Q methodology to determine distinguishing and consensus factors (a case study of university students' ecoliteracy on disaster risk reduction)

Aldila Rahma^{1,2*}, Djati Mardiatno³, and Dyah Rahmawati Hizbaron³

¹Environmental Science Program, Universitas Gadjah Mada, Yogyakarta, Indonesia

²Faculty of Teacher Training and Education, University of Islam Nusantara, Bandung, Indonesia

³Faculty of Geography, Universitas Gadjah Mada, Yogyakarta, Indonesia

Abstract. Q methodology is to measure the diversity of human subjectivity, such as perspective, mindset, preferences, opinions, and attitudes. It is one of the approaches applied in mixed methods research. This paper is part of a preliminary study that aims to examine the application of Q methodology. Fourteen undergraduate students from the Faculty of Geography, Universitas Gadjah Mada, participated in this study. Data collected with Q sorting techniques using the Q-Software, and analyzed with Principal Component Analysis using KADE desktop. Both are software used specifically in Q methodology research. The method implementation aims to determine the distinguishing factors and consensus of students' ecoliteracy components on disaster risk reduction. The research showed that Q-Software tends to facilitate researchers in collecting data. However, it is quite complicated and less practical for participants. Meanwhile, the KADE desktop has an attractive interface with a tutorial that makes it easy to use. By analyzing the case study, the result showed that the students' ecoliteracy component is represented by five distinguishing factors and two consensus statements, with cumulative variations explained by 68 %. This study is, however, limited by the performance of quantitative procedures without in-depth qualitative data interpretation.

Keywords: Consensus, disaster risk reduction, distinguishing factors, ecoliteracy

1 Introduction

In 1935, William Stephenson discovered Q methodology. Traditionally, the use of Q methodology is often associated with quantitative methods because of factor analysis involvement [1]. Furthermore, this method receives much attention in social research because it has a strong statistical component to examine human subjectivity. It plays a major role in combining qualitative and quantitative techniques as well as also one among the significant approaches for mixed methods research [1–3].

Q methodology has been widely used since 1985 in psychology, communication, political science, health, environment, and other related fields [4]. It measures the diversity of subjectivity that represents an individual's feelings, experiences, opinions, perspectives, or preferences. It also identifies similarities, investigates the construction of a phenomenon, and the relationship between categories in a concurrent manner [5]. Yet, this method does not aim at measuring sample or population statistically but rather exploring the possibilities of various perspectives and consensus within the group about specific topics being investigated [6].

What This method does is clustering people and not clustering the variables. Therefore, the researcher must

have sufficient variables to find differences between participants, rather than aiming for a sufficient number of participants to determine the difference between variables as well as in factor analysis [3]. In the factor analysis, the number of samples is useful for controlling errors, while perspective is considered as an external component [7]. Q method, however, only utilizes a small number of participants in order to provide various viewpoints, comments, and opinions from various perspectives [5].

For the application, this method requires a Q grid for placing unit statements according to the participants' preferences. Additional tools are required, either traditional paper-based or software-based. Traditional Paper-based Q Methods (TPQM) typically using cartons and cards with statements [8]. Various studies with Q method approaches have utilized software-based tools such as Visual Q method (VQMethod), Q-Assessor, and FlashQ software for collecting data [8–10]. For data analysis, many studies employ different software such as QUANAL, PCQ for Windows Academic Edition, PQMethod, Ken-Q Online Analysis, and KADE Dekstop [6, 11,12].

*Corresponding author: aldila.rahma@mail.ugm.ac.id

This paper is a preliminary study aimed to examine the Q methodology application to determine the distinguishing factor and consensus using university students' ecoliteracy as a case study. It includes the examination of using Q-software and KADE as software for data collection, and data analysis, respectively. However, the qualitative procedure is not included in this study. Hopefully, this study provides additional methodology information for research in the environmental science and disaster management field.

2 Research methods

2.1 Data collecting technique

This research was carried out in December 2019 – April 2020. Fourteen undergraduate students from the Faculty of Geography, Universitas Gadjah Mada, took part in this study. In the study, Q Methodology employed generally consists of the following six steps [13, 14]:

1. Developing a comprehensive set of statements (Q set) from various sources related to disaster risk reduction, climate change, and ecological understanding topics. This set of statements called concourse [15]. A total of 52 statements were arranged related to these topics, where the statements must be "opinion" and not a fact [4].
2. Sampling (P-set). Q participant is a person selected to do the sorting process. Unlike the respondents represent the population, Q participants were chosen to represent the breadth of opinions in the population target. The most important thing is that the number of Q participants should be fewer than the Q statement. It usually uses a ratio of 3:1, as in three statements require only a single participant.
3. Statements distribution (Q sorting). Each participant was asked to distribute statements into the Q grid, based on their preferences. The ranking process is similar to the Likert scale format. The Q sort design is shaped like an inverted pyramid and applying forced distribution and using quasi-normal distribution (Figure 1) [14]. This step used *Q-Software* online software [16]. Participants were invited by e-mail.

Fig. 1. Q sort table (Q grid) contains 52 columns according to the number of unit statements

4. Conducting interviews with selected participants who have the "strongest" opinions. It is to explore and obtain a deeper understanding of the Q sort profile.
5. Analyzing and interpreting the data obtained.

2.2 Data analysis technique

Data from Q sorting were analyzed using KADE (Ken-Q Analysis Desktop Edition version 1.1.0) [17]. The steps are as follows:

1. Inputting Data
2. Correlation Matrix
3. Extraction Factor (PCA/ Principal Component Analysis)
4. Rotation factor (Varimax rotation)
5. Loadings Factor
6. Output Data

This technique requires factor analysis to analyze groups of people who share their views. Each factor represents a unique viewpoint out of the research topic. Quantitative parts in Q methodology depend heavily on factor analysis and factor rotation. Factor analysis represents the heterogeneity of participant viewpoint. The final step is interpreting interviews result of participants. The results are verified, disassociated, and communicated to the related parties [18].

The reliability test of the Q method usually using a replicability type. However, because it does not aim to generalize the population, the reliability does not become a concern in this method [6, 8]. This study only conducted a single measurement or one-shot. As for PCA, besides being used for factor analysis, it can also be used for the validity test. For one factor to be distinguished by another, the correlation coefficient score between them should be low [19].

3 Result

3.1 Q-Software for collecting data in Q methodology

Collecting data (Q Sorting) in this study was using software called *Q-Software*. The researcher created and designed procedures in advance using the software and then sent the procedures and instructions to each participant's e-mail address of the invited participants, after which the participants obtained a personal link to do O sorting.

There are two main steps in the sorting process using Q-sortware. First, participants classified statements of agree, neutral, and disagree. The aim is they took a "practice" before doing the actual Q sorting.

According to Figure 2, the screen consists of an instruction column on top and statement column below. The statements (in the yellow box) appeared one by one, and then participants were to drag statements one by one to the column below. After all the statements already moved, participants clicked the "continue" icon and proceeded to the second step.

Fig 2. Screenshot of Q-Software first step

In this step, participants sorted the statements that have been classified previously (disagree, neutral, agree), into more detailed categories. The procedure is still the same as the previous step. Figure 3 is a screenshot of Q-Software second step. After finishing the entire sorting process, participants clicked the "continue" icon, and data would be sent immediately to the researchers' account. Data from Q-Software will be exported in .csv type files and analyzed using KADE.

Fig 3. Screenshot of Q-Software second step

3.2 KADE desktop for data analysis in Q Methodology

The data analysis process using KADE is rather simple. Figure 4 shows the interface of KADE. During the process, the researcher needs to input data files and follow the steps. KADE will calculate the correlation. Similar variables represented by correlations between them [20]. A high correlation score will construct a group of factors.

Fig 4. Screenshot of KADE ver. 1.1.0

The second step is factor extraction. KADE provides Centroid and PCA (Principal Component Analysis) methods that commonly used in Factor analysis (R). The Centroid is a preferred extraction method that is not too familiar in general statistical analysis [21]. Traditional factor analysis techniques as Centroid did not automatically provide mathematical solutions, so PCA was developed to assist [22]. The result of PCA is a group of factors. For the optimum model, the percentage of variance must be higher than 60. However, the percentage limit of variance for social research is acceptable between 50 % to 60 %.

The third step is the rotation. There are two rotation options in KADE, varimax, and hand rotation (judgmental). Varimax was chosen because it allowed the researcher to produce simple structures and maximize Eigenvalue for each factor. It can maximize the amount of variance from the factor extraction [22].

The fourth step is factor loading. The researcher selected individuals who can represent each factor called flagging. Representative individuals marked with (X) (correlation score > 0.5). Factor description and analysis are determined by participants who were flagged on certain factors.

Output is the final result of the analysis process. It shows the tables of factors. Researchers conducted interviews with representatives' participants to explain the reasons why they arranged those statements. Lastly, the researchers interpreted each factor and coded it.

3.3 Case study analysis results

The results of the university students' ecoliteracy case study analysis using KADE are shown as follows. Correlation score describes variables that have similarities—high score causing these variables grouping and construct factors [20].

Table 1. Unrotated Factor Matrix

	F1	F 2	F 3	F 4	F 5
Eigenvalues	4.19	1.67	1.35	1.22	1.0009
% Explained Variance	30	12	10	9	7
Cumulative % Expln Var	30	42	52	61	68

Table 1 is the result of the PCA analysis where maximum number of factors that can be extracted are eight factors. Reducing and simplifying multivariate techniques data is applied in PCA. Thus, PCA extracts essential information from data tables and also analyzes observable structures of variables [12,22].

PCA was analyzed based on the Eigenvalue. It indicates how many variables can explain a factor. The default value of Eigenvalue is set on > 1 (less than 1 is ignored because it cannot construct a factor) [4,18, 22]. Therefore, table 1 only shows a factor 1 to 5. The cumulative explained variance is 68 %. Five distinguishing factors represent the ecoliteracy component of the students.

Varimax rotation was applied in this study. Factors that can be rotated by Varimax have eigenvalues > 1 . Varimax is maximizing Eigenvalue on each factor [15]. Table 2 shows a result of varimax rotation.

Table 2. Participants' flagging

*P	*F 1	F 2	F 3	F 4	F 5
P1	0.6608 X	-	0.1217	-0.1523	0.2758
P2	0.5847 X	0.5354	0.0469	-0.1612	0.1095
P3	0.1909	0.1519	0.299	-0.3928	0.608 X
P4	0.2613	0.1736	0.7348 X	-0.2034	0.061
P5	0.4617	0.1351	0.2819	-	0.6084 X
P6	0.0864	-	-0.0988	0.9019 X	0.0718
P7	0.804 X	-	0.1255	0.1116	-
P8	-	0.06	0.7153 X	-0.1188	-
P9	0.2228	0.5192	0.5777 X	0.0891	0.1781
P10	0.1424	-	0.7178 X	-0.1534	0.1838
P11	0.1491	0.6233 X	0.0575	-0.1757	0.3962
P12	-	0.8069 X	0.0812	-0.0513	-
P13	-	0.1133	0.0137	0.1286	0.8758 X
P14	0.4148	0.2254	0.5737 X	0.126	-
*% Exp var	15	13	18	11	11

*P (Participants), F (Loading Factor), % Explained variance

Flagging (X) is for selecting representative participants of each factor. Auto Flagging is based on the high significance of loading factor criteria ($p < 0.05$) [12]. Selected participants are those with scores > 0.5 because it meets the requirements of commonality based on the assumptions of factor analysis [24]. Loading factor is simply the correlation coefficient. A higher score of loading factor indicates a high correlation of participants with their represented factors [7].

3.4 Distinguishing statement

The final results of the factor analysis display the distinguishing statements. According to table 3, the relationship between factors is less than 0.399. Based on the Pearson r correlation coefficient classification, the coefficient interval between 0.2 to 0.399 indicates a low relationship level [24]. Also, a negative mark (-) indicates a negative correlation. The correlation between factors should not be high, because of the higher correlating factors, the lesser distinguishing factors (which can distinguish between factors) [19].

Table 3. Correlation Value Between Factors

*	F1	F2	F3	F4	F5
F1	1				
F2	-0.084	1			
F3	0.3517	0.2234	1		
F4	-0.0717	-0.1564	-0.2568	1	
F5	0.1271	0.2288	0.1994	-0.0609	1

*F: factor

The composite reliability of each factor is above average (0.8). Factor 3 shows the highest score (0.95). A high-reliability above the average indicates the instrument is consistent and reliable. However, validity and reliability, not being a concern in Q because participants' subjectivity is more emphasized in this method [6]. Table 4 compares the characteristics between factors. Five participants represented factor 3. The factor is determined by some participants who "load" in it. One factor that has at least 5 participants who define it, then the reliability, can reach 0.95 [19].

Table 4. Factors Characteristics

	F1	F2	F3	F4	F5
No.of defining Variables	3	2	5	2	2
Avg. Rel.Cof.	0.8	0.8	0.8	0.8	0.8
Composite Reliability	0.923	0.889	0.952	0.889	0.889
	0.277	0.333	0.219	0.333	0.333

Output analysis also shows a table of distinguishing statements ($p < 0.05$). The table will later be useful in interpreting the results. Due to page limitations, this paper only provides 1 example out of 5 total tables.

Table 5. Distinguishing Statements Factor 3

No	Statement	Q Sort value	Z-score
6	I feel the impact of climate change on the environment around me and for myself	5	2.19*
18	The government is the most responsible for conserving nature, overcoming pollution, and taking action to increase ecological awareness	-1	-0.47*
23	Indonesia is a country that is rich in natural resources, but many people still live below the	-2	-0.93

No	Statement	Q Sort value	Z-score
	poverty line. It is part of government responsibility		

Note: (*) significance at $p < 0.01$

Z score indicates the relationship between statement and factor

Table 5 is the distinguishing statement of factor 3. Z score describes the place of each statement in the Q sort table [3]. The most extreme z value (its position increasingly towards the edge of Q sort table) is the most useful in interpreting factors. According to Table 5, 3 statements distinguish this factor from another. A total of 5 participants agreed with statement 6 and disagreed with statement 18 and 23.

3.5 Consensus statement (Agreement)

Consensus statements are useful to determine statements agreed by all participants. Consensus provides insight into how participants perceive the problems, express their logic, and put the essential issues in the context. This method intentionally involves people with different points of view to fill the vital information gap [14].

Table 6. Consensus Statement

Statement 10: <i>I must bequeath a proper, healthy and safe environment for the future generation</i>									
Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
Q	Z	Q	Z	Q	Z	Q	Z	Q	Z
3	1.16	3	1.17	3	1.21	2	0.68	4	1.44
Statement 15: <i>In my opinion, disaster mitigation knowledge must become a compulsory curriculum in every level of the education system.</i>									
Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
Q	Z	Q	Z	Q	Z	Q	Z	Q	Z
2	0.66	2	0.86	1	0.40	2	0.55	3	1.18

Note Q (Q value Sort), Z (Z value).

(*) were not significant at $p < 0.05$ or $p < 0.01$

According to table 6, two statements become consensus. All participants agreed to statements 10 and 15. These two statements were not statistically significant at both $p < 0.01$ and $p < 0.05$.

4 Discussion

Q statements usually consist of 30 to 60 units as representative of a concourse [3]. The concourse obtained from primary sources (direct conversations, participant interviews, and observations) and secondary sources (literature, popular media) [4, 5]. It depends on the researcher's objectives and considers whether it can explain the perceptions of all participants. The unit statements in this study were arranged to explain students' ecoliteracy components on DRR. It consists of affect, knowledge, cognitive, and behavior components [25].

P sets or participants in the Q method, are not the same as the "sample size" in studies with other methods. Participants are chosen because they have a different and well-formed opinion. People with a well-formed

opinion will find it easier to do Q sort and tend to produce reliable results [13].

Q-Software is a software created by Alessio Pruneddu, with intellectual property rights belongs to the University of York United Kingdom [16]. The researcher made registers first on Q-Software website to create an account, then arrange the instrument. Using web-based software ease the researchers in the sorting process, rather than using the Traditional Paper-Based Q Method (TPQM). Practically, TPQM requires direct meetings with participants. Also, it has greater difficulty and requires more resources [8].

On the other hand, Q-Software helps researchers to collect data more efficiently. Once the participants finished the sorting, data can be submitted immediately to be analyzed. Researchers can also get information on participants who have completed the sorting, date of data submission, and time needed for the process.

The disadvantage of using Q-Software is that it may be rather confusing and impractical for some people. This software is web-based, so it requires stable internet access. It is also in the form of a desktop-based software that cannot be operated on a smartphone. In operating the instrument, it requires participants to drag statements into columns. This process cannot be done on a smartphone, where people more often use it for practical reasons. Participants have to use a laptop or PC to conduct the sorting. This software also has a slightly complicated interface. The instruction column is too small, so that most of the participants might not be aware and tend to miss the instructions. It merges instructions and sorting columns in only one screen page. Figure 3 gives the appearance of the interface, which is too "crowded."

Participants also met difficulties in the sorting process. They never did sort with forced distributions before. They were more familiar with the open questionnaire method, such as Likert scale or checklist questionnaire. So even though they were doing online sorting, the researcher must watch the process to minimize the error. The obstacles like this become a challenge in collecting data using Q-software.

Other software that can be used for sorting procedure is Visual Q method (VQMethod). VQMethod allows the addition of new dimensions such as video and audio replacements for statements. VQMethod is more thorough and provides more reliable data compared to TPQM [8]. Another study was carried out using Q-Assessor's paid licensing software. It has complete validation features. This software allows the researcher to send instructions to ease the sorting steps. The researcher obtained data through an e-mail sent by participants [9]. Another study used FlashQ software. This software also consists of a 2-stage process like Q-Software, but it looks much simpler. Participants can directly arrange statement cards into the Q sort table as if they were arranging a puzzle [10]. However, FlashQ software requires particular expertise in creating HTML (HyperText Markup Language) versions. So, it can be concluded that using software is helpful to facilitate the sorting process and making it more interactive.

The data analysis process in this study used KADE. KADE is a desktop-based Q Methodology analysis

application that can run in Microsoft Windows, Apple OS, and Linux OS. KADE interface looks much simpler and easier to operate and also provides interactive features that are helpful for the analysis process[17]. It also provides templates, demo data, and manuals, to help researchers understand its use. This software does not require an installation process (portable) and free from the internet. KADE only analyzes statistically, and the results are in the forms of numbers. The researchers need to interpret the table meaning for qualitative analysis. There is no specific strategy to interpret the formed factors as it depends on the researcher's objective [19].

Other commonly used software for Q method analysis is QUANAL, PCQ for Windows Academic Edition (paid license), PQMethod, Ken-Q online Analysis, and KADE Desktop edition [11]. Among those software, PQMethod and KADE are often used extensively for data analysis. PQMethod has the same analysis steps as KADE, but its operating system based on DOS [6, 12].

Q methodology can determine variations in students' epistemological views. It can also avoid a lengthy interview process, which is typical of qualitative research [15]. This study helps to construct the concourse from the statements. Data interpretation is a part of the hermeneutic process involving an interpretive perspective from the researcher. However, it is still limited by the participant's subjective factors during the sorting process [5].

The result of the Q method is that there are diverse points of view on a topic that may be operant or measurable. The results cannot become a representative of the sample or population. Reliability statistics work to generalize samples to be representative of a population, not become a concern in this method [7].

5 Conclusion

Q Methodology is useful in research involving subjectivity, especially in social research. The steps start from arranging unit statements, determining participants, sorting process, and data analysis. The analysis process consists of correlation, factor analysis, rotation, loading, flagging, and data interpretation steps.

Q-Software facilitates the researcher to collect participant data. However, for the participants, it is quite tricky, less practical as it cannot be operated on a smartphone, and has a "crowded" looks interface. Meanwhile, KADE has a simple interface. KADE also has an easy-to-understand demo data and manual, making it easier for the researcher to do data analysis.

According to the study, there are five distinguishing factors with the composite reliability of each factor between 0.89-0.95 and also two consensus statements agreed by all participants. Cumulative explained variance of extraction factors is 68 %, which is acceptable for social research.

This preliminary study is limited to the procedure for operating Q methodology and its statistical analysis. Also, it reveals a problem that occurs in the process. Factors interpretation, which is the part of the qualitative

analysis, does not include in this paper. The researchers recommend further research to conduct replicability tests to ensure the reliability of Q Methodology. Besides, use comparisons with other tools, either traditional paper-based or software-based Q method for Q sorting process.

Authors would like to thank Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Republic of Indonesia for partially funding this research through Beasiswa Unggulan Dosen Indonesia (BUDI) program.

References

1. S. R. Brown, *Qual. Health Res.*, **6**,4:561–567(1996), doi: 10.1177/104973239600600408.
2. N. Akhtar-Danesh N. Mirza, *Open J. Soc. Sci.*, **5**,3:198–204(2017), doi: 10.4236/jss.2017.53017.
3. A. Tashakkori, C. Teddlie, I. Newman, S. Ramlo, *Using Q Methodology and Q Factor Analysis in Mixed Methods Research*. (2015).
4. S.R. Brown, *Operant Subj.*, **16**,3/4:91–138(1993).
5. P. Shinebourne, *Int. J. Qual. Methods*, **8**,1:93–97(2009), doi: 10.1177/160940690900800109.
6. S.E. Ramlo, *AIP Conf. Proc.*, **1064**:179–182(2008), doi: 10.1063/1.3021248.
7. J.B. Nicholas, *Reliability in Q Methodology: A Case Study*, In: *The Eastern Education Research Association Annual Conference*, 1–20 (2011)
8. S. Nazariadli, D.B. Morais, S. Supak, P.K. Baran, K. S. Bunds, *Methodol. Innov.*, **12**,1, p. 205979911983219, (2019), doi: 10.1177/2059799119832194.
9. B.H. Reber, S.E. Kaufman, F. Cropp, *Operant Subj.*, **23**,4:192–209(2000).
10. P. Delaitre, C. Lavandier, R. Dedieu, N. Gey, *Meaning of quiet areas in urban context through people viewpoints*, In *the Acoustics 2012 Nantes Conference*, 2133–2138 (2012).
11. P. Schmolck, *PQMethod Software*, 2020. [Online]. Available: <http://schmolck.org/qmethod/>. [Accessed: 08-Mar-2020].
12. A. Zabala, *RJ.*, **6**,2:163–173(2014), doi: 10.32614/rj-2014-032.
13. L.S. Bashatah, *J. Res. Method Educ.*, **6**,5:37–43(2016), doi: 10.9790/7388-0605053743.
14. T. Webler, S. Danielson, S. Tuler, *Using Q Method to Reveal Social Perspectives in Environmental Research*. Greenfield, MA, (2009).
15. I. Newman, S. Ramlo, *Using Q Methodology and Q Factor Analysis in Mixed Methods Research*. Thousand Oaks: SAGE Publications, Inc., (2015).
16. A. Pruneddu, *Q-SORTWARE*. [Online]. Available: <http://www.qsoftware.net/home.html>. [Accessed: 16-Dec-2019].
17. S. Banasick, *J. Open Source Softw.*, **4**, 36:1360(2019), doi: 10.21105/joss.01360.
18. M. Nijnik, D. Miller, A. Nijnik, *J. Settlements Spat. Plan.*, **2**:185–190(2013).
19. T.C. Du Plessis, *A Theoretical Framework of Corporate Online Communication: A Marketing*

- Public Relations (MPR) Perspective*, University of south Africa, (2005).
20. H. Abdi L. J. Williams, Wiley Interdiscip. Rev. Comput. Stat., **2**,4:433–459(2010), doi: 10.1002/wics.101.
 21. S. Ramlo, Mid-Western Educ. Res., **28**,1:73–92(2011).
 22. S. Watts P. Stenner, Qual. Res. Psychol., **2**,1:67–91(2005), doi: 10.1191/1478088705qp022oa.
 23. S. P. Mishra et al., International J. Livest. Res., **7**,5:60–78(2017), doi: 10.5455/ijlr.20170415115235.
 24. A. Hidayat, *Asumsi Analisis Faktor dengan SPSS*, Statistikian, 2014. [Online]. Available: <https://www.statistikian.com/2014/03/asumsi-analisis-faktor-dengan-spss.html>. [Accessed: 08-Mar-2020].
 25. B.B. McBride, C.A. Brewer, A.R. Berkowitz, W.T. Borrie, Ecosphere, **4**,5(2013), doi: 10.1890/ES13-00075.1.