USING PERSON FACTORS IN SOCIAL SCIENCE RESEARCH

Gary K. **Burger** *

Resumen

El análisis factorial es el método utilizado frecuentemente para la identificación de las dimensiones y estructuras que constituyen la base de un conjunto de medidas, lo cual es importante para la investigación.

Mientras el *análisis factorial de tipo R*, produce los factores presentes en las variables, es conocido por muchos investigadores que el *análisis de tipo Q* describe los factores presentes en las personas, y ha sido utilizado con menor frecuencia. En el presente trabajo se describe el *análisis factorial de tipo Q* y lo distingue del análisis factorial de tipo *R*. Entonces, se examinan tres usos de factores derivados del análisis factorial tipo *Q*: para describir el perfil de los resultados de pruebas de individuos, para dar más opciones al análisis convencional de los datos y para investigar las cualidades del individuo en los instrumentos de medición. Se propone que los *factores personales* resultan útiles para estos propósitos, en las investigaciones de las ciencias sociales.

Palabras clave: Análisis factorial de tipo *Q* - factores personales - perfiles.

* PhD in Psychology. Professor and Chair of the Department at Psychology at the University of Missouri-St. Louis. United States. E-Mail: GBurger@UMSL.EDU

Abstract

The identification of dimensions and structures underlying a set of measures is often of importance in research, and the method of factor analysis is frequently used for that purpose. While *R*-type factor analysis, which yields the factors present in the variables, is well known to many investigators, *Q-type factor analysis*, which describes the factors present in the persons, has been employed less frequently. In this paper, *Q*-type factor analysis is described and distinguished from *R*-type analysis. Then, three uses of *person factors* (derived from *Q*-tipe analysis) in research setting are discussed, to describe the score *profiles* of individuals, allow more options in conventional data analysis, and to investigate the properties of individual items on measuring instruments. It is argued that person factors can be used profitably for all of these purposes in research.

Key words: *Q*-type factor analysis - person factors - profiles.

Often in the social sciences, indeed in many sciences, it is of interest to investigate the structure of a particular domain by analyzing relevant measures from that domain. For instance, Leumann and Knocke (1987) studied the structure of policy domains, and D'andrade, Quinn, Nerlove, and Romney (1972) studied the structure of meanings regarding disease categories among English and Spanish-speaking students. Isolation of an underlying structure contributes to our theoretical understanding, simplifies the analytical task (by employing a smaller number of dimensions rather than a large number of measures), and also may lead to new ways of conceptualizing strategies for data analysis.

A wide variety of quantitatively based procedures are available to identify such structures, including factor analysis, cluster analysis, and multidimensional scaling. The starting point in most such analyses usually involves obtaining measures of a number of attributes on a number of entities. So, for instance, Filsinger, Faulkner, and Warland (1979) administered

37 items from a religiosity scale to college students, and identified seven different types of religious individuals (e.g., orthodox, culturally religious, moderately religious). This paper will focus primarily on factor analytic approaches to identifying structure.

A fundamental decision to be made when analyzing such a data matrix is whether one searches for relationships among the attributes (columns of the data matrix) or entities (rows of the data matrix). Factor analysis has most frequently been used to study relationships among the attributes. This use of factor analysis is called *R-type*, and is well known in the social sciences. Examples exist in many areas including intelligence (Guilford, 1967; Thurstone, 1938) and personality (Cattell, Eber, & Tatsuoka, 1970; Costa, & McCrae, 1992).

When factor analysis is used to identify relationships among the entities, it is called *Q-type*. *Q*-type factor analysis has been employed with less frequency than *R*-type, and is generally not as well known as *R-type* among social science researchers. The purpose of this paper is to briefly describe the main features of *Q*-type factor analysis and to show how it can be used in social science research.

Q-type factor analysis is of use in situations where the profile of the entity (the pattern of scores that an individual obtains over a relevant set of measures) is of interest rather than focusing upon the analysis or description of a single variable at a time. For instance, Sternberg, and Lubart (1995) have suggested that what distinguishes the highly creative individual from the only modestly creative one is the confluence of multiple factors rather than extremely high levels of any particular factor or even the possession of a distinctive trait. Ackerman, and Heggestad's (1997) theory of intelligence stresses intelligence as a Process, Personality, Interests, and Knowledge (PPIK), and has identified across attribute (ability, interest, personality) trait complexes and related them to various measures. In the area of personality. Magnusson, and Torestad (1993) have stated that in their person-oriented approach, problems for investigation are formulated in terms of individual patterns of values for variables relevant to the problem under investigation. What characterizes the above theories is that it is the entire pattern of measures (the configuration of the scores of individuals on relevant variables) which is of importance.

Often, the entities in *Q*-type factor analysis are persons or individuals, and it is from this perspective that the method will be described. However, entities could be different countries, societies, families, or any other unit over which measurements are obtained. The key concept is that the major purpose of *Q*-type factor analysis is to identify modal profiles or patterns of

responses of the individuals (entities) over the attributes or measures. Presumably (or hopefully), there will exist a relatively small number of these profiles (patterns of responses) that will characterize the majority of the individuals (entities).

The most familiar type of factor analysis, *R*-type, analyzes the correlations among variables (over individuals) to identify clusters of variables called factors. These factors are regarded as dimensions underlying the variables. Ordinarily, there are substantially fewer factors than variables, so the analysis contributes, among other things, a certain parsimony to the situation. The factor matrix, in *R*-type analysis, gives the degree of relationship of the original variables to the underlying factors or dimensions. In *Q*-type factor analysis, the correlations among individuals (over the variables) are analyzed to identify clusters of individuals, *Q*-type factors. These clusters of individuals represent *idealized individuals* or *hypothetical individuals* the profiles of the major subgroups of people in the study. Rimoldi, and Libonatti (2000) have described the relationships between these *person* factors and *variable* factors, suggesting some promising applications of these relationships to the interpretation of test scores, and Burger, and Rimoldi (1997) have applied some of these ideas to the area of personality tests. For this paper, I will employ the term modal profiles (Skinner, Reed, & Jackson, 1976) to describe the profile (over the attributes) of the idealized individuals or hypothetical individuals isolated by *Q*-type factor analysis. Individuals within a particular *Q*-type factor (cluster of individuals) will have similar profiles. The profile, of course, is the pattern of scores a particular person has on the variables. To illustrate this process, Table 1 presents a hypothetical matrix of correlations among six persons (over four variables). As can be seen by inspection, the profiles of persons 1, 2, and 3 are highly correlated, as are the profiles of persons 4, 5, and 6. The intercorrelations between these two clusters are low.

Table 2 presents a hypothetical *Q*-type factor matrix which might describe the correlations in Table 1. Persons 1, 2, and 3 have high loadings on Factor I, while individuals 4, 5, and 6 have high loadings on Factor II.

Finally Table 3 presents the profiles of the *Q* factors (or *idealized individuals* or *modal profiles*). These profiles represent the different patterns or responding of the two *Q*-type factors (or clusters) of individuals over the four variables. Any factoring method that is used in *R*-type factor analysis may be employed in *Q*-type factor analysis. One simply analyzes the correlations among the persons (rows) of the data matrix rather than the variables (columns). Typically this is accomplished in statistical packages (e.g., SPSS) by transposing the data matrix before factoring. It is also possible to

analyze the covariances or average raw score cross products among the persons.

How can these *Q*-type factors (modal profiles), when identified, be employed in research? First, the *Q*-type factors simplify the description of individual profiles, making it easier to compare profiles and measure change of profiles over time. Secondly, the correspondence of an individual's profile to a modal profile can be used as a categorical variable which captures the configural aspects of the profile. The resulting *type* variable can be used in various analyses where using the original measures would be prohibitive. Finally, the *Q*-type factors can be used to examine the role of the individual items (or elements) of scales to see whether they function in the same fashion for all individuals (differential item functioning).

Q-type factors and profile description

The results of a *Q*-type factor analysis is a set of profiles (the score profiles of the *Q*-factors) the modal profiles. These profiles can also be thought of as axes in the multidimensional space of the persons. The profiles of a sample of individuals can be correlated with these modal profiles, and these correlations measure the resemblance of each individual profile to the modal profiles. If an individual profile is highly correlated with one of the modal profiles (say, greater than .50), it can be categorized or classified as a member of that profile type. The correlations of an individual profile with the modal profiles also represent the projections of the profile in the dimensional space defined by the modal profiles.

After the modal profiles are identified and named, an individual profile can then be described as resembling one (or more) of the modal profiles by the correlation of the individual profile with the modal profiles. This offers a way to simplify profile description. Rather than describing all the variables in the profile, one can simply describe the modal profile it most closely resembles. Comparison of two profiles can also be simplified by comparing them in terms of their proximity in the space of the modal profiles. Two profiles can each be plotted (their correlations with the modal profiles are their respective projections on the axes of the modal profiles), and the resulting points could then be compared as to how close they were, or actual distances between the points in the space of the modal profiles could be calculated. The interpretation of a number of personality inventories has been simplified by the isolation of *Q*-type factors which describe the modal profiles that characterize the major profiles obtained by individ-

uals on those instruments (e.g., Burger, & Cross, 1979; Skinner, Reed, & Jackson, 1976). The change of a profile over time could also be described using the above described graphical technique. Plotting the profiles of single individual obtained at two points in time in the space of the modal profiles is a way to conveniently assess movement (toward one modal profile? away from another modal profile?) of the individual's profile over time. As the number of variables in a profile increases, this method of illustrating and measuring change is much more simple than considering the changes in all the variables in the profile. Figure 1 illustrates the position of the profiles of a single individual at two time periods. Alternatively, the two points could represent the profiles of two different individuals measured at the same time. The plot shows that at Time 1, the individual's profile was similar to Modal Profile I, but at Time 2, the individual's profile was more similar to Modal Profile II. An extensive example of how changes in profiles over time can be evaluated using this method is presented in Burger, Calsyn, Klinkenberg, and Morse (2000).

Type membership as a categorical variable

As previously described, the profile of an individual can be compared (correlated) with the modal profiles identified in the *Q*-type factor analysis. Each individual profile (or most profiles) can then be categorized as resembling a particular modal profile (profile types). Now, consider the situation where the relationship of a set of variables to a particular dependent variable is being studied. For instance, this could be a study where the scales of a personality test (e.g., the MMPI) are related to the degree to which an individual acts in an abusive fashion (or takes drugs, or whatever). If the personality scales (13 in the case of the MMPI) were used as independent variables (predictor variables), we would have to examine not only the 13 main effects due to each of the scales, but would also have to test all the interactions between the predictors to determine the particular pattern of scores that might be related to the dependent variable. The potential number of interactions (in this example) would be voluminous, as it would involve all possible *2-way*, *3-way, 4-way,* to *13-way* interactions among the scales. Given the necessity of testing all the interactions, the sample size required to make the statistical tests would be quite large in order to achieve conventional levels of statistical power. An alternative analysis would be to classify each individual as a member of a particular modal profile or profile type (*unclassified* could also be employed) and use these categories as inde-

pendent variables predicting the dependent variable. Since the types summarize the major patterns of responding of the subjects, major relationships, if present, should be detected with more modest sample sizes, as fewer statistical tests are needed. The major configurations of the individuals score patterns are represented in the profile types, so if an interaction between the profile elements and the dependent variable is present, there will be significant differences between the profile types on the dependent variable. It is true that employing profile type correspondence (a categorical variable) rather than the original measurements (continuously measured variables) results in some loss of information. However, the ability to test configural hypotheses (interactions) with smaller sample sizes represents a distinct advantage for the use of profile types identified via *Q*-type factor analysis.

Q-type factors and DIF

Differential Item Functioning (DIF) refers to the question of whether the measurement properties of individual items on a scale are the same for two (or more) groups. While the equivalence of the measurement properties of entire tests over groups of individuals has been studied in the past (test bias, test fairness), more modern approaches have focused on the behavior of individual test items (DIF). If DIF is present in some items in a scale, the implication is that the meaning of test scores differs among the groups, because some of the items of the test have different properties over the groups.

Profile types, identified through *Q*-type factor analysis, represent subsets (groups) of individuals with different patterns of scores over all the scales. If the measurement properties of items on a particular scale are different for the various profile types (DIF), it is possible that individuals may achieve the same total score by responding differentially to the items on the scale evidencing DIF. Some explanations of DIF have suggested that the primary cause is the multidimensionality of the items in question. If a scale is factorially complex, some of the items may measure a common factor and in addition, one or more minor factors. If groups differ on these minor factors, DIF may be present.

There are a variety of procedures that may be used to detect DIF. Two very widely used methods are Item Response Theory (IRT) (Hambleton, Swaminathan, & Rogers, 1991) and the Mantel-Haenszel procedure. Both methods assume subjects are classified into groups (usually two), and the properties of individual items are contrasted over the groups. IRT analysis

involves contrasting the difficulty, discrimination, and guessing (if present) item parameters between the two groups. If an item differs on one or more of these parameters, the item is said to exhibit DIF. Graphically, this would mean that the item characteristic curves of the two groups are not equivalent. Figure 2 illustrates such a case where item characteristic curves of the two groups (Type I and Type II) are clearly different.

The Mantel-Haenszel approach examines the odds of members of each group (at a particular score level) getting a particular item correct. If there is no difference between the two groups (at each score level) in terms of proportion correct for a particular item, the odds ratios between the groups should equal 1.00 (at each score level). Guidelines are available to test departures from this equality (Zwick, & Ercikan, 1989). This procedure is described in Table 4.

Using either of these (or other) procedures, items exhibiting DIF can be identified. Such information may lead to a different interpretation of test scores, depending on the responses identified in the DIF analysis. *Q*-type factor analysis, then, provides us with a way to investigate whether the individual items on a scale have the same properties for all individuals.

Conclusion

Q-type factor analysis has been utilized considerably less than the more well known *R*-type factor analysis, and is perhaps not very well known to social science researchers. However, it is well suited for situations where the description and analysis of the profiles for individuals (entities) are the primary focus of the investigation. The isolation of the modal profiles underlying subjects responses to a particular set of measures simplifies profile description, allows more options for data analysis, and provides a new way to investigate the properties of individual items on scales. It is hoped that this paper may stimulate the use of *Q*-type factor analysis in social science research projects.

A leading competitor to *Q*-type analysis is cluster analysis. Cluster analysis, like *Q*-type factor analysis, can be used to identify clusters of individuals with similar score profiles. A major difference between the two procedures is that cluster analysis adds individuals to various clusters one at a time, while *Q*-type analysis makes use of all of the interrelationships among the individuals simultaneously in order to identify the *Q*-type factors. There is some evidence (Colihan, & Burger, 1995) that *Q*-type is superior to a number of clustering procedures in detecting groupings of individuals'

response profiles, particularly as measurement error becomes more pronounced in the data. With precision of measurement often being an issue in the social sciences, this factor is worth consideration.

Table 1
Hypothetical matrix of correlations* among six persons

PERSONS	1	2	3	4	5	6
1						
2	.82					
3	.92	.79				
4	-.10	.12	.18			
5	.06	-.1	.10	.91		
6	.04	.07	-.07	.73	.82	

* based upon four variables

Table 2
Hypothetical *Q*-type rotated Factor Matrix

PERSONS	*Q* FACTOR I	*Q* FACTOR II
1	.86	-.03
2	.85	.05
3	.89	.02
4	-.04	.89
5	.05	.91
6	.02	.88

Table 3
Hypothetical profiles of the *Q* factors based on the four scales

		Q F A C T O R	
		I	II
S C A L E	1	80	50
	2	50	60
	3	50	35
	4	04	89

Table 4
Using *Q* factors to detect DIF

Mantel - Haenszell method

Type by item score contingency table
(for a particular total score level)

	Item *M*	
Profile type	Right	Wrong
I	R_{IM}	W_{IM}
II	R_{IIM}	W_{IIM}

Test Whether the odds ratio

$$\frac{R_{IM} / W_{IM}}{R_{IIM} / W_{IIM}}$$

Is 1.00 over all total score levels for item *M*

Figure 1
Plot of individual's score profile in the space of the modal profiles at two time periods

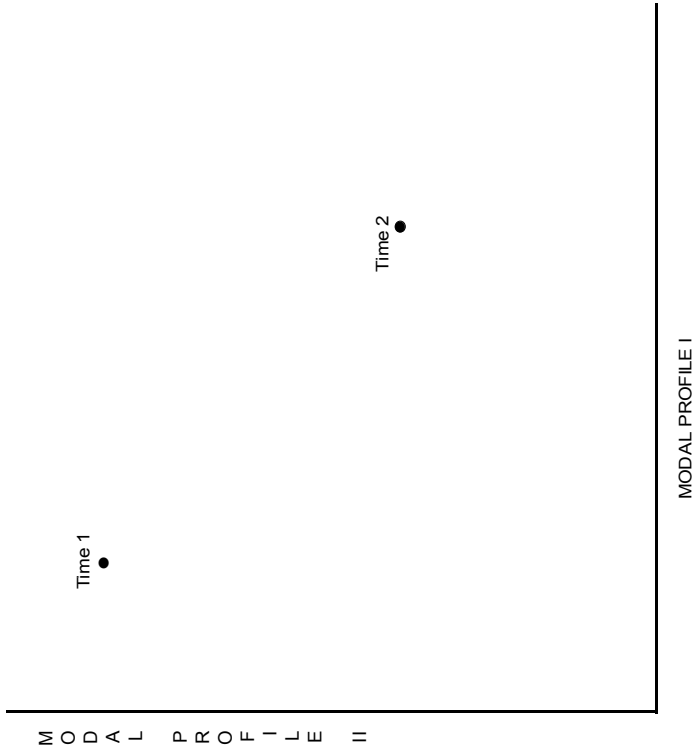

Figure 2
Using *Q* factors to detect DIF Item Response Theory (IRT)

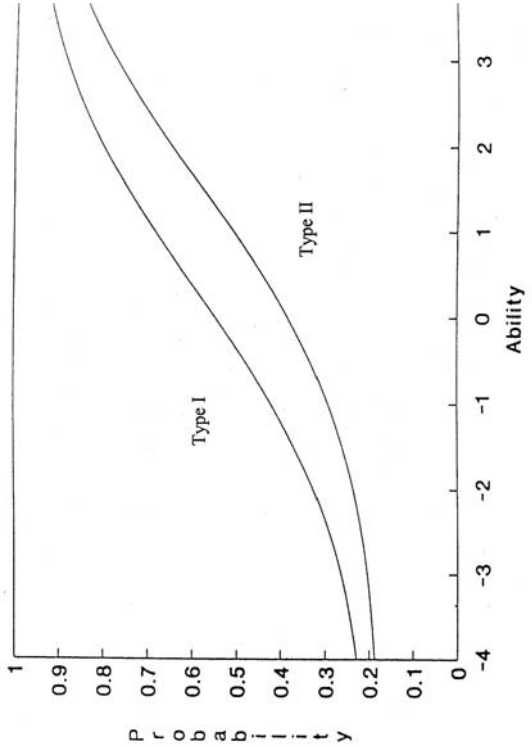

Figure 2

References

Ackerman, P.L., & Heggestad, E.D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*, 219-245.

Burger, G.K., Calsyn, R.J., Klinkenberg, D., & Morse, G. (2000). Prototypical profiles of the Brief Psychiatric Rating Scale. *Journal of Personality Assessment, 75*, 373-386.

Burger, G.K., & Cross, D.T. (1979). Personality types as measured by the California Psychological Inventory. *Journal of Consulting and Clinical Psychology, 47*, 65-71.

Burger, G.K., & Rimoldi, H.J.A. (1997). Using person factors to determine variable factors:An application with the MMPI. *Interdisciplinaria, 14*(1-2), 81-91.

Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970). *The handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Colihan, J., & Burger, G.K. (1995). Constructing job families: An analysis of quantitative techniques used for grouping jobs. *Personnel Psychology, 48,* 563-586.

Costa, P.T. Jr., & McCrae, R.R. (1982). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

D'andrade, R.G., Quinn, N., Nerlove, S., & Romney, A.K. (1972). Categories of disease in American-English and Mexican-Spanish. In A.K. Romney, R. Shepard, & S.B. Nerlove (Eds.), *Multidimensional scaling,* (Vol 2). NY: Seminar.

Filsinger, E., Faulkner, J., & Warland, R. (1979). Empirical taxonomy of religious individuals: An investigation among college students. *Sociological Analysis, 40*, 136-146.

Guilford, J.P. (1967). *The nature of human intelligence*. NY: McGraw-Hill.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage.

Laumann, E.O., & Knocke, D. (1987). *Organization state: Social change in national policy domains*. Madison: University of Wisconsin Press.

Magnusson, D., & Torestad, B. (1993). A holistic view of personality: A model revisited. *Annual Review of Psychology, 44*, 427-452.

Rimoldi, H.J.A., & Libonatti, E.B. (2000). Scores and principal components due to subjects and to variables. *Spanish Journal of Psychology, 3*(1), 68-75.

Skinner, H.A., Reed, P.L., & Jackson, D.N. (1976). Toward the objective diagnosis of psychopathology: Generalizability of modal personality profiles. *Journal of Consulting and Clinical Psychology, 44*, 111-117.

Sternberg, R.J., & Lubart, T. (1995). *Defying the crowd.* NY: Free Press.

Thurstone, L.L. (1938). Primary mental abilities. *Psychometric Monographs,1*. Chicago: University of Chicago Press.

Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP assessment. *Journal of Educational Measurement, 26,* 55-60.

Department of Psychology
College of Arts and Sciences
University of Missouri-St. Louis
8001 Natural Bridge Road
St. Louis, Missouri 63121-4499
United States

Received: April 6, 2001
Accepted: December 21, 2001