Operant Subjectivity: The International Journal of Q Methodology
38/3-4 (2016): 29-36, DOI: 10.15133/j.os.2016.007

Operant Subjectivity
The International Journal of Q Methodology

An Overview of the Statistical Techniques in Q Methodology: Is There a Better Way of Doing Q Analysis?[1]

Noori Akhtar-Danesh
McMaster University, Canada

Abstract: This article critically reviews the statistical techniques used in Q methodology, such as methods of factor extraction and factor rotation. Different techniques of factor extraction including principal component analysis, principal axis factoring, centroid factor analysis and maximum likelihood method will be reviewed. Specifically, centroid factor analysis will be compared against other methods, and some common beliefs regarding this technique, including the concept of factor indeterminacy, will be re-examined. In addition, different techniques of factor rotation will be reviewed. In particular, the pros and cons of judgmental (manual) factor rotation will be discussed. Furthermore, limitations of current statistical packages available for Q methodology, PQMethod and PCQ, are highlighted. Finally, the article concludes with a few suggestions for further improvements of current or future Q programs.

Keywords: centroid factor analysis, indeterminancy, manual (theoretical) rotation, principal components analysis, Q methodology, varimax rotation

Introduction

Q methodology is a research tool that constitutes qualitative and quantitative techniques for data collection, summarization and analysis, with its raw material consisting of subjective viewpoints. Introduced by William Stephenson (Stephenson, 1935a; Stephenson, 1935b), it is used to identify unique as well as commonly shared viewpoints among a group of participants. It is a valuable research method in exploring human perceptions and interpersonal relationships (Dennis, 1986). This methodology uncovers different patterns of thought, rather than their numerical distribution in the population of interest. Its primary objective is to identify a typology, not to test the typology's proportional distribution within the larger population (Brown, 1993).

The quantitative component of Q methodology includes use of a by-person factor analysis in which each extracted factor contains similar Q sorts. The similarity between Q sorts is usually assessed by a Pearson correlation coefficient.

Currently, the two most common programs for analysis in Q methodology are PQMethod (Schmolck, 2014) and PCQ (Stricklin, 1996). These programs include only two factor extraction techniques: centroid factor analysis (CFA) and principal

[1] This is the lead article in a special section of *Operant Subjectivity* devoted to methods of factor extraction and rotation in Q methodology. It is followed by two responses and concludes with a rejoinder from the author. [*Ed.*]

Contact author: daneshn@mcmaster.ca

component analysis (PCA). In addition, there are only two factor rotation techniques implemented in these programs: varimax rotation and manual (judgmental) rotation.

Although Q methodology was introduced in 1935, its use has substantially increased recently because of the above mentioned programs. Perhaps it could have become more common if there had been more techniques for factor extraction and factor rotation in the Q analysis programs, or if there had been suitable codes available in the general and sophisticated statistical programs such as SPSS, R, Stata and SAS.

Another important consideration is that Q methodology users have been using some old methods of analysis, for instance CFA, originating from the inception of Q methodology as suggested by William Stevenson. Seemingly, the first generation of Q users, that is, colleagues and students of William Stephenson, have been too loyal to CFA and judgmental rotation and have resisted any advancement or changes in these methods.

This article critically reviews the statistical techniques used in Q methodology, such as methods of factor extraction and methods of factor rotation. It also reviews factor extraction and factor rotation techniques that are not currently used in Q methodology but might be appropriate to use. Although some of these issues have already been discussed in the literature, this review may be timely and useful to the growing number of Q users.

Factor analysis

Factor analysis is a statistical method used for data reduction. It originated from Charles Spearman (1904) and is based on the correlation matrix of a set of variables. Factor analysis classifies the correlated variables together so that in each group, called a factor, variables are highly correlated with each other but have no correlation or weak correlation with variables in other groups. As a result, a large number of variables are usually grouped into a handful of factors. Techniques of factor analysis can be classified into two major groups: exploratory factor analysis and confirmatory factor analysis (Tabachnick & Fidell, 2001). Exploratory factor analysis is typically used at the early stages of research for data classification and hypothesis generation purposes (Tabachnick & Fidell, 2001). On the other hand, confirmatory factor analysis is used to examine whether variable classifications follow a certain model. The methods discussed in this article fall in the exploratory factor analysis category.

As noted, Q methodology uses a by-person factor analysis. Statistically, the only difference between factor analysis and by-person factor analysis is that factor analysis uses an individual's characteristics as variables, whereas in the latter, Q sorts are used as variables. The following sections discuss some major techniques of factor extraction and factor rotation.

Factor extraction methods

There are different methods of factor extraction. Furthermore, an infinite number of solutions exist for each set of extracted factors, leading to issues of factor rotation (Manly, 2005), which will be discussed in the next section. General statistical programs such as SPSS, R, Stata and SAS embrace various methods for factor extraction including principal component analysis (PCA), principal axis factoring (PAF), maximum likelihood factoring, image factoring, alpha factoring, unweighted and generalized (weighted) least square factoring. Each of these methods calculates a set of orthogonal factors that in general reproduce the correlation matrix between variables (Tabachnick & Fidell, 2001). Although these methods differ in solutions, such as maximizing the variance or minimizing residual correlations, the differences in results are usually negligible with large samples. Moreover, these methods may not provide interpretable results (factors) per se; thus, it is a common practice to rotate factors after extraction.

As mentioned earlier, only two methods of factor extraction are currently implemented in the common Q-method programs such as PQMethod (Schmolck, 2014) and PCQ (Stricklin, 1996)—PCA and CFA. This section reviews PCA, PAF, maximum likelihood (ML), and CFA and some of their statistical properties.

Principal component analysis

PCA is a factor extraction method available in almost every statistical program and is probably used more than any other method in statistical analysis. It is also repeatedly used in Q methodology. It extracts uncorrelated linear combinations of the observed variables (Q sorts in Q methodology). This method analyzes all the variance in the Q sorts; that is, it uses 1's in the diagonal of the correlation matrix for factor extraction. When using PCA, the goal is to explain the maximum variance for each factor from the dataset. As a result, the first factor extracts the most variance from the dataset and the second factor extracts the most variance from the remaining variability among the dataset. This process continues until 100% of the variance is explained by the factors. However, in practice a small number of the extracted factors satisfying some pre-specified criteria are used in the subsequent analyses. Historically, the two commonly used criteria for selecting the number of factors are (1) choosing a set of factors to represent a predetermined proportion, say 70%, of the total variance, and (2) choosing those factors with eigenvalues greater than the average eigenvalue, which is 1 when using a matrix of correlation coefficients. However, it is a common practice in Q methodology to extract factors and then maintain any factor with at least two to three defining sorts after rotation. In addition, when using PQMethod, only a maximum of eight factors can be extracted, which is usually more than the actual number of factors needed and used in Q methodology.

Principal axis factoring

Principal axis factoring is similar to PCA with the only difference being that in the correlation matrix, 1's in the diagonal are replaced with estimates of the communalities. For each Q sort the communality is the proportion of the variance that is shared (or explained) with the other Q sorts. These communalities are usually estimated through an iterative process with the squared multiple correlation of each variable with all other variables as the starting values (Tabachnick & Fidell, 2001). Iteration continues until the changes in the communalities satisfy the convergence criterion for extraction. Principal axis factoring is generally considered best for exploring underlying factors for theoretical purposes (Pett, Lackey & Sullivan, 2003). Principal axis factoring will yield results similar to PCA in either of the following cases (Rencher, 2002): (1) the correlations are fairly large, with a resulting small number of factors, or (2) the number of variables (Q sorts) is large.

Centroid factor extraction

A full description of CFA can be found in Thurstone (1947) and Brown (1980). It is known as an approximation for PAF (Holzinger, 1946; Mulaik, 2009b). The main difference between PAF and CFA is that in PAF, as in PCA, the sum of squares of "loadings" is maximized, whereas in CFA the sum (or average) of the "loadings" is maximized. Geometrically, PAF provides a set of orthogonal factors, but factors extracted using CFA need not be orthogonal (Holzinger, 1946). Indeed, CFA is the only method which extracts nonorthogonal factors. CFA, however, is not implemented in any major statistical program, although it is regarded as the best method of factor extraction by many proponents of Q methodology and is implemented in PQMethod and PCQ.

Maximum likelihood factor extraction

Maximum likelihood (ML) factor extraction is another method in which, like PAF, the correlation replaces matrix 1's in the diagonal with the estimates of communalities. Maximum likelihood is based on the assumption that each variable (Q sort in Q methodology) is distributed normally. Because each Q sort is approximately normally distributed, this might be technically an appropriate method for use in Q methodology. The problem associated with this technique is that it takes longer than PCA and PAF to run and may result in a Heywood case (Pett et al., 2003). A Heywood case can occur in all common factor analysis solutions, including PAF, CAF and ML, where the iteratively estimated communality becomes unity or exceeds unity. In other words, the variance of some factors becomes zero or negative (Heywood, 1931).

A comparison between CFA and the other factor extraction methods

Among the above-mentioned methods, CFA is the only method that can extract nonorthogonal factors, although it may not be known to many Q users. Indeed, PQMethod and PCA programs always provide seemingly orthogonal factors for rotation, regardless of whether the original factors are orthogonal. However, the degree of non-orthogonality is usually small and negligible.

As a result, given that CFA is an approximation of PAF, one might ask why Q methodologists are still so fond of CFA and don't use PAF instead. Indeed, this is the question that is repeatedly asked by new users of Q methodology. Perhaps the main reason for using CFA in the early days of Q methodology was its simplicity in calculation over other methods such as PAF and PCA (Holzinger, 1946; Mulaik, 2009b). Use of CFA among Q methodologists is still common, despite current powerful computers and advanced statistical programs. The first reason for not using PAF in Q analysis may be that it is not implemented in Q method programs such as PQMethod and PCQ. Second, and perhaps the most important reason given in the literature, is the concept of indeterminacy. As Brown explains (1980, pp. 32-33):

> Factor analysts are almost universal in their use of objective rotational procedures, usually varimax or quartimax, but in Q methodology it is often worthwhile to rotate factors judgmentally in keeping with theoretical, as opposed to mathematical, criteria…In these respects, the centroid method, long ago discarded for its indeterminacy, still has its usefulness, for the uniqueness of the centroid method *is* its indeterminacy: There is no correct solution out of the infinite number of solutions [meaning rotations] available, so the investigator is free to pursue his own inclinations, guided by his theory.

However, the truth of the matter is that regardless of what technique we use for factor extraction, there are an infinite number of ways that the factors can be rotated subjectively (Manly, 2005; Rencher, 2002). For instance, PQMethod offers judgmental (manual) rotation for both PCA and CFA. So, indeterminacy, as implied in the previous quote, is no longer unique to CFA, although it may have been in the early days of Q methodology prior to the availability of these statistical programs. Third, historically and perhaps due to the belief in the second reason, use of CFA seems to be a sensitive issue to Q methodologists. The author posted an inquiry on the Q mailing list in 2007 on the reason for using centroid method and a summary of the answers, mainly from the senior members of the list, are presented here:

> It is true that historically centroid factor extraction was used to approximate other methods. Today it is used…because...it is a better theoretical fit for Q method than other factor solutions.
>
> We need to be reminded that the factor analyst is an artist. Interpretations of factor analysis are objective in that they are data bound and can be clearly wrong, but they are also subjective in that it is possible to bring to bear insights that others might miss.
>
> Historically centroid extraction was used as an approximation of PAF, but today in Q studies, PAF is often being used as an approximation of centroid extraction.

These statements clearly show a great deal of sympathy for using CFA, although no clear statistical or theoretical reason is provided. Now that indeterminacy has been put forth as the main reason for using CFA, a review of the topic is provided.

What is indeterminacy?

Indeterminacy has been a long-standing and contentious issue in factor analysis with little agreement on its concept and definition. *The Journal of Multivariate Behavioral Research* devoted a whole issue to this topic in 1996 (vol. 31, issue 4). In summary, the following three definitions/concepts of indeterminacy can be found in the literature. However, none of these definitions exclusively pertains to CFA.

1. For each factor solution, there is an infinite number of rotations (Brown, 1980; Reeve & Blacksmith, 2009). This seems to be the concept of indeterminacy among Q methodologists.
2. In any common-factor analysis (including PAF and CFA), the squared multiple correlation coefficient, ρ^2, between each factor and the observed variables, Q sorts, is less than unity. This means that the common-factor consists of two components, one component that can be estimated from the observed variables and one component that cannot be estimated from the observed variables. Therefore, the common-factor cannot be totally determined from the observed variables (Mulaik, 2009a).
3. No matter what method one uses for factor extraction (and factor rotation), there is an infinite number of ways by which the factor scores can be calculated given the same factor loadings (Grice, 2001; Wilson, 1928).

Methods of factor rotation

Usually, in the original set of unrotated factors, most variables load highly on the first factor or each factor constitutes more than a handful of variables. As a result, unrotated factors are typically not meaningful or easily interpretable. Factor rotation, as originally explained by Thurstone (1947), is a process in which the original factors are rotated about their origin for a simple structure and more interpretability.

Let's assume that $F_1, F_2, \ldots, F_m$ constitute the original set of orthogonal factors and $F_1^*, F^*_2, \ldots, F^*_m$ are a set of linear and uncorrelated combinations of $F_1, F_2, \ldots, F_m$ so that

$$\begin{cases} F_1^* = d_{11}F_1 + d_{12}F_2 + \ldots + d_{1m}F_m \\ F_2^* = d_{21}F_1 + d_{22}F_2 + \ldots + d_{2m}F_m \\ . \\ F_m^* = d_{m1}F_1 + d_{m2}F_2 + \ldots + d_{mm}F_m \end{cases}.$$

Then, $F_1^*, F^*_2, \ldots, F^*_m$ constitute a rotation of the original factors and explain the data as good as the original factors (Manly, 2005; Pett et al., 2003).

As there are different methods of factor extraction, there are different methods of factor rotation as well. Tabachnick and Fidell (2001) listed ten methods of factor rotation, of which five methods (varimax, quartimax, equamax, promax and direct oblimin) are implemented in SPSS. The first three are orthogonal and the last two are oblique.

Varimax

The most common rotation method, varimax is an orthogonal rotation technique that minimizes the number of variables with high loadings, either positive or negative, for each factor (Tabachnick & Fidell, 2001). In other words, it maximizes the variance of each factor loading by making high loadings higher and low loadings lower to simplify factor interpretation.

Quartimax

Quartimax is an orthogonal method that minimizes the number of factors to explain each variable. In other words, each variable (Q sort) is loaded on the minimum number of factors. This is similar to the common practice in Q methodology in which each Q sort is preferred to be loaded only on one factor. Quartimax rotation also tends to allow for a general factor which usually consists of a larger number of variables (Q sorts) compared to the other factors. This method simplifies the interpretation of the observed variables. Indeed, a quartimax rotation does the same thing for variables that varimax does for factors.

Equamax

Equamax is an orthogonal rotation method that combines varimax and quartimax. This method minimizes both the number of variables that load highly on a factor and the number of factors needed to explain a variable.

Direct oblimin

Direct oblimin is an oblique (nonorthogonal) rotation method that minimizes the cross products of loadings to simplify factors. This method permits fairly high correlation between factors, although factors may not necessarily correlate when this method is used.

Promax

Promax is also an oblique rotation, which allows factors to be correlated. This rotation can be calculated more quickly than a direct oblimin rotation, so it is useful for large datasets.

Manual (judgmental or theoretical) rotation (MR)

Manual rotation is a technique only available in Q programs such PQMethod and PCQ. It is an orthogonal rotation method and mostly leads to subjective rather than objective results. This technique allows researchers to rotate the factors on any direction and at any size around the origin. Therefore, it can easily be used to rotate based on some predetermined theoretical context. The danger, however, is to rotate the factors until some convincing solution is found, leading to a subjective data-driven solution. Although some researchers may relate this data-driven process to abductive reasoning (Brown & Robyn, 2004), the connection between MR and abductive reasoning is not clear and needs further discussion. Manual rotation is basically a process in which the researcher rotates the factors to certain degrees and examines the results based on that rotation. This process continues until some convincing result is found. The main problem with this approach is that it can be very subjective as each researcher may find some specific results interesting and convincing. Although each Q sort is a subjective viewpoint from one participant, the analysis of Q sorts needs not to be subjective. Otherwise, there will be no reliability in the findings and the results will hardly be reproducible given the same set of Q sorts. This is obviously contrary to the principle of scientific inquiry that states study findings should be reproducible (Houser, 2008).

Similar to the methods of factor extraction, different methods of rotation tend to give similar results if the pattern of correlations in the dataset is fairly clear or if there is a stable solution for the dataset (Tabachnick & Fidell, 2001). Of these methods, only varimax and manual rotations are implemented in PQMethod and PCQ. In addition, there may be a need for implementing an oblique rotation technique in current Q programs as the authentic factors, if they exist, could hardly be orthogonal. However, calculating factor scores for Q sorts loaded on an oblique factor seems to be more challenging compared to an orthogonal factor.

Discussion

This article revisits some basic statistical components in Q methodology, in particular, factor extraction and factor rotation techniques. Limitations of the current statistical packages available for Q methodology—PQMethod and PCQ—are pointed out. Perhaps the biggest limitation is that PAF and ML, which are theoretically sound methods for factor extraction, are not implemented in the available Q-method programs. Furthermore, use of CFA for the reason of indeterminacy is not justified and no matter how it is defined, is not unique to CFA. This article also concludes that the common use of manual rotation is not scientifically sound as it can easily result in an unreliable and invalid solution.

Some suggestions for future programs

For any further developments in current or future Q programs, inclusion of PAF and ML for factor extraction will be essential as these methods are theoretically more appropriate. In addition, including quartimax for factor rotation is more consistent with the common practice in Q methodology. Furthermore, the available programs do not allow for the oblique or non-orthogonal rotations which may fit better to some datasets. Last but not least, current Q programs allow each Q sort to load significantly on only one factor, which seems to be a limiting step in analysis. If there are some Q sorts that load on more than one factor, the programs should allow for it to happen as this might enrich the interpretation of factors.

Finally, there have been some recent endeavors to improve the statistical analysis of Q methodology by including quartimax and promax rotation techniques in the new *qmethod* program (Zabala, 2014; Zabala & Held, 2015); however, the practicality and acceptability of this program remain to be evaluated by the Q researchers.

References

Brown, S.R. (1980). *Political subjectivity: Application of Q methodology in political science*. New Haven: Yale University Press.

Brown, S.R. (1993). A primer on Q methodology. *Operant Subjectivity, 16,* 91-138.
Brown, S.R., & Robyn, R. (2004). Reserving a key place for reality: Philosophical foundations of theoretical rotation. *Operant Subjectivity, 27,* 104-124.
Dennis, K.E. (1986). Q-Methodology: Relevance and application to nursing research. *Advances in Nursing Science, 8,* 6-17.
Grice, J.W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6,* 430-450.
Heywood, H.B. (1931). On finite sequence of real numbers. *Proceeings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character, 134,* 486-501.
Holzinger, K.J. (1946). A comparison of the principal axis and centroid factors. *Journal of Educational Psychology, 36,* 449-472.
Houser, J. (2008). Precision, reliability, and validity: Essential elements of measurement in nursing research. *Journal of Special Pediatric Nursing, 13,* 297-299.
Manly, B.F.J. (2005). *Multivariate statistical methods: A primer* (3rd ed.). Boca Ration, FL: Chapman & Hall/CRC.
Mulaik, S.A. (2009a). Factor scores and factor indeterminacy. In *Foundations of factor analysis* (2nd ed.) (pp. 369-404). Boca Ration, FL: Chapman & Hall/CRC.
Mulaik, S.A. (2009b). Methods of factor extraction. In *Foundations of Factor Analysis* (2nd ed.) (pp. 139-165). Boca Ration, FL: Chapman & Hall/CRC.
Pett, M.A., Lackey, N.R., & Sullivan, J.J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.
Reeve, C.L., & Blacksmith, N. (2009). A review of current factor analytic practices in the science of mental abilities. *Intelligence, 37,* 487-494.
Rencher, A.C. (2002). *Methods for multivariate analysis* (2nd ed.). Hoboken, MA.: Wiley Interscience.
Schmolck, P. (2014). PQMethod (Version 2.35, adapted from mainframe-program Qmethod written by John Atkinson, 1992) (Computer Software). Neubiberg: University of the Bundeswehr Munich. Retrieved from http://schmolck.org/qmethod/
Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology, 15,* 201-293.
Stephenson, W. (1935a). Correlating persons instead of tests. *Character & Personality, 4,* 17-24.
Stephenson, W. (1935b). Technique of factor analysis. *Nature, 136,* 297.
Stricklin, M. (1996). PCQ: Factor analysis program for Q-technique (Version 3.8) (Computer software). Lincoln, NE: Stricklin, M.
Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.
Thurstone, L.L. (1947). *Multiple factor analysis*. Chicago: University of Chicago.
Wilson, E. (1928). On hierarchical correlation systems. *Proceedings of the National Academy of Science, 14,* 283-291.
Zabala, R. (2014). qmethod: A package to explore human perspectives using Q-methodology. *The R Journal, 6,* 163-173.
Zabala, R., & Held, M. (2015, May 5). Package 'qmethod' (Reference manual). Retrieved from https://cran.r-project.org/web/packages/qmethod/qmethod.pdf