Encyclopedia of Research Design

Q Methodology

Contributors: Steven R. Brown & James M. M. Good
Editors: Neil J. Salkind
Book Title: Encyclopedia of Research Design
Chapter Title: "Q Methodology"
Pub. Date: 2010
Access Date: October 16, 2013
Publishing Company: SAGE Publications, Inc.
City: Thousand Oaks
Print ISBN: 9781412961271
Online ISBN: 9781412961288
DOI: http://dx.doi.org/10.4135/9781412961288.n348
Print pages: 1149-1156

http://dx.doi.org/10.4135/9781412961288.n348
Q methodology is a combination of conceptual framework, technique of data collection, and method of analysis that collectively provides the basis for the scientific study of subjectivity. This is distinguished from R methodology, which provides the basis for the study of what is objective in human behavior. Innovated in the mid-1930s by British physicist–psychologist William Stephenson, Q methodology focuses on opinions and perspectives that are gathered using the well-known Q-sort technique. These data are then submitted to factor analysis, pioneered by Stephenson's mentor Charles Spearman, which reveals the segmentation of subjectivity inherent in the substantive domain under consideration. Given the ubiquity of subjectivity, Q methodology applies to all areas of human endeavor—social attitudes, decision making, administration, the arts and humanities, cultural values, policy, economics, education, and so on, and including the natural and physical sciences insofar as subjective substructures undergird the theories and understandings used to explain objective facts. The focus on subjectivity in Q methodology requires departures from conventional applications of sampling, experimental design, and factor analysis. This entry provides information on the conceptual framework, data collection, and analysis of Q methodology, along with resources for additional information.

Concourse of Subjective Communicability

Q methodology has its roots in *concourse*, a term coined by Stephenson that refers to the universe of subjective communicability surrounding any topic, of the kind found in ordinary conversation, back-fence gossip, commentary deposited on Internet blogs and exchanged in chat rooms, and extending to the high-level discourses of epistemic communities across all the sciences. Facts are invariably interlaced with opinions, and the division between the two turns on the principle of self-reference, its absence in R methodology and its centrality in Q. Hence, it is one thing to state that “free-flying wild birds can be a source of avian influenza,” which is a matter of fact, and another to assert that “scientific surveillance of wild birds should be maintained worldwide,” which is an opinion. The latter is self-referential to those who assert it, and it is apt to be accompanied by other opinions from the same concourse—for example, that “compartmentalization would reduce risk for introduction of AI viruses via trade,” that

“there is an immediate need to promote international research programs to develop better vaccines,” and so forth.

The volume of opinion constitutes the universe of communicability that, in principle, is infinite in magnitude. Unlike sampling in surveys, where population boundaries can be specified and the number of cases is finite, the boundaries of communicability cannot be fixed and its content is limitless. Concourse is not restricted to statements of **[p. 1150 ↓]** opinion. The population of impressionist art also constitutes a concourse, as do the universes of novels, musical compositions, political cartoons, landscapes, and flavors of ice cream; in short, any collection of stimuli, linguistic or otherwise, for which individuals might express preferences. A defining feature of concourse is that its contents are generally untestable and incapable of falsification; however, they remain subject to measurement.

Q-Sample Structuring and Experimental Design

Concourses are typically voluminous, as are person populations in survey research, and one of the steps in Q methodology involves reducing the concourse to a small sample of statements suitable for experimentation. As an illustration, consider the following statements, which were among more than 200 drawn from the media and interviews prior to the 2003 U.S. war with Iraq:

As a counterpart to random sampling in survey research, the statement population in Q methodology is modeled theoretically (typically according to R. A. Fisher's experimental design principles), and statements are then selected in terms of the model. In this instance, examination of the concourse revealed statements compatible with the view of the U.S. administration (such as statement 1 above) and statements that were contrary to the regime's position (statement 2), but also statements that expressed ambivalence or reticence (statement 3), and virtually all of the statements in the concourse found easy placement in one of these three categories.

As a device for increasing diversity and breadth in the Q sample, all statements were reexamined and located in one of the three categories comprising Harold Lasswell's concept of *perspective:* For example, statement 1 expressed a *demand* (from a pro-regime standpoint), statement 2 an *expectation* (from an anti-regime standpoint), and statement 3 an *identification* (from a reticent point of view). All 200 statements were subsequently placed in one of the (3)(3) = 9 cells of the factorial design in Table 1, and $m = 5$ replicates from each cell were selected, for a Q-sample size of $N = (5)(9) = 45$ statements. Each of the statements was then typed on a separate card, resulting in a pack of 45 cards for subsequent Q sorting.

Table 1 Q-Sample Structure (Iraq War)

Main Effects	*Levels*			*n*
(*A*) Regime	(*a*) pro	(*b*) reticent	(*c*) anti	3
(*B*) Perspective	(*d*) demand	(*e*) identification	(*f*) expectation	3

Q Sorting

Q technique is the most widely known feature of Q methodology, and its popularity far exceeds awareness of the more abstract methodology that it was invented to serve. Q sorting consists of ranking the items in a Q sample according to some condition of instruction, typically from *mostly agree* to *mostly disagree.* Table 2 provides an example in terms of the Iraq War study; hence, this person agreed with statements 18, 22, and 45:

18. We have a professional army and the best military technology on the planet. The Iraqi army is no match.

22. The world would be safer without Saddam.

45. The United States has a duty to protect itself and the world.

and disagreed with statements 17, 30, and 40:

17. We're acting like a terrorist regime ourselves and only seem to be in this for self-interested capitalist motives.

30. This is a racist war, a campaign of genocide. We're just taking our 9/11 frustrations out on Iraq.

40. I must admit the idea of war scares me, and frankly I'd rather not hear about it.

[p. 1151 ↓]

This person obviously supports the impending conflict and at least manifestly embraces assertions emanating from the regime while denying with equal intensity propositions issued by counter elites.

The width of the range in Table 2 is relatively unimportant and has no statistical impact (nor does the shape of the distribution), but several principles concerning the Q sort are of considerable importance. First, as should be apparent, there is no correct way to perform the Q sort: It is a wholly subjective affair, which means that issues of validity are largely irrelevant. In a science of subjectivity, there can be no external criterion by which to validate “my point of view.” Second, the distribution should ideally range from *mostly* to *mostly* (e.g., from *mostly agree* to *mostly disagree*), with *nonsalience* in the middle. This principle draws on 19th- and early 20th-century psychophysics, such as the work of Gustav Fechner and J. G. Beebe-Center, as well as observations dating from the earliest times down to Sigmund Freud's principle of pleasure-unpleasure and including the extremes of Likert scales. Third, meaning distends in opposite directions from zero and becomes most vivid at the extremes, as indicated in the tendency for Q sorters to feel most strongly about statements placed at the extremes and to be relatively non-engaged with respect to those located near zero. (Both of the above principles depend on a Q sample that is balanced in terms of the perspectives at issue; e.g., pro-regime vs. anti-regime.) Fourth, the forced distribution in Table 2 has the status of a *model*, in this case, a model of the Law of Error. Virtually all of the literature that has promoted an unforced (free) distribution has done so under the mistaken assumption that the forced distribution distorts how people would actually behave if unconstrained, rather than that it is an intentional constraint (although not without behavioral support) for inducing participants to reveal their preferences. Fifth, the statements in the Q sort are entangled; that is, they interact in a single Q-sort setting, and each is implicitly compared to, and achieves its score in relationship to, all others. (Entanglement is a signal feature of quantum mechanics, with which Q methodology demonstrates

striking parallels.) This is in marked contrast to the situation in R methodology where a score obtained on a measured trait by any single individual is wholly independent of the scores obtained by other individuals for the same or any other trait. Sixth, the subjective character of both concourse and Q sorting necessitates (where feasible) a post-sorting interview so as to permit participants to elaborate upon their Q sorts; that is, to clarify the meanings that they attributed to the statements (which may be other than the meanings assigned in Table 1) and to justify the salience (from +4 to–4) that they assigned to each. This a posteriori determination of meaning is in marked contrast to R methodology, where scale meaning is achieved a priori in terms of validity and reliability tests that are carried out before responses are obtained.

Table 2 Q-Sorting Distribution

Mostly Disagree						*Mostly Agree*		
–4	*–3*	*–2*	*–1*	*0*	*+1*	*+2*	*+3*	*+4*
17	29	3	9	4	2	1	8	18
30	32	13	11	7	6	5	12	22
40	35	21	15	10	14	24	26	45
(3)	37	43	23	16	20	31	28	(3)
	41	44	25	19	33	36	38	
	(5)	(5)	34	27	42	(5)	(5)	
			(6)	39	(6)			
				(7)				

Based on the 3 × 3 factorial design in Table 1, the Q sort in Table 2 could be analyzed in terms of the following variance components:

$$\Sigma d^2 = \Sigma A^2 + \Sigma B^2 + \Sigma AB^2 + \Sigma R^2$$

$$df(44) = (2) + (2) + (4) + (36),$$

where $_{\sigma}d^2$ is the Q-sort total sum of squares, $_{\sigma}A^2$ is the regime effect, $_{\sigma}B^2$ is the sum of squares due to perspective, $_{\sigma}AB^2$ is the regime × perspective interaction, and $_{\sigma}R^2$ is the replication variance (within-cell error). As noted at the beginning, however, the focus on subjectivity in Q methodology necessitates departures from conventional research assumptions, among them that items have an agreed-upon meaning as they do in scaling and as they would be required to have for the calculation **[p. 1152 ↓]** of means and variances necessary for the application of ANOVA. It is for this reason that

the utility of variance designs is restricted to the composition of Q samples where they serve the important purpose of ensuring stimulus *representativeness*, a principle first advanced by Egon Brunswik. Beyond that, variance analysis yields to the operational principle and is superseded by correlation and factor analysis, which respond to the way in which participants actually operated with the statements rather than the way in which the investigator structured them.

It is to be noted that the sample of persons who perform the Q sort (P set) is also usually selected on the basis of experimental design procedures and is typically balanced or semi-balanced for gender, age, party identification, or other salient variables arranged factorially. The P set is normally small (typically in the range of 30–50 participants), but as diverse as possible. The goal is for a representative set of participants to respond to a representative set of stimuli so as to maximize the likelihood that whatever subjective segmentations are in circulation will have an opportunity to reveal themselves.

Correlation and Factor Analysis

Data analysis is illustrated in terms of an experiment on literary interpretation, an area of study rarely examined by other quantitative methods, but easily attended to using Q methodology due to the subjectivity involved. The focus in this instance is on the poem *Piano* by D. H. Lawrence, the reading of which produced a diversity of reactions that were subsequently reduced to a Q sample of size $N = 30$ critical comments including three levels of *reaction* (positive, negative, mixed) cross-classified with three kinds of *poetic concern* (emotion, technique, sense) in a 3 × 3 factorial arrangement. After reading the poem, eight graduate students of literature were instructed to provide their appraisals of the poem by Q sorting the statements from *agree* (+4) to *disagree* (–4), and Table 3 contains an abbreviation of the scores given the 30 statements by two of the students, *X* and *Y.*

Table 3 Q-Sort Scores

	X	Y	d^2	X^2	Y^2
1	1	1	0	1	1
2	−1	2	9	1	4
3	4	4	0	16	16
⋮	⋮	⋮	⋮	⋮	⋮
28	3	−1	16	9	1
29	0	1	1	0	1
30	−1	3	16	1	9
Σ	0	0	82	182	182
σ	2.46	2.46			

The formula for Pearson's correlation when means and standard deviations are the same (as they are when Q sorts adhere to the same forced distribution) is

$$r_{XY} = 1 - \frac{\Sigma d^2}{2N\sigma^2}$$

, where N = 30 statements, σ^2 is the variance of the Q-sort distribution, and σd^2 is the sum of squared differences in statement scores between the two Q sorts. Under these conditions, $2N\sigma^2$ is a constant and equivalent to the composite sums of squares of the two Q sorts (182 + 182 in this case); hence, the correlation between X and Y is

$$r_{XY} = 1 - \frac{82}{364} = .77$$

, which signifies a high degree of similarity between the two interpretations of the poem. The standard error of the correlation is given by

$$\sigma_{r=0} = \frac{1}{\sqrt{N}} =$$

, and 2.58(.18) = .47 indicates that correlations exceeding ± .47 are significant (p < .01).

Given the n = 8 Q sorts, their intercorrelations result in an 8 × 8 correlation matrix, which is then factor analyzed using either of the two dedicated software packages, QMethod and PCQ (both of which can be accessed at http://www.qmethod.org), or by SPSS. Although principal components analysis (PCA) and varimax rotation are the most commonly used procedures in R methodology, investigators who use Q methodology are more inclined toward centroid factor analysis and increasingly toward theoretical

rotation, although varimax is frequently used. Centroid analysis and theoretical rotation are both supported by the QMethod and PCQ programs, as are PCA and varimax.

The reasoning behind these analytic preferences again turns on the phenomenon of subjectivity and **[p. 1153 ↓]** the alteration in conventional analytic practices that this focus demands. The centroid solution is indeterminant (i.e., is one of an infinite number of solutions), and although varimax injects determinacy into the solution, it does so based on a statistical principle (simple structure) that is assumed to apply in any and all situations. Theoretical rotation, by way of contrast, takes context into consideration (in conformity with J. R. Kantor's principle of specificity) and as a result is increasingly relied upon as a way to find solutions that are *operant* (i.e., genuinely functional) as opposed to merely statistical. Implementation of theoretical rotation involves the graphical plotting of factor loadings in two-dimensional Cartesian space and manually rotating them so as to reveal suspected or hypothesized effects, and thereby relies on the abductory logic of Charles Peirce.

Table 4 shows the factor loadings for the eight participants, who offered three different interpretations of the poem, with Participant 1 advancing one interpretation (Factor A), Participants 2–5 advancing another (Factor B), and Participants 6–7 yet another (Factor C); Participant 8's response is mixed. The standard errors for correlations shown above apply to factor loadings, which are therefore significant in this case if in excess of ±47. The Q sorts comprising a factor are merged to provide a single Q sort, and this is effected by first weighting each response in terms of its factor loading:

$$w = \frac{f}{1 - f^2}$$

, where *f* is the factor loading and *w* is the weight. Consequently, in the case of Factor 2, Q sort 2 is weighted w_2 = 1.26 and Q sort 5 is weighted w_5

= .82, the latter therefore weighing only 65% of the former in the calculation of the factor scores. The factor scores are then calculated by multiplying each statement's Q-sort score by the weight and then summing each statement across the weighted Q sorts comprising the factor, with weighted statement sums then being converted into a factor array presented in the form of the original +4 to–4 metric.

Table 4 Factor Loadings

	A	B	C
1	71	27	–04
2	–21	68	–21
3	–12	67	28
4	09	67	21
5	–23	56	–21
6	–19	–03	57
7	–06	31	59
8	50	–52	28

Note: Decimals omitted; significant loadings are in boldface.

Examining Factor A, it is apparent that the persons comprising this factor have responded favorably to the poem, as shown in the following four statements (factor scores for A, B, and C, respectively):

+4–2–2 (*a*) The striking thing is that the poet knows quite well that this reversion to a childhood incident is sentimental, but he does not try to make capital out of the sentiment.

+4–2 +4 (*b*) One is made keenly aware of the strange relationship of past and present experience—one feels the emotion the poet experienced through his identity with and separation from his past self.

–4 +3 +2 (*c*) The triviality of the sentiment is equaled only by the utter puerility of the versification.

–4–1 +4 (*d*) This poem is false. One worships the past in the present, for what it is, not for what it was. To ask for the renewal of the past is to ask for its destruction.

The determination of significant differences among factor scores depends on factor reliability, which depends in turn on the reliability of the individual Q sorts. Adopting

1,2

= .80 as a conservative estimate of average individual test-retest reliability, factor reliability is given by

$$r_{xx} = \frac{(.80)p}{1 + (p - 1).80}$$

, where *p* is the number of persons defining the factor; hence, for Factor B, with *p* = 4 persons defining the factor, *r*
BB

= .94. The standard error of factor scores is given by

$$\sigma_f = \sigma\sqrt{1 - r_{xx}}$$

, where σ is the standard deviation of the Q-sort distribution (in this instance, from Table 3, σ = 2.46) and σ
f

is the standard error; hence, for Factor B, σ
fB

= .60. Given that Factor A has only *p* = 1 defining Q sort, its factor reliability is *r*
AA

= .80 and the standard error of its factor scores is σ
fA

= 1.10. The **[p. 1154 ↓]** standard error of the difference in factor scores between Factors A and B is obtained by

$$\sigma_d = \sqrt{\sigma_{fA}^2 + \sigma_{fB}^2} = \sqrt{1.10^2 + .60^2} = 1.25$$

, which, when multiplied by 1.96 and rounded up to the nearest whole number, indicates that differences of 3 or more between scores for Factors A and B can be considered significant ($p < .05$). Based on this criterion, statements (*a*), (*c*), and (*d*) above can be seen to distinguish Factor A from the other two factors; statement (*b*) distinguishes Factor A from Factor B, but not from Factor C. Statistical information of this kind, which is routinely provided by the QMethod and PCQ software packages, along with postsorting interviews, comprise the main ingredients for the interpretation of factors.

Factor Interpretation

Arguments and inferences in R methodology typically depend on the matrix of factor loadings, which reveals the connections among variables, but the factor loadings in Table 4 are of lesser interest in Q methodology inasmuch as the person sample is small and nonrepresentative by survey standards. Interest in Q methodology instead focuses on the array of factor scores, as with statements (*a*) to (*d*) above, because the factor scores reveal the subjectivity at issue, in this case concerning poetic interpretation. Interpretation is also more intricate in Q methodology because statements can, and frequently do, take on different meanings to different people, and to the same people in different settings, as opposed to scale items carrying meanings that have been established a priori.

In this regard, consider the Q-sort responses of 15 pediatric oncology patients concerning the fatigue that they experienced during treatment. The responses resulted in three factors, the first of which (Factor X) assigned positive scores to the following distinguishing statements (scores in parentheses for Factors X, Y, and Z, respectively):

> I had energy (+4, 0,–3). I felt great (+4, +2,–4). I had enough energy to play as hard as I wanted (+3,–3,–3). I was so tired that I needed to rest during the day (+2, 0,–1). I had enough energy to play sports (+2,–2,–4).

The statement about needing rest provided a clue indicating that these patients had incorporated a rest period into their regimen that enabled them to regenerate and lead

otherwise normal lives. Factor Y, on the other hand, told a different story (scores for Factors X, Y, and Z, respectively):

> Being tired made it hard to keep up with my friends (–1, + 4, +1). Being tired made it hard to concentrate (+1, +4, +1). Being tired made me sad (–3, + 3,–3). My tiredness was hard on my family (–2, +2,–1). I was too tired to watch television (–1,–4, 0). I felt sleepy (+3,–2, +3).

The relative emphasis on concentration and sadness led to a conclusion that fatigue was affecting these patients in cognitive, emotional, and social ways that might require counseling to reactivate them and get them out from in front of the television. Note, however, that their fatigue did not spill over into sleepiness. Finally, Factor Z:

> I feel tired (+2,–1, +4). I felt weak (+1, +1, +4). Being tired made me mad (–3, +1, +3). I had trouble finishing things because I was too tired (0,–1, +2).

These patients felt weak and not just tired, and unlike Factor Y, their emotional reaction was one of anger rather than sadness, hence providing them with a motive force for improvement that Factor Y lacked.

The above are bare caricatures of the three factors, which contained interesting nuances that could be drawn out only through careful and prolonged examination of the factor arrays. The results led to suggestions for crafting a questionnaire, based on the Q factors, that could serve as a screening device aimed at distinguishing patients who were experiencing one of these three reactions to fatigue. Factor X patients already have control over their lives and require little in the way of assistance, whereas the Factor Y and Factor Z patients require different therapeutic interventions. The interpretive phase of a Q study has much in common with hermeneutics as well as narrative and discourse analysis, and it is at this point that Q methodology resembles qualitative methods, despite its roots in variance design and factor theory. As a practical matter, the factors that it produces represent the vital signs of human **[p. 1155 ↓]** subjectivity and serve as decision structures that point in promising directions of implementation.

Resources

Additional information about Q methodology can be found at http://www.qmethod.org, including information on an electronic discussion group and the International Society for the Scientific Study of Subjectivity. Three journals specialize in Q-related studies: *Operant Subjectivity, Journal of Human Subjectivity*, and the Korean language *Q-Methodology and Theory.* Books on Q methodology are also now available in Korean, Persian, Romanian, and Thai.

Steven R. Brown and James M. M. Good

http://dx.doi.org/10.4135/9781412961288.n348
See also

Further Readings

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

Brown, S. R., Durning, D. W., & Selden, S. C. (2008). Q methodology. In K. Yang, ed. & G. J. Miller (Eds.), Handbook of research methods in public administration (2nd ed., pp. pp. 721–763). New York: Taylor & Francis.

Brown, S. R., & and Robyn, R. Reserving a key place for reality: Philosophical foundations of theoretical rotation. Operant Subjectivity, (2004). vol. 27, pp. 104–124.

McKeown, B. F., & Thomas, D. B. (1988). Q methodology. Newbury Park, CA: Sage.

Stenner, P., Watts, S., & Worrell, M. (2007). Q methodology. In C. Willig, ed. & W. Stainton Rogers (Eds.), The Sage handbook of qualitative research in psychology (pp. pp. 215–239). London: Sage.

Stephenson, W. (1953). The study of behavior: Q-technique and its methodology. Chicago: University of Chicago Press.

Websites

Q Methodology: http://www.qmethod.org