Estimating Agreement between the Q Block and the Q Tool

Hung Kyu Kim
Hankuk University of Foreign Studies

Abstract: *The purpose of this paper is to explore the agreement or consistency between the Q Block and the Q Tool, both of which were constructed to identify typologies based on the factors found initially with Q methodology. Both tools are very useful to examine the characteristics of people who belong to a specific type, and to test the difference by types in terms of demographics and other related variables. This study adopted two examples to test the agreement rate, which is defined by the percentage of people indicated as having the same type by both tools. The first example, with three factors, showed 62.20 percent agreement, while the second example comprising four factors showed 70.08 percent agreement. Researchers should be aware of the advantages and limitations of each tool when they choose one.*

Purpose of the Study

The purpose of this study is to examine the agreement or consistency between the Q Block (Talbott, 1963; reprinted in this issue) and the Q Tool (Kim, 1999). The two devices were contrived to identify the Q typologies based on the structure of the Q factors and its characteristics throughout Q analysis. Researchers sometimes continue to test their hypotheses or working theories. Some researchers might also want to compare the demographics and topic-related variables among Q typologies with a sizable number of people. These Q-based R studies might not be economically and technically feasible if we administer standard Q sort to every respondent. Both Q Block and Q Tool would allow researchers to easily assess the factors that people belong to.

This agreement study between two assessment tools is to examine the equivalency element of reliability of the measurements. It might also provide an opportunity to test the predictive and constructive validity of the operational measurements.

Comparison of the Q Block and the Q Tool

From the perspectives of a quantitative/qualitative dichotomy, Q methodology can be viewed differently. The first perspective is that Q

Contact author: doctorq@hufs.ac.kr
Operant Subjectivity: The International Journal of Q Methodology, 2010, 34(1): 25–37

might be an example of a mixed model, because it uses numbers and employs correlation coefficients and factor analysis while adopting more qualitative procedures in interpretation. The second perspective is more radical. It disregards the usefulness of quantitative/qualitative dichotomy. It holds that almost all steps in the analysis already form part of quantitative and qualitative methods. The preparation of Q samples involves both quantity and quality. Researchers consider both quantitative and qualitative information and use abductory technique when conducting theoretical or judgmental rotation. When researchers decide where they have to stop in varimax rotation, judgment is also involved.

A Q factor is, after all, the result of interaction between the researcher and the data used. A factor array is not a mere average of the Q sorts similarly sorted, but the complex gestalt or a patterned whole. In this sense, Q methodology is more than a complex method, and a question such as where does Q locate in the structure of quantitative and qualitative dichotomy should be approached in the nature of classification and measurement in human science.

Nevertheless, it would be possible to enhance the usefulness of Q methodology in a broader perspective. A working theory or hypothesis, for instance, can be tested further, and the typologies found through Q analysis, in fact, can be identified by the factors found through Q analysis, revealed with the tools. For example, before asking huge number of voters their political opinions, an election campaign manager wants to find out Q factors. The manager can produce efficient campaign strategy through correlating voting-related variables and the factors.

Researchers sometimes want to know the proportion of people with each Q-factor opinion and their geographical and demographical distribution in order to produce more efficient strategies (as shown in election-campaign example). Thus, we develop an assessment tool to identify these questions, since it is not practical to conduct a Q study on a large scale. This approach adopts a linking role to connect typologies found in a Q study to an R study.

Q Block

This assessment device was first developed by Talbott (1963). From the arrays of statement *z*-scores demonstrating belief system about fallout shelters and radiation, he then selected a number of item sets. Each set contained four items—because he had ended up with four factor-solutions—to meet two criteria as shown in his four blocks (see pages 21–23 in this issue). First, each set included a statement from the array of each of the four shelter types at about the same level of acceptance. Second, each statement was one that the other types less substantially accepted. Such a set of four items constitutes a Q Block.

Talbott's sets show an example of the scoring procedure that could be used. Four scores would be derived for each respondent, one for each of the four types. Each score would be the sum of the ranks assigned to the four statements—one from each Q Block—associated with that particular type. The four scores for each subject would provide the basis for assignment to a type of orientation toward fallout shelters. One possible criterion for assignment could be on the basis of the highest score, 16. In the scoring example, the respondent would be assigned to Type A.

One of the disadvantages of the Q Block is that researchers can find it hard to select enough statements having discriminating power to construct Q Blocks. Another possible problem is that the way in which we construct a block with a specific combination might affect the internal consistency of the measurement.

Q Tool

The Q Tool was developed to alleviate the difficulty of Q Block construction, keeping internal consistency and higher validity. Since Q Tool is a short description about each type, researchers easily prepare core sentences based on the interpretation of Q-factor analysis. These statements used in Q Tool have, by nature, a comparative characteristic, because the interpretation of Q factor employs several input data such as a specific typal array, typal differences, interviews, demographics, and their related variables. This means that we can avoid additional comparison (sorting) to identify the typology as Q Block does. And, more importantly, researchers do not have to select the statements discriminating factors to construct the Q Block. Nor do respondents have to rank the statements in a Q Block. Respondents simply choose the one among the typologies provided that they think are most alike or close. The Q Tool also prevents tie situations, which the researchers frequently encounter in reality when using the Q Block. The Q Tool was first used in Kim's study on the characteristics of the Korean consumer by its value and lifestyle typology (Kim, 1998) as shown in Table 1.

Table 2 shows a high correlation between time-lagged measurements of the Q Tool: the proportion of five typologies each in 1998 is very similar with that in 2005, which demonstrates the Q Tool's high reliability. It also tests a theory that terminal value systems in a society do not change in a short period of time. It was also reported that there were no difficulties in self-designation when respondents are forced to choose only one type they belonged to, while the Q Block showed numbers of tied Q-type scores.

Table 3 is another example that supports this result. In the study of TV news-viewing types, the proportion of each type shows similar pattern. Since news-viewing behavior should be more susceptible to

change than the terminal value system, a one-year time interval was used to examine the measurement reliability.

Table 1: Example of Q Tool for Terminal Value

Value type	*Characteristics*	*Mark (X) only one I belong to*
Family-centered	I value family security the most and a comfortable life with my family or loved one. I am willing to disregard or sacrifice even my self-esteem or accomplishment when my family-centered values come up.	
Self-sacrificing practice	I am willing to relinquish my life of freedom and excitement for the sake of my children's success. Through children's success in education and getting a socially respected job, my family acquires social respect and recognition.	**X**
Emotional stability	I seek happiness or contentment through emotional stability such as peaceful mind, inner harmony, and wisdom. So, I value beauty, peaceful world, and true friendship that are related to emotional, psychological aspects.	
Personal goal achiever	I most value an achievement that brings a sense of accomplishment and self-respect. For a goal achievement, a world of beauty or equal life loses its meaning. I'm not interested in peaceful world and salvation either.	
Hedonic value-oriented	I seek hedonic values like freedom, beauty, pleasure, and exciting life. My goals would be acquired and enhanced mainly with the hedonic value. I always prioritize this value. I don't really care about family values.	

Source: Kim, H. K. (1999, p. 176).

Q Flow

Meanwhile, researchers sometimes need to handle a small number of Q samples because the Q population itself is limited in nature. For instance, Kim used 20 aroma essential oils as a Q sample to find out olfactory factors, and then provided a handy assessment tool, Q Flow (Kim, 2001). As shown in Figure 1, by using only a few aroma oils like lemon, ylang-ylang, and fennel, this tool easily could identify people's preference type. (Refer also to Table 4). In the first stage, if one likes lemon, a respondent falls in either type A or B, and if he or she continuously likes ylang-ylang, then we classify the respondent as Type A. If a respondent dislikes ylang-ylang in the second test, and likes fennel, then he or she falls in Type B. In the third stage, if a respondent

Table 2: Frequency Distribution for each Terminal Value

Terminal value type	*1998**		*2005***	
	Frequency	*Percent*	*Frequency*	*Percent*
The family-centered	964	51.2	772	51.5
The self-sacrificing practice	318	16.9	172	11.5
The emotional stability	186	9.9	241	16.1
The personal goal achiever	190	10.1	223	14.9
The hedonic value-oriented	224	11.9	92	6.1
Total	1882	100	1500	100.1

*Sources: * Kim, H. K. (1999, p.176; ** Hyundai Research Institute (2005, p. 11).*

does not like fennel, then his or her preferences may be confounded, or he or she may not belong to any type we found. By the same token, a person who previously disliked lemon but in the first test likes fennel can be identified as Type C. To select these representative Q items having factor-discriminating power, researchers should pay attention to the items in the arrays of typal *z*-score differences between factors including two factors, and a specific factor and all other factors. Researchers easily identify the typology when we have a small number of Q samples, like 20 or fewer, and people show relatively strong preference to some Q samples over others. A set of selected Q samples conducts a Litmus role in each test stage.

Table 3: Frequency Distribution of TV News-Viewing

Viewing type	*June 2008*		*October 2007*	
	Frequency	*Percent*	*Frequency*	*Percent*
The careless viewer	106	26.8	885	22.9
The news briefs	153	38.7	186	48.2
The actively engaged	108	27.3	93	24.2
The attention to periphery	28	7.1	18	4.7
Total	395	100.0	38	100.0

Source: Kim, S. H. (2008, p. 69).

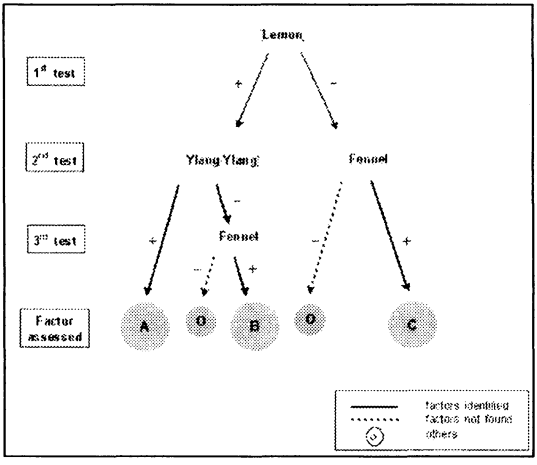
As shown in Figure 1, Q Flow consists of three test stages using only three essential aroma oils, and a solid line means positive, or like orientation of a certain type, while a dotted line means negative, or not

like a type, which is a contaminated and atypical type of preference.

Table 4. Aroma Assessment Tool

Type A (Ecstasy)	Type B (Meditation)	Type C (Comfort)
Lemon (+)	Lemon (+)	Lemon (–)
Ylang-Ylang (+)	Fennel (+)	Fennel (+)
	Sandalwood (+)	Eucalyptus (+)

Figure 1. Q Flow of Aroma Factor



Research Method

To test the degrees of agreement between the two assessment tools for factor identification, this study used two different measurement sets for comparison, as shown in Tables 5–8. Tables 5 and 6 indicate examples of fruit preference analysis that resulted in three factors, and Tables 7 and 8 show examples from a four-factor TV news-selection study.

For the analysis, a total of 150 responses were collected from May 28, 2008 to June 5, 2008, but 23 cases were excluded because Q block produced the same type scores across at least two types in either one of the two tests. The reliability formula developed by Holsti (1969) was adopted for the test.

$$\text{Reliability} = \frac{2M}{N_1+N_2}$$

(M is the number of coding decisions on which two coders agree; N_1 and N_2 refer to the total number of coding decisions by the first and second coder, respectively.)

Since this study compares the typal identification by two sets of measurement, N_1 and N_2 must be the same number. So, the agreement

formula can be modified as follows:

Agreement = M/N

where N refers to the number of cases, 127 in our test, and M is the number for which two measurements identify the same type.

For example, in the first comparison for fruit preference, if 70 out of 100 respondents were identified as the same typology by both Q Block and Q Tool, then agreement would be 0.7.

Table 5: Example of Q Tool for Fruit Preference

Type	*Characteristics*	*Choose one*
A	I like fruits tasting sweet and sour, but neither small ones like oriental cherry and plum nor sweet only like mango and persimmon.	
B	I like sweet and sour fruits, but mostly fruits having high sugar content. I don't like fruits having too much water such as water-melon or Asian pear.	
C	I like fruits having soft pulp and water. I don't like thick and sour taste.	X

This method is straightforward and easy to apply, but it can be criticized because it does not take into account the occurrence of some agreement strictly by chance, an amount that is a function of the number of categories in the analysis. To take this into account, Scott (1955) developed the Pi index, which corrects for the number of categories used and also for the probable frequency of use as follows:

$$\text{Pi} = \frac{\text{observe agreement (\%)} - \text{expected agreement (\%)}}{1 - \text{expected agreement (\%)}}$$

Table 6: Example of Q Block for Fruit Preference

Type	*Fruit*	*Most like*	*Neutral*	*Least like*	*Preference*		
					A	*B*	*C*
A	Banana	3	2	[1]			
B	Mango	[3]	2	1	1	3	2
C	Watermelon	3	[2]	1			
A	Tangerine	3	2	[1]			
B	Hallabong	3	[2]	1	1	2	3
C	Melon	[3]	2	1			
A	Pear	3	[2]	1			
B	Pineapple	3	2	[1]	2	1	3
C	Cherry	[3]	2	1			
				Type Scores	4	6	[8]

Table 7: Example of Q Tool for TV News-Viewing

Type	*Characteristic*	*Choose one*
Careless viewer	I'm a bit indifferent to news and information. Thus, I rarely watch television news by myself. I like a news channel that I'm used to.	X
News briefs	Owing to busy schedule, I watch television news when I'm available. Thus, I prefer news in brief.	
Actively engaged	With my own perspective, I actively watch news judging its objectivity, fairness, accuracy, and depth.	
Attention to periphery	Rather than news content, I'm more interested in peripheral cues like anchor's or reporter's feature or voice. I'm more influenced by visual effects like studio set, graphics, and so forth.	

Results

Test of Agreement in the Typal Proportion

Table 9 shows very similar proportional distribution of three typologies in fruit preference by two assessment tools. Type C appears as the biggest portion followed by Types B and A, respectively, in both measurements. Incidentally, Type B achieves 24.4 percent measured by the two tools. Type A in Q Block and Type C in Q Tool had larger estimates when compared with the other tool. These results imply that the assessment tools are reliable in predicting the relative proportion of the typologies.

Table 10 supports the same result: in both the tools used in the TV news selection study, the proportion of Type B appeared the biggest share, 41.7 percent each, and then Type C, Type A, Type D followed, respectively. Even though the percentage of each cell in the four types does not look very similar, the proportional order in both the measurements appeared identical, which means high correlation between the two.

Table 8: Example of Q Block for TV News-Viewing

Type and Statement	*Most agree*	*Agree*	*Disagree*	*Most disagree*	*Viewing Types*			
					A	*B*	*C*	*D*
A I just watch TV news that I am used to.	4	3	2	1				
B I don't like the anchor's exaggerated or sensational expression.	4	3	2	1				
C I prefer a news channel that reports something that others don't.	4	3	2	1				
D I watch TV news where its anchor looks comfortable.	4	3	2	1				
B I like to watch brief news airing between programs.	4	3	2	1				
C I believe that the news I'm watching is more complete and communicative.	4	3	2	1				
D The news that I mainly watch has many outstanding reporters.	4	3	2	1				
A I usually watch news channel that I am used to. I feel strange when I watch a new channel.	4	3	2	1				
C The news I usually watch is relatively fair.	4	3	2	1				
D I am impressed with news I usually watch because of good visual presentation.	4	3	2	1				
A I watch news continually after soup operas.	4	3	2	1				
B I understand what's going on from scroll news.	4	3	2	1				
				Type Scores				

Table 9: Estimated Proportions of Q Tool and Q Block of Fruit Preference Typology

	Q Tool	*Q Block*	*Average (%)*
Type A	25 (19.7%)	40 (31.5%)	25.56
Type B	31 (24.4%)	31 (24.4%)	24.4
Type C	71 (55.9%)	56 (44.1%)	40.0

Table 10: Frequency of Q Tool and Q Block of TV News-Viewing

	Q Tool	*Q Block*	*Average (%)*
Type A	33 (26.0%)	23 (18.1%)	22.05
Type B	53 (41.7%)	53 (41.7%)	41.70
Type C	36 (28.3%)	35 (27.6%)	27.95
Type D	5 (3.9%)	16 (12.6%)	8.25

Test of Agreement in the Typal Identification

As shown in Table 11, the shaded cells show the same estimate by the two measurements, and the sum of the shaded cells divided by the total valid cases, 127 in this example, indicates the degrees of agreement. In other words, the agreement between the two measurements in its typal identification would be calculated by the number of same classification divided by the total cases. Thus, the agreement here is:

$$\left(\frac{15+20+44}{127}\right) = \frac{79}{127} = 0.622$$

Table 11: Agreement Matrix for Fruit Preference

		Q Block			*Total*
		Type A	*Type B*	*Type C*	
Q Tool	*Type A*	15	4	6	25
	Type B	5	20	6	31
	Type C	20	7	44	71
Total		40	31	56	127

This relatively low (62%) agreement might be attributed mainly to the Type A assessment: 15 cases fall in Type A by both measurements leaving 5 and 20 cases in Type B and Type C. Out of 40 cases, only 15 cases were agreed as Type A. 25 people assessed as Type A by Q Block were differently assessed as Type B (5) and Type C (20) by Q Tool.

Meanwhile, Table 12 shows higher agreement in the study of TV news-viewing. Out of 127 total cases, agreed number in the typal identification is 89 (19+40+25+5), and, thus, the agreement rate is

$$0.7008 \left(= \frac{89}{127}\right).$$

This agreement appeared higher than that of the fruit study even though TV news study has more categories, 4, compared with 3.

Table 12: Agreement Matrix for TV News-Viewing

		Q Block				*Total*
		Type A	*Type B*	*Type C*	*Type D*	
Q Tool	*Type A*	19	4	6	4	33
	Type B	3	40	4	6	53
	Type C	1	9	25	1	36
	Type D	0	0	0	5	5
Total		23	53	35	16	127

However, this agreement between the two measurements can occur by chance, and we have to consider that this chance depends on the number of types we are trying to identify. For example, a three-type system should obtain 1/3 (33.33%) agreement by chance; a four-type system would generate a 1/4 (25%) agreement by chance as shown in this study and so on. The percentage of expected agreement can be calculated by the sum of the squared percentages of all types.

In the fruit preference study, % expected agreement = $(1/3)^2 + (1/3)^2 + (1/3)^2 = 1/3$ (0.33), and, in the TV news-viewing study, % expected agreement = $(.25)^2 + (.25)^2 + (.25)^2 + (.25)^2 = 0.25$. Thus, Pi can be calculated in both studies as follows, respectively:

$$\text{Pi (1)} = \frac{0.622 - 0.33}{1 - 033} = 43.58(\%)$$

$$\text{Pi (2)} = \frac{0.7008 - 0.25}{1 - 0.25} = 60.11(\%)$$

Discussion and Conclusion

This study examines the agreement between Q Block and Q Tool in its typal identification through two cases, one having three types and the other having four types. Q Block is basically constructed with blocks consisting of Q items having discriminating power, and identification process is identical to Q sorting. It is a simplified Q sort, in this sense. Therefore, to construct Q-Block sets, researchers need enough Q items, say more than 40; otherwise there might be only several items to be selected for the construction of Q blocks. And, if there are high correlations between the factors found, it is not easy to select the items having factor-discriminating power. Another limit of Q Block would be

that respondents should sort the items even if it is simple, and that, from the statistical point of view, people who are loaded with low factor weight are more likely to be identified in wrong typologies, or it leads the tie scores in typal identification as shown in this study: 23 out 150 cases (15.3%) were revealed as ties by either one of the measurement.

Meanwhile, Q Tool would be comprehensively constructed after considering all the information from theory, interviews, demographics, questionnaire as well as Q-factor arrays. Q Tool is made through the interpretation of Q factors, which is primarily abductive. Since Q Block mainly includes the Q items from the arrays of typal comparison, the input data for the assessment tool is limited in nature. In that sense, Q Tool probably has higher validity than Q Block. Respondents are assessed by Q Block operationally, while respondents decide their own typology after reading the synopsis of the factor interpretation, operantly. Thus, typal identification by Q Tool is mutually exclusive and decisive in most cases, while typal identification by Q Block cannot always discriminate a type because of tied type scores.

However, this study does not guarantee that Q Tool has a sufficient validity to identify the typology, even though it has several advantages such as convenience, comprehensiveness, and decisiveness. Researchers who want to use Q Tool for further study should be cautious, and keep in mind that the interpretation of Q factors requires insight and abductive mental framework based on the data collected. In Q Tool, the validity wholly depends on the factor interpretation.

The two cases analyzed in this study reveal that the agreement between Q Block and Q Tool was not high enough, but the typal proportion assessed by the two measurements exactly agreed. Tests of reliability and validity would be continually conducted through the agreement test between Q Tool and Q Block, particularly under the conditions of various combinations of Q samples.

Note: *This article is a revised version of a paper presented at the 24th Annual Conference of the International Society for the Scientific Study of Subjectivity, Hamilton, Ontario, Canada, Oct. 2–4, 2008. An earlier version appeared in 2008 as "Agreement between the Q-Block and the Q-Tool",* Q Methodology and Theory, 16, 5–*16 [in Korean]. This work was supported by Hankuk University of Foreign Studies Research Fund of 2009.*

References

Holsti, O. (1969). *Content analysis for the social science and humanities.* Reading, MA: Addison Wesley.

Hyundai Research Institute (2005). A report of market segmentation for KT BcN Service.

Kim, H. K. (2008). *Q methodology: Philosophy of science, theory, analysis and application.* [In Korean]. Seoul: Communication Books.

Kim, H. K. (2002). An exploratory study of aroma preference. *Korea Japan Congress of Aromatherapy, 2.*

Kim, H. K. (1999). The characteristics of the Korean consumer by value and lifestyle typology. *The Korean Journal of Advertising, 10*(2), 173–197.

Kim, S. H. (2008). Television news-viewing patterns and characteristics. Doctoral dissertation, Hankuk University of Foreign Studies Graduate School, Seoul.

Scott, W. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly, 17,* 321–325.

Talbott, A. D. (1963). The Q-block method of indexing Q typologies. Presented at the AEJ Conference, Lincoln, Nebraska. [reprinted in this issue, pp. 6–24]

Wimmer, R. D., & Dominck, J. R. (1987). *Mass media research: An introduction,* (2nd Ed.). Belmont, CA: Wadsworth