positive perspective. In other words, if there are psychological or contextual reasons to believe people are biased to agree, positive packing increases variance in responding. Unpublished studies by John Hattie have shown that certain key terms tend to be space in intensity reasonably equally. These are 'slightly', 'moderately', 'mostly', and 'strongly'.

Naturally, if you expect people to be negatively biased, then packing negatively makes entirely good sense. However, in my experience in surveying students and teachers around educational phenomena, positive packing seems appropriate and valid.

A short reading list of papers on positive packing:

- Brown, G. T. L. (2004). Measuring attitude with positively packed self-report ratings: Comparison of agreement and frequency scales. *Psychological Reports*, 94(3), 1015-1024. doi:10.2466/pr0.94.3.1015-1024
- Hancock, G. R., & Klockars, A. J. (1991). The effect of scale manipulations on validity: Targetting



EDUCATIONAL MEASUREMENT, 23(2), 85-96. doi: 10.1111/j.1745-3984.1982.tb00294.x

- Lam, T. C. M., & Klockars, A. J. (1982). Anchor point effects on the equivalence of questionnaire items. *Journal of Educational Measurement*, 19(4), 317-322. doi:10.1111/j.1745-3984.1982.tb00137.x

In my own research I have found that

- positive packing increased mean scores,
- the response options are well-ordered according to Andrich's rating scale,
- in a completely balanced version of the packed-scale, the 4 negative options had thresholds that crossed over each other and had to be reduced to just 2 negative options.

My colleague Prof. Boaz Shulruf (UNSW) and I have written a chapter on rating scale design in survey research. It will appear later in 2023. Brown, G. T. L., & Shulruf, B. (2023). Response Option Design in Surveys. In L. R. Ford & T. A. Scandura (Eds.), *The SAGE Handbook of Survey Development and Application*. Sage. In this we argue that response scales should not have a neutral mid-point, should be between 5 and 7 options, and that positively or negatively packed options work really well, when you know in which direction participants are likely to be biased.

Alternative to Summated Rating Scale

Needless to say, the summated rating scale (whether balanced or packed) is not the only way to ascertain the intensity and direction of attitudes. A popular option is the Rasch rating scale. My own view is that most people do not construct rating scale items to take advantage of the features of the Rasch continuum. In essence, what is needed is to write items (not response options) that capture different places on the continuum from very negative to very positive. While this can be similar to the 'ideal point' approach, it differs because the Rasch approach still presumes that agreeing with an option at the middle of the scale also means you agree with the items lower in the scale. Prof. Larry Ludlow has done some prize-winning work in constructing ordered items using Rasch scale analysis:

Ludlow, L. H., Matz-Costa, C., Johnson, C., Brown, M., Besen, E., & James, J. B. (2014). Measuring engagement in later life activities: Rasch-based scenario scales for work, caregiving, informal helping, and volunteering. *Measurement and Evaluation in Counseling and Development*, 47(2), 127-149. doi:10.1177/0748175614522273

