# Free Text Factor Analysis

Nigel Guenole

Variables in conventional psychometric factor analysis are homogenous and ordered item responses. Free text is instead heterogeneous and may or may not be ordered. Free text factor analysis needs a harmonization layer to scale text along common continua. Using a simulation measuring *leader absorptive capacity*[1] I show how convex hulls might provide such a layer.

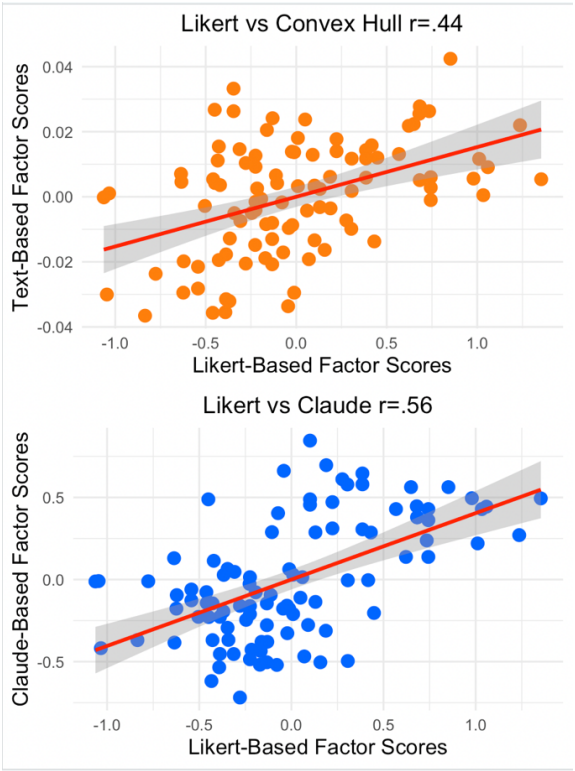
Factor analysis is currently used with text[2]. Pseudo-factor analysis for item analysis uses embedding vectors to approximate real responses, but it is unclear, perhaps unlikely, that any single encoding of free text responses across respondents could function similarly due to response heterogeneity in free text.

PCA can be used to directly factor heterogeneous response embeddings but for dimensionality reduction rather than interpretable measurement models. Components are linear combinations of embedding dimensions or whole embeddings with no easy interpretation and where different topics might be discussed even in the same contexts.

This test of the ability of convex hulls to provide a harmonization layer shows the approach comes in marginally behind direct LLM scoring in terms of person score recovery. However, convex hulls are at a disadvantage because the LLM generating and scoring are the same method, which is not true for convex hulls. Convex hulls also offer advantages that LLMs do not.

**Method.** Simulate continuous factor scores with Lavaan. Specify thresholds, discretize. Model fit is perfect $\chi^2$= .42, *df*=2, *p*=.89. Create a prompt instructing GPT Turbo to generate responses to match Likert levels across subscales: *assimilation*; *acquisition*; *transformation*; *exploitation*. Use a second LLM, Sonnet 3.5, to score responses. Claude score fit is perfect $\chi^2$= 1.095, *df*=2, *p*=.58. Likert correlations range from .67 to .81. We embed responses, reduce dimensionality, form a convex hull, calculate centroid distances, set referent, calculate observed-referent differences, calculate dot product between differences and referent, multiply centroid distance by dot product. Model fit is perfect $\chi^2$= .91, *df*=2, *p*=.63. Likert correlation absolute range is.14 to .47. Likert with convex hull factor score *r* is .44 and .56 with Claude scores.

**Fig 1. Correlations with Likert scales.**




**Conclusion.** LLM scoring led to better recovery of scores but the results were close and the design favored LLMs. There are also many convex hull configurations that need cataloguing and testing, this is one replication of one configuration. Convex hulls are more complex than LLMs but make fewer assumptions about what is measured and are potentially more interpretable, transparent, and replicable and do not require API calls. The methods can be complementary.

[1] Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, *35*, 128-152.

[2] Topic modeling, a closed dictionary approach, does not yield measurement properties comparable to factor analysis in the author's experience.