

# The Analytic Process of Q Methodology

**Siti Maftuhah Damio**

*Faculty of Education,  
UiTM Selangor, Kampus Puncak Alam  
maftuhah@salam.uitm.edu.my*

Received: 17 May 2018

Accepted: 23 May 2018

Published: 30 June 2018

## ABSTRACT

*The purpose of this article is to describe the analytic process of a method of data collection known as Q Methodology. This method is an alternative method in collecting data especially suited to research on “points of views” (Coogan & Herrington, 2011, p. 24). The analytic process of Q methodology involves factor analysis, a mathematical technique that reveals underlying explanations for patterns in a large set of data (Webler, Danielson and Tuler, 2007). This is known as Q technique factor analysis which “look for groupings of similar Q-sorts which represent similar viewpoints” (Bradley, 2007). To identify the factors, a statistical program known as Method, a program which has been tailored to meet the requirement of Q Methodology is utilised. This is a free downloadable program which can be accessed from the web. Method Version 2.11 was used to exemplify the process of Q Methodology analysis in this article. The analysis process involves three main stages. The first is data entry where encompasses 6 steps of keying in data. Data exploration is the second stage where a number of factors are produced by looking at patterns from the Q sorts. The final stage is data interpretation of the factors which is guided by the desire to remain true to what the data showed. This paper, however, looks into the first stage of Q analysis process which is data entry.*

**Keywords:** *analytical process, Q-Methodology, data entry, factor analysis*

## INTRODUCTION

Q Methodology is an alternative method of data collection that researchers can utilise. This is especially suitable for researchers who are working or intending to carry on issues related to “points of views” (Coogan & Herrington, 2011, p. 24). The introduction to what Q Methodology entails has been published in Asian Journal of University Education, Vol. 12. Therefore, this article on “The analytic process of Q Methodology” is a continuation to the introduction of Q Methodology.

The analytic process of Q Methodology involved rigorous techniques in extracting key findings by looking for major evidence without excluding the minor findings. In extracting these key findings, the Q methodology analysis adopted a statistical program, Method Version 2.11 for DOS, a freeware version of Q-method software tailored for the analysis of Q-sort data. (The software is available as a free download from [www.lrz-muenchen.de/~schmolck/qmethod](http://www.lrz-muenchen.de/~schmolck/qmethod)). Researchers interested to have a further look on the PQ programmes can access it using the following web addresses:

[http://schmolck.userweb.mwn.de/qmethod/pq\\_page](http://schmolck.userweb.mwn.de/qmethod/pq_page)  
<http://schmolck.userweb.mwn.de/qmethod/downpqwin.htm> for windows  
<http://schmolck.userweb.mwn.de/qmethod/webq/webQ>  
<http://schmolck.userweb.mwn.de/qmethod/pqmanual.htm> pq manual

This Method program is tailored to the requirements of Q-methodology. The software enables “factor analysis”, which is a mathematical technique that reveals underlying explanations for patterns in a large set of data. As explained by Webler, Danielson and Tuler:

*In the case of Q method the factor analysis looks for patterns among the Q sorts. The analysis produces a number of 'factors', which are particular arrangements of the Q statements – they are Q sorts. These are called 'idealized sorts' since they are produced by the analysis averaging together the Q sorts of several people. The job of the researcher is to read the idealized Q sorts and write a narrative for each one. These narratives summarize shared perspectives (2007, p. 19).*

Interpretation of the findings was guided by the desire to remain true to what the data stated, unaffected as much as possible by assumptions, personal perceptions and expectations of other interested parties. The purpose of this article is therefore, to explain and exemplify the Q analysis process basing it on a study of Autonomy in Language Learning. This explanation and description is closely related to the article “Q Methodology: An Introduction” by the same author.

### **Q-Analytic Process**

Thirty one completed Q-sorts during the fieldwork were the basis for factor analysis of the data. In each of the completed Q-sort, participants set out their perspectives regarding 40 statements on the issue of autonomy in language learning. As discussed in the introduction to Q Methodology article, this was done by sorting the statements from “most like what I think” to “least like what I think” in a Fixed Quasi-Normal Distribution grid. The ranking values range from +5, through zero to -5.

Q-Methodology employs exploratory factor analytic procedure to identify the factor structure or model for a set of variables. This often involves determining how many factors exist, as well as the pattern of the factor loadings (Stevens, 1996, p. 389). Fundamentally, it is a tool that summarizes the variables into distinct pattern of occurrences. This pattern would then be used in the subsequent stages, which, according to Rummel (1967, p. 445), address the question of “What are the patterns of relationship among these data?” Kerlinger (1979, p. 180) acknowledged that factor analysis is “one of the most powerful methods yet for reducing variable complexity to greater simplicity”.

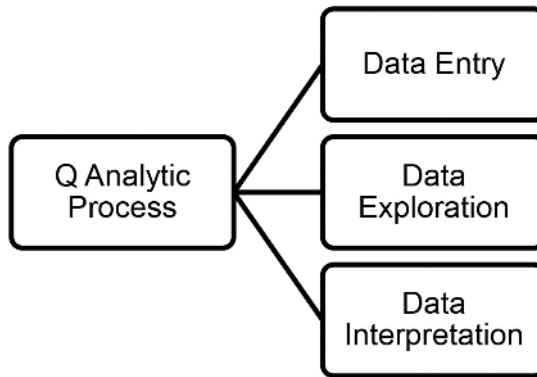
Cattell (1966) (in Thompson, 2004, p. 83) recapitulated that Q-Methodology focuses on the participants in the columns defining entities to be factored, and variables in the rows defining the patterns of association. Cattell (1966) further elaborated that Q-Methodology uses the Q Technique Factor Analysis which identifies people factors, rather than variable factors. The Q technique is used to address these three questions: How many types of people are there?, Which people belong to the different types?, and Which variables were the bases for delineating the different person factors?

Q Sorts are the basis for this Q Factor Analysis. The whole Q Sort, or overall configuration, a term coined by Stenner and Watt (2005, p. 80), are intercorrelated and factor analyzed. The analysis involves the correlation of one Q-sort to the Q-sorts of other participants. This is known as a 'by-person' correlation. This by-person correlation matrix is subjected to a Q technique factor analysis, from which factor loadings that represent the participants' different viewpoints are extracted. What it does is "to look for groupings of similar Q-sorts which represent similar viewpoints" (Bradley, 2007, p. 99). In this study, the aim was to then identify the number of different types of people and which people belong to which group in relation to their viewpoints on autonomy in language learning.

Q-sort analyses are made more appropriate with the availability of dedicated Q-methodological packages available like the PCQ for Windows (Stricklin and Almeida, 2001) and Method (Schmolck, 2002). These programmes, according to Stenner and Watt (2005, p. 81), "facilitate data input, automatically generate the initial by-person correlation matrix, and make processes of factor extraction, rotation and estimation very straightforward". These two packages were used to conduct the analyses effectively. One advantage of the later package is that it can be downloaded from the internet for free.

To exemplify the Q analysis process, Method was used. Even with the inclusion of some automatization process, the role of the researcher was still prominent in making decisions as the analysis progressed. This is seen as one of the strengths of Q-Methodology. Nonetheless, Q-Methodology can be challenging. It takes time and effort to familiarize oneself with the system. It was made more manageable in my case by some guidance from more experienced Q-Methodology researchers to show the way. Nonetheless, it was positively rewarding once mastered.

There are three main stages in Q technique factor analysis as presented in Figure 1. This article will be focussing more on the first stage of the Q analytic process, The Data Entry. The other two stages are to be touched upon briefly.



**Figure 1: The Three Main Stages of Q Analytic Process**

## The Data Entry

Entering the data into the Method programme is done by following a few steps. Basically, it involves six steps and they are explained accordingly in the following sections.

### Step 1: Entering statements

Statements used in the Q sorts were entered line by line. This entering of statements has to follow the numbering given to the statements during the preparation stage of the instrument. For example, statement one is entered as line one and this went on until statement 40 for line 40. These statements were saved in the .STA in the project folder.

### Step 2: Entering Q-sort

This step is known as the QEnter in the Method. There are a number of items of information needed at this stage, such as the number of statements (40), the distribution information (ranking value) based on the Q grid (-5 to 5) and the number of items that can be entered to each of the ranking value ( 2 items for -5, 3 items for -4 etc). The completed Q sort is used at this point when entering the individual Q sort into the programme. Apart from that, each Q sort was given identification. In this study, the participants were identified as L (DPLI) and S (trainee teachers from three different cohorts) followed by their group and initials. An example of the identification is S8NA. These data were saved in the .DAT in the project file.

### **Step 3: Extracting factors**

Once the statements and the Q-Sorts have been entered (40 statements and 31 Q sorts for this study), Centroid Factor Analysis is used to extract factors. This is known as QCent in the programme. A centroid refers to a kind of grand central average of the relationships between all the sorts, because they are represented by their correlation coefficients (Brown 1980a, p. 40). Kline (1994, p. 3) explains that correlation is “a numerical measure of the degree of agreement between two sets of scores”. This correlation runs along a continuum from +1, indicating full agreement through 0 in the middle, indicating no relationship, to -1, indicating full disagreement. The product of this factor extraction is factor loadings, which refer to the values expressing each sort’s relationship with the centroid. A column of numbers is generated, one for each of the Q-Sorts, which represents the extent to which the Q sort is associated with each factor, as shown in Table 1.

Table 1: Correlation Matrix between Sorts

SORT	TS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	LAZ	10	72	28	28	39	33	49	59	64	64	46	50	51	76	41	54	47	28	45	57	55	56	47	61	50	45	51	59	50	62
2	LMM	72	10	35	39	31	30	53	53	66	63	50	47	44	56	36	34	40	47	36	61	61	48	44	59	47	54	47	52	41	60
3	S2HM	28	35	10	21	-21	61	54	46	50	26	-2	58	43	40	45	-13	26	37	48	47	43	-10	36	27	52	65	16	50	1	35
4	S2HA	28	39	21	10	19	39	47	28	49	44	31	38	43	50	48	34	36	71	39	34	39	22	42	45	40	51	22	50	51	48
5	S2SR	39	31	-21	19	10	-5	8	20	13	55	57	1	9	47	21	47	-1	2	17	25	19	73	16	45	3	13	43	15	49	16
6	S4NN	33	30	61	39	-5	10	60	45	37	30	25	75	39	39	49	-5	43	44	53	50	50	7	48	34	66	59	10	63	16	51
7	S4NS	49	53	54	47	8	60	10	59	54	48	45	51	43	52	52	13	26	40	46	46	55	25	51	53	48	64	29	55	28	54
8	S4FA	59	53	46	28	20	45	59	10	55	52	30	55	57	60	32	9	12	36	40	61	61	37	53	62	59	65	35	54	35	45





18	S8LA	28	47	37	71	2	44	40	36	62	19	24	48	53	46	43	29	27	10	36	40	42	5	41	45	46	57	29	49	31	54
19	S8VA	45	36	48	39	17	53	46	40	46	40	17	57	36	43	72	10	22	36	10	65	59	16	36	62	73	43	14	67	38	50
20	S8NH	57	61	47	34	25	50	46	61	60	59	41	59	34	53	42	21	27	40	65	10	63	37	43	68	63	54	30	48	40	70
21	S8NA	55	61	43	39	19	50	55	61	66	47	35	52	53	43	57	1	24	42	59	63	10	20	38	65	63	57	34	61	18	43
22	S4MA	56	48	-10	22	73	7	25	37	26	79	59	16	13	51	17	52	14	5	16	37	20	10	24	53	27	20	52	29	66	39
23	S4MH	47	44	36	42	16	48	51	53	47	38	29	48	56	61	38	25	38	41	36	43	38	24	10	46	42	49	26	40	30	59
24	S4MH	61	59	27	45	45	34	53	62	52	59	65	42	41	61	63	35	31	45	62	68	65	53	46	10	53	47	54	51	63	52
25	S4RO	50	47	52	40	3	66	48	59	60	47	13	80	49	49	54	10	34	46	73	63	63	27	42	53	10	53	29	80	38	58
26	L5S	45	54	65	51	13	59	64	65	65	47	17	57	47	64	50	11	36	57	43	54	57	20	49	47	53	10	29	59	28	49

27	SAMA	51	47	16	22	43	10	29	35	34	48	46	27	36	54	27	43	34	29	14	30	34	52	26	54	29	29	10	30	44	43
28	LKS	59	52	50	50	15	63	55	54	54	45	18	70	49	59	59	22	38	49	67	48	61	29	40	51	80	59	30	10	43	52
29	LFA	50	41	1	51	49	16	28	35	32	52	51	31	23	49	39	39	29	31	38	40	18	66	30	63	38	28	44	43	10	46
30	LFW	62	60	35	48	16	51	54	45	57	58	43	62	52	61	41	42	45	54	50	70	43	39	59	52	58	49	43	52	46	10
31	LM	60	48	30	40	30	39	44	59	64	46	32	40	43	59	41	16	12	42	55	61	63	34	36	60	46	61	21	49	43	46

The next stage in the automatization is to produce the correlation coefficient. Correlation coefficient refers to the strength of a relationship between two variables and runs along the same continuum of correlation, with +1 indicating perfect positive relationship, to -1, indicating perfect negative relationship (Berenson & Levine, 1996, p. 732). This resulted in the Unrotated Factor Matrix (Table 2). The factor matrix or factor loadings “indicates the initial association, or correlation, of each Q sort with each factor” (Watts & Stenner, 2012, p. 103).

**Table 2: Unrotated Factor Matrix**

		1	2	3	4	5
1	LAZ	0.7894	-0.2727	-0.053	0.03	0.0937
2	LMM	0.7491	-0.1369	-0.0756	0.0111	0.1976
3	S2HM	0.495	0.5472	-0.0584	0.1786	0.156
4	S2HA	0.6079	0.0451	-0.0703	0.009	-0.2673
5	S2SR	0.3571	-0.5982	0.2677	0.178	0.0885
6	S4NN	0.6068	0.4568	0.0418	0.1181	-0.1397
7	S4NS	0.6962	0.1822	0.0051	0.0218	0.121
8	S4FA	0.7054	0.1417	0.086	0.0128	0.4469
9	S4MK	0.7637	0.1543	-0.2491	0.0619	0.1542
10	S4NA	0.7224	-0.348	0.1581	0.0476	0.1006
11	S8NS	0.5147	-0.4435	0.2281	0.0906	0.0623
12	S4NAA	0.7242	0.3285	-0.0768	0.0715	-0.2309
13	S4NH	0.6244	0.1864	-0.2621	0.0727	0.1314
14	S4NAM	0.8221	-0.1469	-0.1897	0.0334	0.1221
15	S8AS	0.654	0.157	0.138	0.0179	-0.2208
16	S8IE	0.4009	-0.5248	-0.3738	0.2194	-0.1602
17	S8KL	0.4634	-0.033	-0.3458	0.0783	-0.2829
18	S8LA	0.5992	0.2348	-0.3115	0.1024	-0.1097
19	S8NA	0.6572	0.2944	0.3669	0.0961	-0.2672
20	S8NH	0.7521	0.0355	0.1612	0.0044	0.0257
21	S8NAM	0.7127	0.2683	0.2058	0.0518	0.239
22	S4MAS	0.5134	-0.6868	0.2722	0.241	0.0302
23	S4MI	0.6277	0.0529	-0.1818	0.0297	0.0245
24	S4MH	0.804	-0.1432	0.309	0.0313	0.0131
25	S4RO	0.7378	0.3419	0.1303	0.0689	-0.2118
26	LSS	0.7282	0.3153	-0.0747	0.0665	0.1869
27	S4MAA	0.5279	-0.422	-0.1098	0.0835	0.0518
28	LKS	0.7587	0.237	0.0488	0.0333	-0.238

29	<b>LFA</b>	0.5857	-0.3671	0.1433	0.0523	-0.2305
30	<b>LFW</b>	0.7703	-0.0203	-0.1674	0.0227	-0.0935
31	<b>LMN</b>	0.6785	0.0887	0.1642	0.0097	0.1977
Ei	<b>genvalues</b>	13.5347	3.125	1.2407	0.2627	1.046
%	<b>expl.Var.</b>	44	10	4	1	3

The Q-Methodology researcher has to examine the unrotated factors to determine whether it is worthwhile to retain all of the factors or to try other factor extraction if the present result is unsatisfactory. Examination of the Eigen values, which “provide information to the communality ... in relation to each factor rather than to each Q sort” (Watts & Stenner, 2012, p. 104), is usually used in making this decision. In Q methodology, a higher Eigen value is achieved when more participants loaded on that factor. This represents an important viewpoint of the issue at hand. Nonetheless, in a Q-methodology study, which is interested in looking at the variety of viewpoints, a factor with smaller Eigen values could prove to be important as well. Therefore, instead of accepting the first rotation and its’ Eigen values, it is advisable to explore several different rotated factors solutions to see which of them is significantly loaded by the participants. Table 2 shows the unrotated factor matrix for 5 factors solutions. Factors with Eigen values of more than one are then extracted. The larger the Eigen values, the more variance is explained by the factor.

The researcher has to decide on how many factors are to be extracted and this is one of the unique aspects of Q-Methodology. The analysis is not totally computerised and the researcher plays a vital role in making decisions at certain stage of the process. According to Watts (2009 Q workshop), as a generic rule, it is best to “explain as much study variance and as many Q-Sorts as possible in the fewest number of factors”. The automated factors extraction is set at 7 by the programme, but the researcher could extract more or fewer than this suggestion. Objectively, there is no one correct number of factors to use as any number of factors will provide insights into how the participants (Q-Sorters) think. Nonetheless, Watts (2009) suggested that, as a rule of thumb, it might be a good idea to have the equation of 6 Q-sorts to one factor to start off with. It is advisable for the researcher to carry out a few factor extractions (Table 1) before choosing the data for the study. For this study, the researcher carried out 6 factor extractions, from 2 to 7 factors.

### Step 4: Rotating Factors

The factor rotation is computed by means of QVarimax which performs a varimax rotation of all the extracted factors. Once the rotation has been done, the loading of each Q-sort will be displayed. The researcher then has to calculate a significant factor loading manually using this formulation of  $2.58(1/\sqrt{\text{No of items}})$  for a 0.01 significance. For this study the significant factor loading is 0.38 ( $2.58(1/\sqrt{40}) = 2.58(0.144) = 0.372 = 0.38$ ). In order to check which Q-sort load significantly (on only one factor), non-significantly (not significantly associated with any factor) or is confounded (associated with more than one factor), the researcher has to manually go through each q-sort to determine them. Based on a 5 factor solution at the significance of .42, 24 out of 31 Q-sorts were significant, 7 confounded (they are loaded significantly on more than one factor) and 0 non-significant. This is shown in Table 3.

**Table 3: Significant Q-Sorts for Each Factor**

Factors	Significant Q-Sorts for each factor								% expl. Var.
	1	2	3	4	5	6	7	8	
1	4	6	12	15	18	19	25		17
2	5	10	11	22	27	29			17
3	17	18	30						9
4									0
5	3	7	8	9	13	23	26	31	18

### Step 5: Flagging Factors

Creating factor arrays for interpretation is the reason why the factors were flagged in the first place. The researcher has to enter the Q sorts associated with the factor (Table 10) and then decide which is to be included in the file. As for this study, factor 4 was not included as no Q sort was associated with it. Child (1970, p. 45) pointed out that “a factor loading in the factor analysis is worth considering for interpretation when it represents about 10% or more of the variance”. The significance of each factor for this study are 17 (Factor 1), 17 (Factor 2), 9 (Factor 3) and 18 (Factor 4). In total, 4 out of the 5 factors were taken for the next level of analysis.

### Step 6: Analysing Factors

This is the final step in the data entry process. All the data entered from step three, four and five were correlated and similitude among the participants were analysed. The information was saved in the .LIS in the project file. The above steps in the data entry process were repeatedly done with different significant factor loadings of .38, .40 and .42. Only the result of .42 is presented here as it is the best possible factor loadings data for this study, as seen in Table 4.

**Table 4: Significant factor loadings of .42**

Factor Extraction	F1	F2	F3	F4	F5	F6	F7	Total	N. Sig.	Confnd	Consensus State	Exp Var	Cor Max %
2 Fac	18	7						25	1	5	13		
3 Fac	14	6	1					21	0	10	13		
4 Fac	14	6	1	0				21	0	10	20		
5 Fac	7	6	3	0	8			24	0	7	5	61	
6 Fac	6	6	2	0	7	0		23	0	8	7		
7 Fac	4	6	1	0	5	0	2	19	0	12	9		
Decision	High	High	High	High	High	High	High	High	Low	Low	low	High	Low

In making a decision on the final set of factors that is to be used in the analysis, Jaffares (2010) suggested a few criteria. There should be a high number of loaders for each of the factors, for the total of the Q sorts and total explained variance of the factors. Meanwhile, there should be a low number of loaders for non-significant sorts, confounded sorts, the number of consensus statements and the maximum degree of correlations. Webler, Danielson & Tuler (2009, p. 31) proposed a similar formula, with the exception of naming the criteria. The four criterias are:

- **Simplicity:** fewer factors are better as it makes the viewpoints at issue easier to understand.

- Clarity: the factor which each sorter loaded highly on
- Distinctness: lower correlations between factors are better as highly correlated factors are saying similar things
- Stability: certain groups of people tend to cluster together

Based on the suggested criteria and the comparison of the different rotations, the best combination was from the five factor extraction. It consisted of the second highest total loaders, the second lowest of confounded and the lowest of consensus statements. Though it is a five factor extraction, only four factors were used as they have significant number of loaders. The other one factor is insignificant as there is no loader for the factor. This five factor solution with four significant factors at the value of .42 is the basis for the next stage of data exploration, before they are passed on to the third stage, data interpretation.

## CONCLUSION

This article has delved into Q analytic process, a process in analysing the data for a Q methodology study. Though there are three main stages in the process, namely data entry, data exploration and data interpretation, only the data entry stage has been discussed extensively. The six steps in the data entry stage were explained through an exemplification of a study on autonomy in language learning. The findings of the data entry in the form of factors are sought which are then considered for exploration and interpretation. The use of specific program, namely Method, though can be challenging, is very fruitful upon mastering it. It is hoped that this article has shed further light on how researchers can utilise Q Methodology in their research.

## REFERENCES

- Coogan, J & Herrington, N., (2011). Q Methodology: An overview. *Research in Teacher Education*, 1(2), 24-28.
- Berenson, M. L., & levine, D. M. (1996). *Basic Business Statistics: Concepts and Applications*. Prentice Hall Trade.
- Bradley, J. (2007). *Q methodology*. Unpublished PhD Thesis, University of Sheffield, United Kingdom.
- Brown, S. R. (1980). *Political subjectivity: applications of Q methodology in political science*. New Haven: Yale University Press.
- Cattell, R. B. (1966). The Scree Test for the Number of Factors. *Multivariate Behavioral Research*, 1(2)
- Child, D. (1970). *The Essentials of Factor Analysis*. London: Holt
- Jaffares, S. (2010). *Q methodology workshop*.
- Kerlinger, F.N. (1979). *Behavioral research: A conceptual approach*. New York: Holt, Rinehart, and Winston
- Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.
- Rummel, R. (1967). *Understanding Factor Analysis. Journal of Conflict Resolution*, v11, n 4, pp 444-480. Retrieved from: [http://scholar.google.co.uk/scholar\\_url?hl=en&q=http://94.23.146.173/ficheros/0ceeb142b589e93fd332c6fe795f3430.pdf&sa=X&scisig=AAGBfm0AJCyAVUR5p3dryp](http://scholar.google.co.uk/scholar_url?hl=en&q=http://94.23.146.173/ficheros/0ceeb142b589e93fd332c6fe795f3430.pdf&sa=X&scisig=AAGBfm0AJCyAVUR5p3dryp)
- Schmolck (2002). *Method*. Retrieved from : <http://www.lrz.de/~schmolck/qmethod/>
- Stenner, P., Watts S., and Worrell M., (2008). 'Q Methodology', in Carla Willig and Wendy
- Stainton-Rogers (eds.), *The Sage Handbook of Qualitative Research in Psychology*, (pp. 215-239)..Los Angeles, CA: Sage



- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stricklin, M., & Almeida, J. (2001). *PCQ: Analysis software for Q-technique* [Computer software]. Retrieved from <http://www.rz.unibw-muenchen.de/p41bsmk/qmethod>
- Watts, S., & Stenner, P., (2005). *Doing Q Methodology. Qualitative Research in Psychology*, 2: 67-91
- Watts, S. (2009). *Q methodology workshop*. London: University College London.
- Webler, T., Danielson, S., and Tuler, S., (2007). *Guidance on the use of Q Method for Evaluation of Public Involvement Programs in Contaminated Sites*. Report Prepared Social and Environmental Research Institute, 278 Main Street, Suite 404 Greenfield, MA 01301 (413) 773-9955. [www.seri-us.org](http://www.seri-us.org). Retrieved from: [www.epa.gov/publicinvolvement/pdf/policy2003.pdf](http://www.epa.gov/publicinvolvement/pdf/policy2003.pdf)

