Factor Stability, Number of Significant Loadings, and Interpretation: Evidence from Three Studies and Suggested Guidelines

John R. Fairweather, Ph.D.
Lincoln University

Abstract: *Factor stability is an important issue for Q methodological studies that seek to identify viewpoints in a population, since it is possible that the addition of significant loadings can change the factor array and the consequent interpretation. This article examines change in the character of factors as the number of significant loadings changes and the subsequent effects on interpretation. The literature does not give firm guidance on the appropriate number of significant loadings. Data from three case studies are presented showing the extent of changes in the distribution of items in the selected factor arrays that occur as the numbers of subjects and significant loadings increase. In some cases, changes in item position altered factor interpretation. The results show that a number of research dimensions affect factor stability, so there are no uniform rules to guide researchers. The general applicability of the results is discussed along with suggestions relevant to the issue of factor stability.*

Introduction

In Q methodology the researcher usually aims to understand and describe the subjective viewpoints among the individuals being studied (Addams 2000). Viewpoints are represented by factors, each of which is based on the Q sorts of those subjects who load significantly on it. The number of significant loadings is variable and may be quite low (e.g., two or three subjects). Usually, factors found and interpreted are part of a theoretical enquiry and serve a useful purpose even with small numbers of significant loaders. Such research is typical of what McKeown and Thomas (1988) call *intensive* studies, in which one or only a few subjects may be studied. Intensive studies are well illustrated in Brown's *Political Subjectivity* (1980).

In other applications of Q methodology researchers seek to identify the different viewpoints in a population, usually of a large size rather than the small populations found in a classroom or a work organization, for example. These larger studies often have end users who want to be sure that the factors described are generally representative of the population of interest and are

Author's address: Agribusiness and Economics Research Unit, Lincoln University, P O Box 84 Canterbury, New Zealand; Fairweat@lincoln.ac.nz.

Operant Subjectivity, 2001 (October), 25 (1): 37-58.

relevant rather than esoteric and idiosyncratic. Such studies are of the type described by McKeown and Thomas (1988, 37) as *extensive* "because the intent is to determine the variety of views on an issue" and have a sample size ranging from 50-100 subjects[1]. Viewpoints have traditionally been identified using R methodology, that is, with formal instruments, random sampling, statistical testing, and inference from the sample to the population. That approach proceeds best when there is good prior knowledge of the characteristics being studied. However, Q methodology can also be used to differentiate and understand viewpoints in a population. In this application it is axiomatic that individuals can respond to a Q set and a condition of instruction in ways that reflect their subjectivity without constraint by the researcher. Q sorting reflects the structure of subjective response, and by-person factor analysis and inductive interpretation make that structure explicit. Assuming that human subjectivity has some structure is not the same as knowing what the structure is.

This application of Q methodology to discover the broad subjective viewpoints of a population may not be typical of its usual use in extensive Q studies. The approach is not R methodological because the research objective is to identify and understand the main contours of subjectivity without concern for the prevalence of each viewpoint within the population. The emphasis is on identifying and representing the qualities present, not measuring the quantities of an attribute of the population. In identifying such contours of subjectivity the researcher is taking a methodological position that is neither R methodology nor intensive Q methodology. Such an approach, however, is still a valid form of Q methodology, because it is based on self-reference.

As the number of significant loadings increases the character of the factor array is an issue, in particular the stability of the factor scores for each item in the array. For Q methodologists seeking to determine the variety of views in a population, the order of the items in the array is important in factor interpretation. In some studies the use of additional Q sorts may increase the number of significant loadings on a factor, which may alter the order of items in the factor array. Dramatic changes in item factor scores may, in turn, alter factor interpretation.

On technical grounds, item position might be expected to influence interpretation. When there are a small number of significant loadings, each additional loading contributes a relatively great amount to the character of the

[1] McKeown and Thomas do not give a detailed account of the differences between intensive and extensive person samples and they say that the terms are defined contextually. It seems reasonable from what they do say that the approach described in this article is consistent with extensive studies. There may be other approaches that are extensive that have research objectives different from those stated here.

factor. As the number of significant loadings increases, each new addition has a relatively smaller effect, and the factor scores may become more stable. This process raises the question of the nature of factor stability and its relationship to the number of subjects loading on it. If, for argument's sake, the number of significant loadings doubles from three to six, does this have much bearing on the character of the factor? What happens when there are ten or more significant loadings on a factor? Only in extensive Q studies, when the research objective is to identify the different viewpoints in a large population does this become relevant — i.e., when the researcher needs confidence that the factor and its interpretation are an accurate reflection of the viewpoint as it exists in the population.

The goal to provide stable, detailed descriptions of population-based viewpoints for some research employing Q technique may appear to be somewhat antithetical to the spirit of the extensive mode of Q methodology. In the author's opinion, the application of Q methodology to identify population viewpoints via the descriptive interpretation of factors is a small but important part of Q methodology, especially in applications that seek to identify and understand the main contours of subjectivity in a population. Consequently, researchers who use Q methodology to characterize widely held population views need to give attention to the stability of the factor as the number of significant loadings changes. The examples considered in this article arose from a program of research at Lincoln University using photographs as the Q set. The use of photographs for Q sorting requires relatively large numbers of significant loadings to obtain sufficient comments on which to base factor interpretation. The focus on the number of significant loadings led to the observation that factor characteristics can sometimes change as the number of significant loadings changes, and this may be particularly true when respondents have Q sorted rich, complex stimuli such as a Q set composed of photographic images.

The literature does not give firm guidance on the optimal number of significant loadings, and there is good reason for this. Many studies are theoretical in nature, and do not have as their main objective the more applied aim of accurately identifying the viewpoints widely held in a population. As long as the factors are distinct and advance theory or are based on an abduction, then it is not so important to be concerned about how precise and complete the descriptions are of the factors that represent the viewpoints in a population. In this vein, Brown (1980, 92) argues that "It is inconsequential, for example, that 27 persons were purely loaded on Factor A, for it only requires two or three variates to establish a common factor, the remainder merely serving to fill up factor space without altering the scores to any significant degree." He goes on to state, "All that is necessary, therefore, is to

include enough persons, typically no more than 40, to assure the comprehensiveness of the factors and the reliability of the factor arrays." Clearly, Brown is referring to the use of Q methodology in its extensive mode through the use of non-exhaustive P sets; i.e., there is no attempt to use a P set so large to satisfy any claim that it might be representative of all the population viewpoints.

Most Q studies have emphasized subject selection rather than the number of significant loadings. Subject selection is conducted so that the P set includes all relevant groups of people, ensuring that maximum diversity is obtained (Brown 1980, 191-4). It is expected that from this diversity a collection of relatively stable factors will emerge to define the landmarks of the subjectivity. With this orientation there is not as much need to consider the number of significant loadings, and once the factors are extracted interpretation proceeds accordingly. McKeown and Thomas (1988) and Addams (2000) do not specifically address the question of the appropriate number of significant loadings, thereby supporting the view that attention to subject selection is adequate. At present we do not have firm guidance about factor stability in relation to the number of significant loadings. Factor stability is an important issue, since it is possible that the addition of significant loadings can change the factor array and the consequent interpretation.

Thus, the research questions arise. 1) Does item position change as the number of significant loadings is increased? The results show that in some circumstances this does happen. The derivative question then is: 2) What is the effect of item position change on factor interpretation? There may well be changes in factor characteristics, but perhaps the basic character of the factor is unchanged. For example, an extrovert factor may still be extrovert even if there are changes in the position of items in the array. The results presented here are mixed because in some cases factor interpretation changed little as the number of significant loadings increased, but in others factor interpretation changed considerably. The final question addressed is: 3) Can any consistent criteria be derived to guide researchers using Q method that would indicate the factors are stable? Again the results are mixed, because there were many dimensions of research that bear on this question.

Method

The results from a 1999-2000 research program were utilized to develop answers to the research questions. The program focused on forestry and tourism issues, using Q methodology to distinguish subtle but important differences in the perceptions of natural character among the study populations. These differences in perception have implications that are important in planning. The studies are identified by the names of the three locations in New Zealand where they were conducted. The general research

objective in each study was to identify viewpoints among the population. Photographs were used as the items to be Q sorted. PQMethod (version 2.06)[2] was used to analyze the Q sort data and Varimax rotation was used in all cases to simplify the structure of the factor space.

Photographs used as Q sort items cannot be the sole basis of factor interpretation because the meaning of each photograph is highly contingent upon the respondent. While each written statement in a Q set can have a number of meanings or subtleties of meaning, it is still possible to construct a viable Q sort interpretation using only the statements. In contrast, photographs have diverse meanings deriving in large part from the rich variety of content and the responses each engenders in the subject. Comments, therefore, are essential for factor interpretation. The post-Q sort interview helps the investigators understand more precisely (in words) what part of a picture or what effect of an image was most influential in causing the sorter to assign a particular Q sort position to the photograph. Factor interpretation was based mainly on the recorded comments made regarding each respondent's six top-ranked and six bottom-ranked photographs. When the number of significant loadings is low (about three or four), there may be photographs in the factor array for which there are no available comments. Consequently, the smaller number of factors with a correspondingly larger number of significant loadings was preferred[3].

Coromandel

The first study was in the Coromandel Peninsula region of the North Island of New Zealand (Fairweather and Swaffield 1999). The research objective was to document public perceptions of natural character for a range of landscape features (water, vegetation, patterns, artifacts, and cues for care) in the main landform categories (ranges, foothills, estuarine, beach, and headland). In the Coromandel study a total of 88 locals and visitors, selected to represent all the main stakeholder groups, completed two separate Q sorts. The first was a "full range" Q sort composed of 25 images showing extremes of natural and unnatural character in the Coromandel Peninsula. The second was a "focus range" Q sort composed of 26 images showing more subtle gradations of changes within the extremes of natural character. The condition of instruction for each Q sort was to sort from most

[2] PQMethod is an MS-DOS program that was adapted, revised and maintained by Peter Schmolck (p41bsmk@unibw-muenchen.de). The Fortran code on which it is based was originally written by John Atkinson at Kent State University. Freeware copies of PQMethod are downloadable as a self-extracting zipped archive at www.qmethod.org.

[3] The criterion used to determine the number of factors selected for rotation was the presence of at least two significant loadings on the unrotated factor matrix. This ensured the number of factors was between two and five rather than a larger number obtained using the eigenvalue criterion.

natural to least natural[4].The forced choice pattern of the Q sort and associated item scores was[5]:

Coromandel Q sort distribution

No. in pile	1	2	3	4	5	4	3	2	1
Score	-4	-3	-2	-1	0	1	2	3	4

For the "full range" Coromandel Q sort, four factors were extracted and rotated. Only two were interpreted, because there were only two subjects loading on Factor 3 and one on Factor 4 (compared to 42 on Factor 1 and 23 on Factor 2). These secondary factors were distinctly minority viewpoints and were unimportant in the context of trying to determine public perceptions of natural character. The two main factors were quite similar with a correlation of 0.82. Despite the high correlation, there were important differences between the factors. While they agreed with what had natural character, they disagreed on what was unnatural. This pattern of two viewpoints was also found at two other locations in New Zealand, and these natural character results are reported in Newton et al., (*in press*). For the "focus range" Coromandel Q sort, three factors were extracted and rotated. Only two factors were interpreted, because there were only two subjects on Factor 3 (compared to 42 on Factor 1 and 25 on Factor 2). The two main factors for this "focus range" Q sort were less similar than in the "full range" Q sort with a correlation of 0.46.

Westland

The second study was located in the Westland District of the South Island of New Zealand (Fairweather et al. 2001). The research objective was to identify and characterize the subjective reactions of tourists and locals regarding a range of landscapes including general infrastructure and tourist infrastructure. There were two Q sorts: one for landscapes and general infrastructure and one for tourist infrastructure. In the Westland study 111 subjects, including locals, domestic visitors, and overseas visitors, were asked to sort 26 photographs in each Q sort. The condition of instruction was to sort from "most like" to "least like." The pattern of the Q sort and associated item scores was:

Westland Q sort distribution

No. in pile	1	2	3	4	6	4	3	2	1
Score	-4	-3	-2	-1	0	1	2	3	4

[4] The items scoring -4 can be considered to mean "most unnatural' but are expressed as "least natural" to make it easier for respondents to understand.

[5] The "focus range" Q sort with 26 photographs had a total of 6 photographs on the middle pile.

Only the first Westland Q sort on landscapes and general infrastructure is considered here, because it is slightly simpler. There were 32 significant loadings on Factor 1, 18 on Factor 2, and 27 on Factor 3. The correlations between the three main factors were: Factors 1 and 2 = 0.70, Factors 1 and 3 = 0.85, and Factors 2 and 3 = 0.67.

Rotorua

The third study was located at Rotorua in the central North Island of New Zealand (Fairweather et al. 2000; Fairweather and Swaffield *in press*). The research objective was to identify and characterize the subjective reactions of tourists and locals to a range of landscapes attractions, natural sites, and activities. In the Rotorua study a total of 66 subjects, including locals, domestic visitors, and overseas visitors, were asked to sort 30 photographs. The condition of instruction was to sort from most like to least like. The pattern of the Q sort and associated item scores was:

Rotorua Q sort distribution

No. in pile	1	2	3	5	8	5	3	2	1
Score	-4	-3	-2	-1	0	1	2	3	4

Five factors were extracted and interpreted. For Factors 1 to 5 there were respectively 18, 7, 13, 3, and 7 significantly loading subjects. The correlations among the four main factors were all less than 0.43. Only Factors 1 and 3 are considered here, because they had the largest numbers of significant loadings. As an aside, Factor 4 with only three significant loadings was interesting in that it represented a "picturesque landscape experience" which emphasized the pictorial and scenic qualities of the landscape and an affinity for architecture, a viewpoint found in an earlier study (Fairweather et al. 1998; Fairweather and Swaffield 2001), but not widely appreciated as important among visitors. The research in Westland and Rotorua did not seek to find only the dominant viewpoints but sought to describe all views found using the chosen criterion for selecting factors for rotation.

Analysis

For the purposes of the analysis presented here, a complete PQMethod analysis was done at regular intervals during the data input process and after all Q sorts were entered. The interval was nearly always ten cases. Each factor in the final printout was selected and then the intermediate printouts examined to identify the factor that corresponded to the selected final factor. Once all the intermediate factors had been identified it was then possible to report, for each ten subjects entered, the factor array and the number of significant loadings. It was also possible to calculate the correlation coefficient between the

intermediate factor and its corresponding final factor in order to measure the degree of similarity between them. These data were then included in a table that was used to highlight changes in the position of each item in the array.

The challenge in this technique was to identify which intermediate factor corresponded to the final factor. For the Coromandel study this was easy because there were only two main factors and the factors in all the intermediate analyses could be linked confidently to the final factors. For the Westland study it was more difficult, especially since the order subjects were interviewed reflected attempts to cover certain groups of people. For example, on one day six farmers were interviewed and on another day six local business people were interviewed. The order generated by this "lumpy" sampling meant that the intermediate factor characteristics changed considerably during the sequence of data entry. It was difficult and, at times, impossible to link an intermediate factor, on which there were very few significant loadings, to a final factor. In response to these difficulties, the 111 Q sorts were re-entered in random order[6]. In addition, given that the original data for the 111 Q sorts had a three-factor solution, three factors were specifically selected and rotated in the Q analyses after every ten Q sorts. This approach facilitated the accurate linking of the intermediate factors with the final factors.

There may be a case for not restricting the number of factors in the earlier Q sorts on the grounds that there may have been a different number of factors that could have been rotated for that particular number of subjects. That is, the data may have had an inherently different number of factors at that time. While this could be true, it remains the case that for the research objective of identifying viewpoints in a population, the final three-factor solution has much to recommend it. If that solution is the best representation of viewpoints then for our purposes it is acceptable to assess each earlier sample in those terms. Further, keeping to a three-factor solution applied a consistent rule across the data series and made it easier to determine how intermediate factors vary as the number of significant loadings increases. Because it is important to use similar methods when comparing different studies, the data for both the Coromandel and Westland studies also were entered in random order and intermediate analyses were examined for the minimum number of factors. In contrast, the Rotorua data were entered in original order and for an eight-factor solution because there was no difficulty in linking intermediate factors with final factors, and in order to minimize the amount of data reentry.

Comparing the results each time ten new cases are added to the analysis can demonstrate factor stability, but it is also necessary to examine whether factor interpretation changes. When comparing the factor arrays corresponding to n cases and n-10 cases, interpretation is not likely to be altered if an item in

[6] Random numbers between one and 111 were generated using Excel (Tools, Add-Ins, Data Analysis Toolpak).

one array moves to an adjacent column in another array. With more extreme movements, i.e., across an adjacent column to the next column, there may be changes in interpretation. The more extreme movements were counted and reported here. Multicolumn movements were inspected closely, and an assessment was made about their impacts on factor interpretation.

The method of counting movements parallels the correlation coefficient, but it identifies individual item changes more precisely. A cautionary note is necessary. As mentioned earlier, full interpretation when using Q sorts with photographs is dependent on the comments made about the photographs as recorded during individual post-Q sort interviews. This requires many significant loadings in order to get sufficient cases with comments. These comments are not available for the arrays examined here, and only a general indication can be made of any interpretation.

Results

Table 1 presents the results for Factor 1 of the "full range" Q sort in the Coromandel study. The table shows the increasing number of subjects and the corresponding number of significant loadings, along with the correlation coefficient between the array for that column and the final array with 51 significant loadings. With ten subjects there were eight significant loadings and a correlation coefficient of 0.99. There is a high level of similarity among all columns with correspondingly high correlation coefficients and no movements in item position to adjacent columns.

Table 1: Coromandel "Full Range" Factor 1

Column No.	1	2	3	4	5	6	7	8	9
Respondents	10	20	30	40	50	60	70	80	88
Significant Loadings	8	14	18	23	28	34	43	46	51
r	0.99	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00

Table 2 shows the results for Factor 2 of the "full range" Q sort in the Coromandel study. For this factor there were fewer significant loadings in the early stages, but for 30 subjects there were 12 significant loadings, at which point the correlation coefficient was 0.95 (Table 2, Column 3). The data available also report movement in items when each intermediate factor array is compared with the final factor array. For example, for the intermediate factor formed from two significant loadings there were six movements across two columns, three movements across three columns, and two movements across four columns (Table 2, Column1). Additional data show that the interpretation based on the array in column 1 is different from the interpretation based on the

array in column 9. One photograph that moved across four columns showed a townscape of shops and cars. An important feature of the interpretation of the final factor array was that these respondents consistently interpreted the townscape as most unnatural. Consistent with this interpretation is a photograph of a concrete relic amid regenerating bush: for the column 1 array it received a score of -3 but in the final factor array it received a score of +1 where it was interpreted as a sign of progression in the right direction.

The factor formed by six significant loadings was quite similar to the final array and its interpretation is similar except for a minor discrepancy caused the photograph of the townscape of shops and cars receiving a score of only -1. That is to say, subjects loading onto the factor with two significant loadings are still influencing the array based on six significant loadings. However, these results suggest that even with a modest number of only six significant loadings the corresponding factor gives an interpretation similar to that based on the final factor.

Table 2: Coromandel "Full Range" Factor 2

Column No.	1	2	3	4	5	6	7	8	9
Respondents	10	20	30	40	50	60	70	80	88
Significant Loadings	2	6	12	17	22	26	27	34	37
r	0.53	0.86	0.95	0.96	0.95	0.98	0.98	0.98	1.00
±2	6	2	1	1	1	0	0	0	—
±3	3	1	0	0	0	0	0	0	—
±4	2	0	0	0	0	0	0	0	—
Total ±	11	3	1	1	1	0	0	0	—

Table 3: Coromandel "Focus Range" Factor 1

Column No.	1	2	3	4	5	6	7	8	9
Respondents	10	20	30	40	50	60	70	80	88
Significant Loadings	6	13	15	22	28	33	41	48	50
r	0.93	0.96	0.96	0.98	0.98	0.98	0.98	1.00	1.00
±2	1	0	0	0	0	0	0	0	—
Total ±	1	0	0	0	0	0	0	0	—

Results from the Coromandel "focus range" Q sort are shown in Tables 3 and 4. Both tables show that the factor was relatively stable from the outset. There was only one item that moved by more than one column, and the interpretation of the factor is very similar to that for the final array.

Table 4: Coromandel "Focus Range" Factor 2

Column No.	1	2	3	4	5	6	7	8	9
Respondents	10	20	30	40	50	60	70	80	88
Significant Loadings	3	6	14	17	21	25	27	30	36
r	0.91	0.94	0.95	0.99	1.00	0.99	0.99	0.99	1.00
±2	1	0	0	0	0	0	0	0	—
Total ±	1	0	0	0	0	0	0	0	—

The results from the Coromandel studies show that overall there were few changes in the character of the factors as the number of significant loadings increased. For three of the four factors (Tables 1, 3, and 4) the interpretation did not change much when the first intermediate factor array was compared to the final factor array. For the remaining factor (See Table 2, Coromandel "full-range" Factor 2.), the interpretation based on the first intermediate array, for which there were only two significant loadings (Column 1), was different from the final array. The next array (Column 2) with six significant loadings had a similar interpretation to the final factor.

The Coromandel study was perhaps unusual in that the two "full range" factors were highly correlated at 0.82. This Varimax factor solution was very useful in distinguishing subtle but important differences in the perceptions of natural character among the study population and these differences in perceptions have important implications for planning. The distinguishing items highlighted the differences between factors. There was much in common for the two factors since they largely agreed on what had natural character but differed on what was less natural. The results presented here suggest tentatively that greater factor stability occurs where there are factors with moderate rather than high correlations.

Table 5 shows the results for the Westland study. Factor 1 is "Pure Nature" characterized by settings with relatively unmodified bush, mountains, and water. The presence of a rubbish dump and a gold mine in some photographs resulted in these being sorted as least liked photographs. The intermediate

factor in Table 5 formed by two significant loadings (Column 1) has nine items in different positions when compared to the final factor. In that array the photograph showing a rubbish dump receives a score of -1 rather than -4.

In contrast, the three least-liked photographs in the final factor (Column 11) were of built structures (dairy factory, electricity substation, and petrol station). These settings showed modification of an orderly character, unlike the chaos of the gold mine and rubbish dump. For this final factor, the randomness of the gold mine and rubbish dump is preferable to the order demonstrated by buildings and the array (Column 11) necessitates a different interpretation: one that favors unmodified natural settings and dislikes ordered human interventions. The intermediate factor formed by three significant loadings (Column 2) is similar in interpretation to the first intermediate factor because the photographs with buildings have the lowest scores. The intermediate factor formed by ten significant loadings (Column 3) is a closer match to the final factor (Column 11) with the rubbish dump taking its place as least liked. However, a pastoral farming scene receives a score of +2 compared to a score of 0 in the final factor array (Column 11). This result is not consistent with the pure nature interpretation. Consequently, the interpretation of this intermediate factor is similar, but not the same as, the final factor.

Table 6 shows the results for Westland Factor 2, "Living in Nature." Settings with buildings were given positive scores, in particular a photograph of the town center with hotel, and a photograph of the Glacier Hotel, receiving +3 and +1 respectively. The factor expresses appreciation of nature, but at the same time allows some built structures to be part of the experience. The intermediate factor formed by the first four significant loadings (Column 1) has some other pure nature photographs receiving high scores. Town center and Glacier Hotel each received a score of 0.

This means, if no additional Q sorts were available, the interpretation would have to change more toward the "pure nature" view of Factor 1, because none of the positively rated photographs have any built structures. The intermediate factor formed by three significant loadings (Column 2) bears a similar interpretation to the final factor (Column 11), with town center now receiving the highest score.

Table 7 shows the results for Westland Factor 3 "Pastoral Nature." Among the natural photographs liked by Factor 3 loaders were two similar pastoral farming scenes. The gold mine received a score of -1. The color green was apparently important to the Westland Factor 3 loaders, who liked the picturesque character and connection to the past; and they exhibited higher tolerance of infrastructure than others. The intermediate factor formed by the first three significant loadings (Column 1) shows only one of the farmland settings as being "liked," while the other farmland image received a

Table 5: Westland Factor 1: "Pure Nature"

Column No.	1	2	3	4	5	6	7	8	9	10	11
Respondents	10	20	30	40	50	60	70	80	90	100	111
Significant Loadings	2	3	10	4	12	22	18	18	27	28	32
r	0.73	0.82	0.91	0.85	0.88	0.94	0.98	0.97	0.99	0.97	1.00
±2	8	7	2	5	4	1	0	0	0	0	—
±3	1	0	0	0	0	0	0	0	0	0	—
Total ±	9	7	2	5	4	1	0	0	0	0	—

Table 6: Westland Factor 2 "Living in Nature"

Column No.	1	2	3	4	5	6	7	8	9	10	11
Respondents	10	20	30	40	50	60	70	80	90	100	111
Significant Loadings	4	3	7	11	13	11	12	11	11	12	14
r	0.77	0.85	0.87	0.94	0.94	0.95	0.97	0.98	0.98	0.99	1.00
±2	4	4	3	0	0	0	0	0	0	0	—
±3	2	0	0	0	0	0	0	0	0	0	—
Total ±	6	4	3	0	0	0	0	0	0	0	—

Table 7: Westland Factor "Pastoral Nature"

Column No.	1	2	3	4	5	6	7	8	9	10	11
Respondents	10	20	30	40	50	60	70	80	90	100	111
Significant Loadings	3	6	3	12	12	6	13	24	23	30	31
r	0.83	0.94	0.86	0.95	0.93	0.88	0.96	0.96	0.98	0.97	1.00
±2	1	1	4	0	1	0	0	0	0	0	—
±3	2	0	0	0	0	1	0	0	0	0	—
Total ±	3	1	4	0	1	1	0	0	0	0	—

neutral score. This intermediate result would weaken the "pastoral nature" interpretation. The intermediate factor formed by six significant loadings (Column 2) bears an interpretation similar to the final factor (Column 11).

Tables 5-7 also show that item position varies within the factor array as the number of significant loadings changes. The movement of items is greater compared to that in Tables 1-4, and more subjects are needed before a factor stabilizes. Tables 5-7 also show that the number of significant loadings can decrease as the number of subjects increases. For example, Table 5 shows that the number of significant loadings decreased from ten (in Column 3) to four (in Column 4) as the number of subjects increased from 30 to 40. Similarly, in Table 7 the number of significant loadings decreased from twelve (in Column 5) to six (in Column 6) as the number of subjects increased from 50 to 60.

As in the Coromandel study, the factors in the Westland study are highly correlated. In both studies this occurred because respondents tended to agree about what they considered to be natural or what they liked. In many cases what respondents in the Westland study liked was what they considered to be natural. Despite the similarities among factors there were statistically significant differences in the way some items were scored that formed the basis for distinguishing different preferences. In the context of planning for tourism it is useful to make these distinctions rather than defining how people like nature in general. The results show that when there were three significant loadings the arrays had seven, four, and three movements with correlation coefficients of 0.82, 0.85, and 0.83 in Tables 5, Column 2; 6, Column 2; and 7, Column 1 respectively. These factors were similar in structure to the final factor. Generally the interpretation of each of these intermediate factors with three significant loadings was only somewhat similar to the final factor: the meanings were similar but some important details were different. As larger numbers of significant loadings defined the factor the interpretation better matched that of the final factor. The Westland results, in contrast to the Coromandel results, have high correlations and modest levels of instability.

The last data set is from the Rotorua study. These Q sorts were entered in original order for a final solution containing eight factors, only five of which met the criteria for inclusion in the study. There were a small number of significant loadings on each factor. Tables 8 and 9 show that with the lower number of significant loadings there is great variation among the intermediate factor arrays, and this is reflected in the low correlations. Table 8 shows a final Factor 1 array with 18 significant loadings for a five-factor rather than an eight-factor solution. Table 9 also shows widely varying intermediate factor arrays for the final solution of Factor 3 with correspondingly low correlations.

Table 8 shows the results for Rotorua Factor 1, "Sublime Nature," Respondents loading on the final factor liked the natural features of bush and stream, large trees, and thermal activity (e.g., hot springs, geysers etc) but

in settings not necessarily devoid of human activity. The natural settings generated positive feelings of peace and quietness along with feelings of awe and power. The "most liked" settings show streams surrounded by bush. The intermediate factor formed by three significant loadings (Column 1) has two of three photographs showing lake and bush as "most liked," but one of these photographs is "neutral". At the other end of the array, a photograph of a recreation site is disliked by this intermediate factor but is neutral in the final factor. Similarly, two photographs showing tourist sites (thermal pool, Maori meeting house) are liked by the first intermediate factor but disliked in the final factor. The first intermediate factor in the Rotorua study has a recreational and tourist activity orientation rather than an appreciation of sublime nature. The contrasting positions of key photographs would necessitate a different interpretation. The general pattern for the factors formed by five, six, and seven significant loadings respectively (Columns 2-4), is for a gradual change in the position of those photographs just mentioned from the first intermediate factor in the direction of the final version of the same factor. The intermediate factor formed with eight significant loadings (Column 5) is similar to the final factor, with the six top-ranked photographs receiving the same scores and the three bottom ranked photographs receiving the same scores.

Table 8: Rotorua Factor 1 "Sublime Nature"

Column No.	1	2	3	4	5	6	7
Respondents	20	30	40	51	60	66	66
Significant Loadings	3	5	6	7	8	9	18
r	0.59	0.79	0.75	0.83	0.95	0.86	1.00
±2	5	4	4	3	1	2	—
±3	4	1	2	1	0	1	—
Total ±	9	5	6	4	1	3	—

Table 9 shows the results for Factor 3 "New Zealand Family." This factor prefers places that provide activities and attractions for the family and places that are natural, but not with the same sense of awe as Factor 1. Several intermediate versions of the factor show items that move as much as three columns as the number of significant loadings increases for the first five factor versions (Columns 1-5). For example, the photographs of an urban park and mud pools start as liked, then move to neutral or disliked and settle as neutral in the final factor. The photograph of Lake Tarawera moves from neutral

through strongly like to strongly dislike before settling as neutral. The Maori meetinghouse moves from slightly dislike through strongly like to slightly like. The mixed exotic forest moves from strong dislike through neutral to strong dislike. These large changes in the intermediate factor arrays make it unlikely that the interpretations based on fewer than thirteen significant loaders (Column 8) would be similar to the final factor array. Even if a case could be made that the factor is still essentially about family recreation, variations among the intermediate factors reveal differences in family preferences. The varieties among the interpretations of the intermediate factors are quite different when compared with the interpretation made for the final version of the array for Factor 3.

Table 9: Rotorua Factor 3 "New Zealand Family"

Column No.	1	2	3	4	5	6	7	8
Respondents	10	20	30	40	51	60	66	66
Significant Loadings	1	1	3	3	5	4	6	13
r	0.66	0.66	0.65	0.63	0.77	0.89	0.83	1.00
±2	7	7	9	10	3	2	3	—
±3	4	4	2	2	2	0	1	—
Total ±	11	11	11	12	5	2	4	—

The Rotorua results are interesting in that even with low correlations between factors there are many changes in item position across the different factors. For the first factor, 60 Q sorts finally yielded eight significant loadings that produced a stable factor array essentially the same ($r = 0.95$) as the final version of the "Sublime Nature" factor. For the "New Zealand Family" factor, even in the penultimate intermediate version (Column 7, $r = 0.83$), there were still four items more than one column different when compared to the final factor array.

Discussion

The results from all the studies reported above show that the character of factors sometimes changes as the number of significant loadings increases. The results from the Coromandel study with a relatively simple two-factor structure showed few changes in item position, while the results from the Westland and Rotorua studies showed greater movement of items. In the case of the Coromandel "full-range" study, Factor 1 was stable from the point of the first analysis with only ten subjects and eight significant loadings. The

Coromandel dataset was perhaps unusual in that there were only two factors, and Factor 1 was especially dominant. This meant that a high proportion of subjects loaded on Factor 1 from the outset. In all the other studies the number of significant loadings for the first intermediate analysis with ten subjects was small, ranging from two to six. Consequently the character of the factors changed as more subjects loaded on them.

The results also address the question of possible change in interpretation as factor characteristics change. In the Coromandel study three of the four factors had very similar interpretations, regardless of the number of significant loadings. In the other studies, however, there were changes in intermediate interpretations as the number of significant loadings increased. In all cases the interpretations gradually changed to better match that of the final factor as significant loadings increased in number. It cannot be concluded that the intermediate factor interpretations were always essentially the same as for the final factors. In some cases, a fundamentally different interpretation was needed.

It must be emphasized however, that in some cases an intermediate interpretation based on the factor having three to six significant loadings was similar, but not identical, to the interpretation of the final factor. In these particular studies there were 10 to 21 subjects in total. Thus it is possible for very small samples to identify factors in a population that are relatively unchanged even by the addition of many more subjects. Clearly it is quite possible for Q method studies to yield vital insight into the character of a factor with only a small number of significant loaders. The main findings of this paper shows that this general result, however, did not obtain for all factors or across all studies.

At least four points deserve further consideration in order to assess the general applicability of some of the results from the three studies. These would include the use of photographs in the Q set, the related issue of salience of the topic, the use of Varimax statistical rotations, and the appearance of highly correlated factors.

Perhaps Q studies using photographs produce atypical results. The studies reported here were all similar, using Q sets comprised of photographs of landscapes. Subjects sorted the pictures according to what they "liked" or what they thought was "natural." Full interpretation was based largely on comments obtained from post-Q sort interviews. These studies were, however, identical in their essential structure with other common types of Q studies, so it seems unlikely that these results are unique. Further, many Q studies have used images of some sort, and there has been no suggestion that they represent a different type of application of Q method. It must be acknowledged that in usual applications of Q, different results regarding factor stability might be

obtained even though that seems unlikely, since the same process of analysis occurs regardless of the type of items. The minor question raised here could be addressed by examining data from Q studies using statements.

The salience of the topic might be another important consideration. Studies reported here perhaps were not as emotionally powerful as issues such as abortion or politics, for example. Consequently the factors may not have polarized participants into distinct patterns with low correlations between factors. This argument may have some merit, but it needs to be tempered with the observation that some people have strong views about land, nature, and the environment. Many of the subjects in these studies were tourists visiting New Zealand with well-developed expectations about what the countryside should provide for them. Further, they have invested in considerable expense to be a tourist — going to New Zealand is not cheap — so they have a stake in what they experience. Similarly, local people living and working within an environment have strong attachments to their landscape and there are always tensions around conservation and production via farming or forestry. The evidence presented here indicates that intermediate factors would not necessarily be more stable in a Q study involving stronger emotions and with low correlations between factors. The Coromandel study had two factors with moderate correlation and reasonable stability, but the Rotorua study also had two factors with moderate correlation, and still the intermediate factors did not have stability.

Perhaps Varimax rotation influenced the results. There is some evidence that Varimax can lead to factor instability. The Q methodology network discussion list of 13 May 1996 reports a note from Steven Brown to the effect that the addition of two late-arriving Q sorts changed the pattern of loadings among three factors in a five-factor solution that was being inspected. The data themselves show that the factor with five significant loaders still had the same significant loaders after the additional Q sorts were entered. In the second analysis, changes occurred where there were only one or two significant loaders. Varimax appears to give variable results because it is responding to the low number of significant loaders. The addition of each new significant loader can be expected to have a large impact because it contributes a relatively great amount to the character of the factor. This change in the pattern of loadings is an example of instability caused by the low numbers of significant loaders, and it is not caused by the use of Varimax. If more significant loaders were found for Factors 1 and 2 then a stable factor would develop.

Further, the suggestion in this context is to use judgmental rotation to overcome the apparent problem derived from using Varimax. That strategy requires having some sense of what the researcher is looking for, that is, some

theoretical hunches to test. In many applications of Q method this is exactly what is recommended. But where the research objective is to identify the different viewpoints among a large population, it is not appropriate to build upon the researcher's expectations about what the pattern of subjectivity may be. The researcher does not have specific ideas to be tested other than the expectation that there may be different ways of experiencing some phenomena.

Another criticism of Varimax is that it spreads variance across all the selected factors. Because of this, the use of Varimax is ideal for the research objective here. Spreading the variance across all the selected factors results in identification of the different types of subjectivity rather than reflection of the numerical dominance of one very frequent viewpoint. Finally, Varimax is an accepted technique and it was reasonable to use it consistently.

Perhaps the presence of high correlations influenced the results. Here the objection is that high correlations indicated that the factors were not distinct, and therefore there was a tendency for items to move across columns. The results from the Coromandel study suggested that as correlations decrease, the intermediate factors are more stable. However, this suggestion was contradicted by the Rotorua results, which showed that even with moderate correlation there was a high level of intermediate factor instability. The evidence from the studies examined here is not conclusive because there were no factors with a correlation near zero. When the factors have a very low correlation, do factor characteristics change when more significant loaders are used? It must be acknowledged that in many Q studies there are no factors with very low correlations.

Conclusion

Three suggestions are made to tentatively guide researchers seeking stable factors representing populations. These results have important implications for those using Q methodology to identify viewpoints in a population. It does not mean that the emphasis on selection of subjects is misplaced but suggests that there are additional considerations that may be relevant in some applications of Q methodology. There are no simple rules that can be sustained on the basis of the data presented here, and additional research is needed before general claims can be strongly supported. It is appropriate, however, to make some suggestions. These are applicable only to studies where the research objective is to identify the main viewpoints in a population larger than would be found in a classroom or a work organization, for example.

The first suggestion is to inspect results as Q sorts are entered on a sequential basis and observe the level of intermediate factor stability. If there are few items moving by more than one column when the last factor is

compared to the one formed by n-10 subjects, then it is likely that the factor is stable. Pay attention to the interpretation supported by the penultimate factor and then compare this to the interpretation of the final factor. This suggestion means that the penultimate factor solution would be specified as having the same number of factors as the last solution.

The above suggestion does not indicate at which point the researcher should initiate inspections of factor characteristics; should comparisons begin when there are results from 30 Q sorts, or 50? In the absence of an end result that has stable factors, there would also be uncertainty as to when to stop collecting additional Q sorts. Further, with many Q sorts and apparently stable factors, the definitive test would be to repeat the random re-order analysis presented here. It is a time-consuming and tedious process to re-enter all Q sorts in random order and inspect intermediate results for every ten cases. This paper suggests a minimum number of significant loaders, and two more alternative suggestions can be made: One is for situations where there is a modest need for factor stability, the other is for when there is a particularly high need for factor stability. An example of the latter would be applied work for a city council in which there was a requirement in law to base planning on public preference. Quite precise requirements are needed because the location of an item can have a direct bearing on the policies and rules made as part of the legislative process.

Table 10 shows the results from all the studies reported here when moderate stability with similar intermediate factor interpretation is achieved and also for when high stability with very similar intermediate factor interpretation is achieved. The table suggests that in the first case, from three to eight significant loadings would give stable factors. While there were four instances of only three significant loadings giving moderate stability, for the purposes of suggesting a general guide it is necessary to be conservative and use a number that is likely to be adequate in most studies. Accordingly, from six to eight significant loadings would give moderately stable factors. In the second case where high factor stability is needed, from eight to 22 significant loadings would give the desired results. The data show only one occurrence of 22 significant loadings while most are very close to 12 significant loadings, a number that can be used as a reasonable guide.

For moderate factor stability, if a researcher needs to have six to eight significant loadings, then for a Q study with a five factor solution, 30 to 40 subjects will be needed, assuming an even spread of loadings. Based on studies reported here, it may be expected that approximately 70 per cent[4] of subjects load significantly, the remaining subjects either loading on more than

[4] For the Coromandel study it was 77 per cent, for the Westland study it was 69 per cent, and for the Rotorua study it was 73 per cent.

one factor or not loading on any. Consequently, in the case of a five-factor solution, the sample will need to include at least 47 to 57 subjects, but more subjects may be necessary, because it is very unlikely that the significantly loading Q sorts will be spread evenly across all the factors. For high factor stability, if 12 significant loadings are desired, then for a Q study with a five-factor solution, at least 60 significant loaders might be necessary or a total of at least 86 subjects.

Table 10: Summary of Key data

Moderate Factor Stability			High Factor Stability		
Significant Loadings	**Correlation**	**Changes**	**Significant Loadings**	**Correlation**	**Changes**
8	0.99	0	8	0.99	0
6	0.86	3	12	0.95	1
6	0.93	0	13	0.96	0
3	0.91	0	14	0.95	0
3	0.82	6	22	0.94	1
3	0.85	4	11	0.94	0
3	0.86	4	12	0.95	0
7	0.83	4	8	0.95	1
6	0.83	4	13		

In the end, however, since the number of factors in any given data set is indeterminate, even these suggestions may not be definitive. This means that the number of subjects needed in any study is a contingent matter, and this is appropriate and unsurprising for qualitative research. Such a contingent approach allows for the inherent variation in study results deriving from the interplay of all the considerations mentioned above.

The first suggestion of a minimal range of at least 47-57 subjects indicated for a five-factor study and the second suggestion of at least 86 subjects are within the range of 50-100 mentioned by McKeown and Thomas (1988) in relation to extensive studies. Generally then, the recommendations suggested here are consistent with the available guidelines. However, they are more precise and highlight the issues of factor stability and factor interpretations. On the basis of the results from three separate studies it is reasonable to conclude that care needs to be given to sample size when factor stability is important. Such care can help improve Q methodological research for those who seek to identify and characterize accurately the viewpoints in a population.

Acknowledgements

The New Zealand Foundation for Research, Science, and Technology provided funding for this research. Bronwyn Newton provided assistance with the data analysis associated with the research included in this article. She patiently re-entered much of the data and this helped the project considerably. The comments from five referees were helpful in clarifying the ideas contained in this article. Professor Simon Swaffield provided useful comments for the discussion of the results. Ongoing advice from Steven Brown and his general support of our program at Lincoln University is acknowledged.

References

Addams, H. 2000. Q Methodology. Chapter 2 in *Social discourse and environmental policy: An application of Q methodology,* edited by H. Addams and J. Proops. Cheltenham, UK: Edward Elgar Publishing Ltd.

Brown, S.R. 1980. *Political subjectivity: Applications of Q method in political science.* New Haven: Yale University Press.

Fairweather, J.R., Swaffield, S., and Simmons, D.G. 1998. Understanding visitors' experience in Kaikoura using photographs of landscapes and Q method. *Tourism Research and Education Centre (TREC) Report No.5.* Canterbury, NZ: Lincoln University.

Fairweather, J.R. and Swaffield, S.R. 1999. Public perceptions of natural and modified landscapes of the Coromandel Peninsula, New Zealand. *AERU Research Report No. 241.* Canterbury, NZ: Lincoln University.

Fairweather, J.R., Swaffield, S.R., and Simmons, D.G. 2000. Understanding visitors' and locals' experience of Rotorua using photographs of landscapes and Q method. *Tourism Research and Education Centre (TREC) Report No.13.* Canterbury, NZ: Lincoln University.

Fairweather, J.R. and Swaffield, S.R. 2001. Visitor experiences of Kaikoura, New Zealand: An interpretative study using photographs of landscapes and Q method. *Tourism Management* 22(3): 219-28.

Fairweather, J.R., Newton, B., Swaffield, S.R., and Simmons, D.G. 2001. Visitors' and locals' experiences of Westland, New Zealand. *Tourism Recreation Research and Education Centre (TRREC) Report No. 23.* Canterbury, NZ: Lincoln University.

Fairweather, J.R. and Swaffield, S.R. (Accepted for publication.) Visitors' and locals' experiences of Rotorua, New Zealand: An interpretative study using photographs of landscapes and Q method. *International Journal of Tourism Research.*

McKeown, B. and Thomas, D. 1988. *Q Methodology.* Newbury Park: Sage.

Newton, B.M., Fairweather, J.R., and Swaffield, S.R. (Under review.) Public perceptions of natural character in New Zealand: Wild nature versus cultured nature. *New Zealand Geographer.*