DOI: 10.47743/ejes-2023-0201 • December 2023 • VOLUME 14 ISSUE 2

Generalized Q analysis as a new tool in social science research – a pedagogical introduction

Tomaz Ponce Dentinho[a*], Karima Kourtit[b], Peter Nijkamp[b]

[a]University of the Azores, Portugal; [b]Open University, Heerlen, The Netherlands

Abstract
Q analysis is a frequently used, multivariate exploratory technique in the social sciences which seeks to identify and analyse commonalities and differences in the respondents' rankings of a set of relevant individual qualitative statements. It has some intrinsic weaknesses which may be overcome by widening and deepening the underlying multivariate statistical approach. This new technique, termed a Generalized Q method, is able to handle an enlargement of the number of ranked combined statements based on a structured re-combination of the rankings of simple statements. This method will be presented in this paper. We employ a simple illustrative data set in our study. In comparing the traditional and the Generalized Q analysis, the same questionnaire data are used to apply and assess both methods, to compare the results, and to highlight the advantages of the Generalized Q method. We find that the latter technique is able to take account of many simple questionnaires with trustful responses, allows for an expansion of the number of respondents, facilitates the naming and interpretation of the extracted multivariate components, and is able to test the consistency of the responses.

Keywords: Q analysis, qualitative analysis, generalized Q method, Principal Component Analysis

1. Setting the scene

The social sciences (e.g. economics, geography, psychology) have over the past decades shown a rapid transition from traditionally qualitative research methods to quantitative analysis techniques. This 'quantitative revolution' comprises of various statistical methods, such as regression analysis, multivariate statistical analysis, logistic regression, principal component analysis, and so forth. All these methods serve to 'measure the unmeasurable' (Nijkamp, 1985), and are also increasingly employed in the spatial sciences (geography, regional science, transportation science, ecology, political science etc.). An important part of quantitative research techniques in the social science is devoted to comparative studies, in which

*✉ Tomaz Ponce Dentinho, Professor, University of the Azores, Portugal; email: tomas.dentinho@uac.pt.

numerical indicators of relevant subjects or objects are examined in order to trace or assess the relative performance of such phenomena. Examples can be found in correspondence analysis, similarity analysis, multi-criteria analysis or data envelopment analysis. The social sciences have indeed become largely quantitative data analysis engines. An important class of relative performance analysis is formed by Q analysis, which aims to identify common patterns in statements by respondents or actors. The aim of the present contribution is to depict briefly the state of affairs around Q analysis and next to lay down the foundations of a new – generalized – Q analysis and to show its relevance for social and spatial sciences by means of some illustrative examples.

Q analysis is a research technique used in social sciences to analyse the commonalities and differences in the respondents' subjective points of view on a topic of their concern. It was initially developed by Stephenson (1953) and subsequently often used to explore research fields like educational attitudes (Gawron 2016), auto-ethnographic issues (Ellis, 2004; Pepeka et al., 2022), credibility studies (Metzger & Flanagin 2013), healthcare studies based on survey data (Churruca et al., 2021), job satisfaction (Guastello et al., 2019), urban sustainability (Fuentes et al. 2021). Nowadays we witness a rising set of new fields that use the method to transform subjective evaluations into objective results.

The traditional Q method involves: (1) a collection of qualitative statements on a topic of the respondents' concern; (2) a systematic ranking of many disagreements that follow usually an approximated normal distribution; (3) a transposition of the collected data defining individuals ('respondents') as key variables and subjective statements as observations; (4) an implementation of Principal Component Analysis to reduce the responses profiles into synthesised and orthogonal responses, and (5) an analysis of the synthesised orthogonal responses relating them with the typology of statements and with the respondents' features. This may then lead to interpretable outcomes.

We note that the Q method differs from the prevailing R method in multivariate analysis in the social sciences where individuals are the units of analysis rather than variables in a transposed matrix allowing the possibility to study subjective phenomena such as opinions, attitudes, and values of respondents (McKeown & Thomas, 2013).

There are three main limitations inherent in the application of a traditional Q method. First, it assumes that respondents are able to rank many statements which, according to (Miller, 1956), is not plausible ('the magical number seven'). Second, the number of non-redundant respondents (variables) is limited by the number of statements (observations), constraining therefore the number of respondents and their relative representativity. Finally, the traditional Q method does not provide objective information to name the extracted multivariate attitudes, so that the results can lead to different interpretations (Brown, 1993).

The traditional Q method places much emphasis on the meaningful creation of relevant statements and their combinations. The idea to work with composed statements associated with the Q methodology has been advocated mainly to obtain an appropriate set of statements, normally around 30, originating for instance from five responses to six consistent questions on a given topic (McKeown & Thomas, 2013). This number of combined statements is still limited, however. The new Generalized Q methodology proposes, instead of getting 6x5= 30 statements, to create 6^5= 7776 combined statements from the same data set.

The history of Generalized Q analysis is very recent. The idea came from the application of a data-analytic requirement to apply the conventional Q method to study the attitudes of people from the Indonesian island of Timor Leste regarding the occurrence of natural hazards related to climate, based on the standard Q method but with 500 questionnaires (D'Haeyer et al., 2022). The research question seemed strange, because in that case there was the need to find more than 500 statements to support the use of the Q method exercise for the 500 respondents or otherwise to promote a hierarchical Q analysis (Sánchez et al., 2021), where the components related to the lower level of a territorial unit become the Q sorts for higher levels of territorial aggregation. The solution, that may lead to the creation of the Generalized Q method, was different: with a limited set of statements, viz. 6 hazards + 8 assets + 6 responses + 6 expectations = 22 simple statements, on climate hazards, it was possible to get 6 x 8 x 6 x 6 = 1728 combined responses enabling the Q method exercise to be pursued for the sample of 500 questionnaires which allowed to find two characteristic prominent attitudes in Timor Leste regarding natural hazards. The Generalized Q method was also used in other recent studies (Dentinho, 2023; Simionov, 2023).

Thus, Q analysis allows the enlargement of the number of ranked re-combined statements on a relevant topic based on a structured re-combination of the rankings of simple statements, assuming respondents are consistent in their sequential rankings of simple statements. The new Generalized Q method has important benefits, because: (1) it allows an expansion of the number of respondents by overcoming the redundancy of many respondents in the usual Q Method; (2) it facilitates the nomenclature of the extracted components that are representative responses, and (3) it allows to test the consistency of the various responses.

This short paper seeks to systematize advances in multivariate Q methods, so that it can be used by a wide range of social scientists interested in understanding people's attitudes with a replicable statistical method. Section 2 will offer details on the method, highlighting the key advantages of the Generalized Q analysis compared to the traditional Q technique; Section 3 presents few case studies of the Generalized Q analysis application in various research domains, for a better understanding of how and with which kind of added value we may capitalise this method in social sciences research.

2. From traditional Q analysis to Generalized Q analysis

2.1. An illustrative case

The Generalized Q analysis tries to overcome the main limitations of the traditional Q method by working with graded re-combined statements based on combinations of basic statements ranked by small groups, assuming that respondents are consistent in their sequence of choices. Its basic principle will now concisely be described.

Suppose there are (q) questions with (r) alternative qualitative responses each. We may then have (q*r) basic statements where the (r) responses can be ranked for each one of the (q) questions. In the traditional Q analysis, the respondents have to rank the (q*r) basic statements, whereas in a Generalized Q analysis the respondents have to make (q) rankings of (r) responses, with the view to obtaining (q^r) combined and ranked responses. This will be illustrated by means of the following pedagogical example.

Suppose we are faced with an urban planning problem, where in the context of an inner-city rehabilitation plan several distinct choice options for urban rejuvenation are foreseen. These development possibilities for the inner city are each characterized by different features, such as accessibility, ecological quality etc. A group of respondents (e.g. stakeholders) has then to evaluate these various planning options by giving scores.

Let us assume a fictitious rejuvenation plan for a city, with 3 questions (planning issues):

Q1: number of inner-city parking places;
Q2: pedestrianisation of the inner city;
Q3: implementation of a 15-minute city concept.

We also assume that the respondents' support for each plan related to the questions Q1-Q3 is approximated by 3 preferential responses as follows:

For Q1:
R11: 3 – Parking facilities for all incoming cars;
R12: 2 – Parking facilities for residents;
R13: 1 – No parking facilities.

For Q2:
R21: 3 – No cars;
R22: 2 – Only cars for residents;
R23: 1 – Free access for all cars.

For Q3:
R31: 3 – Service supply to residents within 15 minutes;

R32: 2 – Service supply to all within 15 minutes;
R33: 1 – Reasonable service supply to all within 15 minutes.

The 4 features related to each of the respondents are:
F1: Age.
F2: Distance to shopping facilities.
F3: Car ownership.
F4: Distance to work.

The fictitious information provided by each respondent and the overall data are next illustratively included in Table 1 and 2. Table 1 provides an assessment of the comparison between three distinct sets of responses (R11, R12, R13; R21; R22, R23; R31, R32, R33) for a trio of questions (Q1; Q2; Q3) posed to ten individual respondents (I1, I2, I3, I4, I5, I6, I7, I8, I9, and I10). The four attributes (F1 to F4) associated with these respondents are detailed in Table 2, utilizing dummy variables. This data matrix will be used now to illustrate the Generalized Q method.

Table 1. Ranking of respondents I1-I10 on simple statements R11… R33

		I1	**I2**	**I3**	**I4**	**I5**	**I6**	**I7**	**I8**	**I9**	**I10**
Q1	R11	1	2	1	2	1	1	2	2	1	2
Q1	R12	3	3	2	1	1	2	1	2	1	1
Q1	R13	3	2	1	1	2	1	3	1	1	3
Q2	R21	2	2	2	1	2	2	2	2	2	1
Q2	R22	2	2	1	2	1	3	3	2	3	2
Q2	R23	3	3	3	3	3	2	3	1	2	3
Q3	R31	2	2	2	1	3	1	3	3	1	2
Q3	R32	3	1	1	2	1	3	2	3	1	3
Q3	R33	2	1	2	1	3	3	1	2	2	3

Source: authors' representation

Table 2. Features F1… F4 of the respondents I1-I10

c	F1 (Age)	F2 (Shop)	F3 (Car)	F4 (Work)
I1	1	0	1	1
I2	1	0	0	0
I3	1	1	0	0
I4	0	0	1	0
I5	1	1	0	0
I6	0	0	1	1
I7	0	0	1	0
I8	0	0	0	0
I9	0	0	0	1
I10	0	1	1	0

Source: authors' representation

2.2. Traditional Q analysis

In line with the classic Q method for multivariate data exploration, the application of Principal Component Analysis to Table 3 yields 5 significant components (C1 to C5), accounting for 25%, 20%, 15%, 14%, and 12% of the explained variance, totaling to 86% (as shown in Table 3). After performing a Varimax Rotation of the axes, the total explained variance of the 5 components appears to remain rather consistent, albeit with more evenly distributed weights across the various components. Table 3 is based on a connection between the 10 individual respondents and the 5 Principal Components.

Table 3. Traditional Q Method - ranking of respondents I1-I10 on simple statements R11… R33

	C1	C2	C3	C4	C5
NR	25	20	15	14	12
R	19	18	18	18	14
I1	0,405	0,102	-0,002	0,202	0,812
I2	0,636	-0,572	0,213	-0,405	0,165
I3	0,687	-0,14	0,551	0,348	-0,07
I4	0,503	0,447	-0,196	-0,446	0,048
I5	0,557	0,032	0,088	0,716	-0,406
I6	-0,18	0,815	0,493	-0,057	0,22
I7	0,497	0,066	-0,603	-0,249	-0,31
I8	-0,708	0,061	0,046	0,197	-0,11
I9	0,259	0,585	0,448	-0,381	-0,395
I10	0,29	0,624	-0,55	0,394	0,116

Source: authors' representation

Table 3 also illustrates that the majority of respondents align with Component C1, with the exceptions being respondents I6 and I8, who have in common a low age, living close to shopping facilities and using a car. Component C2 is opposed by respondents I2 and I3, that have in common a higher age, living close to work and not in the possession of a car, while Component C3 is opposed by I7 and I10, with lower age, car and close to work. Component C4 appears to align solely with respondent I5, while Component C5 corresponds to respondent I1 with high age, car and long distance from work.

It is also noteworthy that, as shown in Table 4, Component 1 exhibits a positive correlation with statements R23 (free access for all cars), R13 (no parking facilities), and R22 (only cars for residents), while showing a negative correlation with statements R12 (parking facilities for residents) and R33 (no services within 15 minutes). A similar interpretation can be obtained for all other Components presented in both Table 3 and Table 4. For instance, Component C2 corresponds to Response/Question R12 (parking for residents), Component C3 to R23 (free access for cars) and R33 (no service supply within 15 minutes), Component C4 to R22 (only cars for residents), and Component C5 to R32 (service availability within15 minutes).

Table 4. Traditional Q Method – Principal Component Factors on Simple Statements R11… R33

	C1	C2	C3	C4	C5
NR	25	20	15	14	12
R	19	18	18	18	14
R11	0,42	0,14	-0,99	-0,74	-1,26
R12	-1,52	1,59	-0,55	-0,24	1,03
R13	0,9	-0,32	-0,18	-1,42	0,78
R21	-0,76	0,71	0,33	0,35	-0,86
R22	0,85	0,16	-1,02	1,6	-0,81
R23	1,52	0,87	1,29	0,6	1,01
R31	-0,13	-0,36	0,99	-1,25	-0,97
R32	-0,25	-1,51	-1,14	0,22	1,17
R33	-1,01	-1,28	1,26	0,87	-0,07

Source: authors' representation

The above set of tables constitutes the analysis delivered through a traditional Q approach. With 10 respondents responding to 9 questions/responses, we arrive at 7 distinct, non-redundant questions/responses and 5 significant components. Clearly, the labelling of these components is still a challenge.

Table 5. Regressions of Extracted Values of Traditional Q Principal Components per Respondent on Features (F1, F2, F3, F4) of respondents

	Component 1				*Component 2*				*Component 3*				*Component 4*				*Component 5*			
	R	0,44	*Sig*	0,49	*R*	0,94	*Sig*	0,00	*R*	0,69	*Sig*	0,14	*R*	0,66	*Sig*	0,18	*R*	0,86	*Sig*	0,02
	Coef.	*Std E*	*Stat t*	*p*	*Coef.*	*Std E*	*Stat t*	*p*	*Coef.*	*Std E*	*Stat t*	*p*	*Coef.*	*Std E*	*Stat t*	*p*	*Coef.*	*Std E*	*Stat t*	*P*
C	-0,04	0,29	-0,14	0,89	0,08	0,10	0,81	0,46	-0,26	0,21	-1,23	0,27	-0,26	0,21	-1,23	0,27	-0,40	0,19	-2,04	0,10
F1	0,53	0,32	1,64	0,16	-0,62	0,11	-5,84	0,00	0,10	0,23	0,43	0,68	0,10	0,23	0,43	0,68	0,46	0,21	2,15	0,08
F2	0,11	0,35	0,30	0,78	0,46	0,12	3,94	0,01	0,67	0,26	2,63	0,05	0,67	0,26	2,63	0,05	-0,18	0,23	-0,79	0,47
F3	0,28	0,31	0,92	0,40	0,16	0,10	1,55	0,18	0,02	0,22	0,10	0,93	0,02	0,22	0,10	0,93	0,46	0,20	2,26	0,07
F4	-0,16	0,34	-0,48	0,65	0,53	0,11	4,76	0,01	0,14	0,24	0,55	0,60	0,14	0,24	0,55	0,60	0,15	0,22	0,67	0,53

Source: authors' representation

We will finally apply a regression analysis on our findings. The regression results involving the Extracted Values of traditional Q Principal Components for each Respondent against the Respondents' Features (F1, F2, F3, F4) are detailed in Table 5. The regression results suggest that, while only Components 2 and 5 exhibit significant associations with the respondents' features, it is noteworthy that Component 2 aligns notably with Factors 1 (Age), 2 (Distance to shopping), and 4 (Distance to work) for the respondents, while Component 5 correlates with Factors 1 (Age) and 4 (Distance to work). This simple example demonstrates the analytical power of traditional Q analysis.

2.3. Generalized Q analysis

To illustrate the potential and the limitations of conventional Q analysis, we will resort again to the previous simple example. Table 6 displays the conversion of the rankings provided by 10 respondents for the 9 questions/responses in Table 1, incorporating the features (Fi) from Table 3. This transformation results in 27 combined ranked questions/responses for the 10 respondents. There exist 9 dummy variables, signifying all the (3^3 = 27) potential combinations involving the three questions and three corresponding responses.

Table 6. Generalized Q Method – Combined Rankings of Respondents I1-I10 on Simple Statements R11… R33

			I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
V1	Q1	R11	1	2	1	2	1	1	2	2	1	2
V2	Q1	R12	3	3	2	1	1	2	1	2	1	1
V3	Q1	R13	3	2	1	1	2	1	3	1	1	3
V4	Q2	R21	2	2	2	1	2	2	2	2	2	1
V5	Q2	R22	2	2	1	2	1	3	3	2	3	2
V6	Q2	R23	3	3	3	3	3	2	3	1	2	3
V7	Q3	R31	2	2	2	1	3	1	3	3	1	2
V8	Q3	R32	3	1	1	2	1	3	2	3	1	3
V9	Q3	R33	2	1	2	1	3	3	1	2	2	3

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
F1	1	1	1	0	1	0	0	0	0	0
F2	0	0	1	0	1	0	0	0	0	1
F3	1	0	0	1	0	1	1	0	0	1
F4	1	0	0	0	0	1	0	0	1	0
Average	7	6	5	5	6	6	7	6	5	7
StdDev.	1	1	1	1	1	1	1	1	1	1

	D1	D2	D3	D4	D5	D6	D7	D8	D9	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
S1	1	0	0	1	0	0	1	0	0	5	6	5	4	6	4	7	7	4	5
S2	1	0	0	1	0	0	0	1	0	6	5	4	5	4	6	6	7	4	6
S3	1	0	0	1	0	0	0	0	1	5	5	5	4	6	6	5	6	5	6
S4	1	0	0	0	1	0	1	0	0	5	6	4	5	5	5	8	7	5	6
S5	1	0	0	0	1	0	0	1	0	6	5	3	6	3	7	7	7	5	7
S6	1	0	0	0	1	0	0	0	1	5	5	4	5	5	7	6	6	6	7
S7	1	0	0	0	0	1	1	0	0	6	7	6	6	7	4	8	6	4	7
S8	1	0	0	0	0	1	0	1	0	7	6	5	7	5	6	7	6	4	8
S9	1	0	0	0	0	1	0	0	1	6	6	6	6	7	6	6	5	5	8
S10	0	1	0	1	0	0	1	0	0	7	7	6	3	6	5	6	7	4	4
S11	0	1	0	1	0	0	0	1	0	8	6	5	4	4	7	5	7	4	5
S12	0	1	0	1	0	0	0	0	1	7	6	6	3	6	7	4	6	5	5
S13	0	1	0	0	1	0	1	0	0	7	7	5	4	5	6	7	7	5	5
S14	0	1	0	0	1	0	0	1	0	8	6	4	5	3	8	6	7	5	6
S15	0	1	0	0	1	0	0	0	1	7	6	5	4	5	8	5	6	6	6
S16	0	1	0	0	0	1	1	0	0	8	8	7	5	7	5	7	6	4	6
S17	0	1	0	0	0	1	0	1	0	9	7	6	6	5	7	6	6	4	7
S18	0	1	0	0	0	1	0	0	1	8	7	7	5	7	7	5	5	5	7
S19	0	0	1	1	0	0	1	0	0	7	6	5	3	7	4	8	6	4	6
S20	0	0	1	1	0	0	0	1	0	8	5	4	4	5	6	7	6	4	7
S21	0	0	1	1	0	0	0	0	1	7	5	5	3	7	6	6	5	5	7
S22	0	0	1	0	1	0	1	0	0	7	6	4	4	6	5	9	6	5	7
S23	0	0	1	0	1	0	0	1	0	8	5	3	5	4	7	8	6	5	8
S24	0	0	1	0	1	0	0	0	1	7	5	4	4	6	7	7	5	6	8
S25	0	0	1	0	0	1	1	0	0	8	7	6	5	8	4	9	5	4	8
S26	0	0	1	0	0	1	0	1	0	9	6	5	6	6	6	8	5	4	9
S27	0	0	1	0	0	1	0	0	1	8	6	6	5	8	6	7	4	5	9

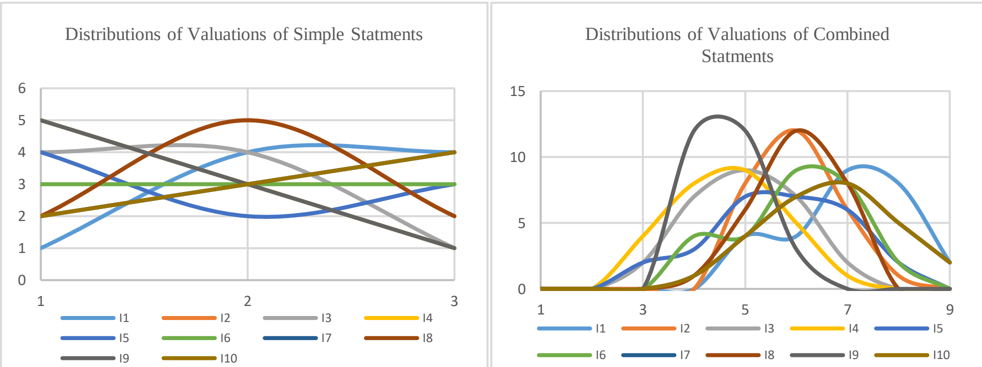
Source: authors' representation

The entry {[S1;I1]=5} in Table 5 is the result of multiplying the dummy values from vector S1, represented as (1,0,0,1,0,0,1,0,0), by the respondent evaluation vector V1, which is (1,3,3,2,2,3,2,3,2). This multiplication process is repeated for all Si dummy vectors and all Ij respondent evaluation vectors, leading to (i^k) evaluations of combined 'question/responses' (S1...27; I1...I10) that are organized here by column. These evaluations, according to the Central Limit Theorem, exhibit a Normal Distribution, rendering them amenable to processing using Principal Component Analysis techniques. We note here that:

$S_{ij} = \sum_{i=1}^{9} D_i \sum_{k=1}^{9} V_{kj}$ for all combinations (i^k) e respondent (j).

Figure 1 presents the frequency distributions for the individual and combined statements among the 10 respondents. This comparison highlights that the process of combining statements and their corresponding evaluations results in a normal distribution of the valuations.

Figure 1. Distribution of Valuations for Simple and Combined Statements



Source: authors' representation

Subsequently, we can employ Principal Components Analysis to assess the evaluations of amalgamated 'questions/responses' as presented in Table 5. This analysis yields the Extracted Values of Principal Components for each respondent, as presented in Table 6, and the Principal Component scores, as outlined in Table 8. Tables 7 and 9, on the other hand, encompass regression analyses aimed at elucidating the content of Tables 6 and 8, respectively.

The regression analyses, showcasing the Extracted Values of Principal Components for each respondent against the features of the respondents (F1, F2, F3, F4), are presented in Table 7. The findings reveal that Component 1 correlates with F1 (Age), Component 2 with F2 (Distance to shopping), Component 3 with F4 (Distance to work), and Component 4 with F3 (Car ownership). It is worth noting that these features can encompass variables represented by dummies, such as place of origin, gender, and income groups, as well as numerical variables like age, distance to a specific location, and income. The objective here is to discern which factors do or do not account for the amalgamated synthetic responses identified through Principal Component Analysis.

Table 7. Generalized Q Method – Principal Component Factors on Combined Questions/Responses R11… R33

	C1	C2	C3	C4		F1	F2	F3	F4
NR	31,7	22,9	17,8	14,4					
R	24,8	22,7	20,4	18,8					
I1	0,5	0,13	0,27	0,57		1	0	1	1
I2	0,89	0	-0,16	0,03		1	0	0	0
I3	0,87	0,4	0,05	-0,16		1	1	0	0
I4	0,01	-0,03	-0,11	0,8		0	0	1	0
I5	0,41	0,81	-0,34	-0,24		1	1	0	0
I6	-0,27	-0,11	0,92	0,25		0	0	1	1
I7	-0,14	0,1	-0,82	0,38		0	0	1	0
I8	-0,06	-0,96	-0,07	-0,25		0	0	0	0
I9	-0,6	0,32	0,51	-0,2		0	0	0	1
I10	-0,26	0,63	-0,11	0,72		0	1	1	0

Source: authors' representation

Table 8. Regressions of Extracted Values of Principal Components per Respondent on Features (F1, F2, F3, F4) of Respondents.

	Component 1				*Component 2*				*Component 3*				*Component 4*			
	R	0,93	*Sig*	0,004	*R*	0,62	*Sig*	0,227	*R*	0,71	*Sig*	0,130	*R*	0,86	*Sig*	0,021
	Coef.	*Std Error*	*Stat t*	*P Value*	*Coef.*	*Std Error*	*Stat t*	*P Value*	*Coef.*	*Std Error*	*Stat t*	*P Value*	*Coef.*	*Std Error*	*Stat t*	*P Value*
Constant	-0,108	0,119	-0,909	0,405	-0,371	0,271	-1,369	0,229	-0,162	0,235	-0,692	0,520	-0,166	0,137	-1,211	0,280
F1	0,964	0,132	7,334	0,001	0,189	0,299	0,634	0,554	-0,163	0,259	-0,628	0,558	0,087	0,151	0,571	0,593
F2	-0,224	0,143	-1,558	0,180	0,789	0,326	2,421	0,060	0,196	0,283	0,694	0,519	-0,042	0,165	-0,253	0,810
F3	0,084	0,124	0,673	0,531	0,205	0,282	0,727	0,500	-0,176	0,245	-0,719	0,504	0,767	0,143	5,363	0,003
F4	-0,389	0,137	-2,841	0,036	0,286	0,311	0,920	0,400	0,903	0,270	3,346	0,020	-0,166	0,158	-1,052	0,341

Source: authors' representation

Table 9. Principal Component Scores per Combined Statement and Dummies of Combined Statements

	C1	C2	C3	C4		D1	D2	D3	D4	D5	D6	D7	D8	D9
1	0,004	0,799	1,493	1,451		1	0	0	1	0	0	1	0	0
2	0,627	1,364	0,214	0,001		1	0	0	1	0	0	0	1	0
3	0,778	-0,33	-0,31	1,448		1	0	0	1	0	0	0	0	1
4	0,932	0,781	1,228	0,497		1	0	0	0	1	0	1	0	0
5	1,555	1,346	-0,051	-0,952		1	0	0	0	1	0	0	1	0
6	1,706	-0,348	-0,576	0,494		1	0	0	0	1	0	0	0	1
7	-0,719	-0,158	1,661	-0,175		1	0	0	0	0	1	1	0	0
8	-0,096	0,407	0,382	-1,624		1	0	0	0	0	1	0	1	0
9	0,055	-1,287	-0,143	-0,178		1	0	0	0	0	1	0	0	1
10	-1,415	0,953	0,215	1,555		0	1	0	1	0	0	1	0	0
11	-0,792	1,518	-1,064	0,106		0	1	0	1	0	0	0	1	0
12	-0,641	-0,176	-1,589	1,552		0	1	0	1	0	0	0	0	1
13	-0,486	0,935	-0,051	0,601		0	1	0	0	1	0	1	0	0
14	0,136	1,5	-1,33	-0,848		0	1	0	0	1	0	0	1	0
15	0,287	-0,195	-1,855	0,598		0	1	0	0	1	0	0	0	1
16	-2,137	-0,004	0,383	-0,071		0	1	0	0	0	1	1	0	0
17	-1,515	0,561	-0,896	-1,52		0	1	0	0	0	1	0	1	0
18	-1,364	-1,133	-1,421	-0,074		0	1	0	0	0	1	0	0	1
19	-0,191	-0,213	1,472	1,026		0	0	1	1	0	0	1	0	0
20	0,432	0,352	0,193	-0,423		0	0	1	1	0	0	0	1	0
21	0,582	-1,342	-0,332	1,023		0	0	1	1	0	0	0	0	1
22	0,737	-0,231	1,206	0,072		0	0	1	0	1	0	1	0	0
23	1,36	0,334	-0,073	-1,377		0	0	1	0	1	0	0	1	0
24	1,511	-1,36	-0,598	0,069		0	0	1	0	1	0	0	0	1
25	-0,914	-1,17	1,64	-0,6		0	0	1	0	0	1	1	0	0
26	-0,291	-0,605	0,361	-2,049		0	0	1	0	0	1	0	1	0
27	-0,14	-2,299	-0,164	-0,603		0	0	1	0	0	1	0	0	1

Source: authors' representation

Next, the results of regression analyses for each of the Principal Component score vectors per combined statement against the corresponding dummy variables for combined statements (D1, D2, D3, D4, D5, D6, D7, D8, D9) are presented in Table 9.

Our findings indicate that Component 1 aligns with Response (2) to Question (2) (R22- Only cars for residents), Component 2 corresponds to Response (2) of Question 3 (R32 – Services within 15 minutes), Component 3 associates with Response (1) to Question (3) (R31-All services to Residents within 15 minutes)), while Component 4 corresponds to Response (1) of Question (2) (R21 -No cars).

Moreover, the results in Table 9 reveal that Component 1 favors Response 2 to Question 2 while opposing Response (2) to Question (1) (R12-Parking for residents). Component 2 exhibits disagreement with the majority of responses to the questions. Component 3 disagrees with responses to Question (1) and aligns with response R31- All services to residents within 15 minutes. Component 4 is in opposition to responses to Questions (1) and (3) and favors response R21-No cars.

Table 10. Coefficients of Regressions of 4 Principal Component scores per Combined Statement and Dummies of Combined Statements.

		C1	C2	C3	C4
R11	D1	0,055	-1,287	-0,143	-0,178
R12	D2	-1,364	-1,133	-1,421	-0,074
R13	D3	-0,140	-2,299	-0,164	-0,603
R21	D4	0,778	-0,330	-0,310	1,448
R22	D5	1,706	-0,348	-0,576	0,494
R23	D6	0,055	-1,287	-0,143	-0,178
R31	D7	-0,719	-0,158	1,661	-0,175
R32	D8	-0,096	0,407	0,382	-1,624
R33	D9	0,055	-1,287	-0,143	-0,178

Source: authors' representation

3. Overview of some empirical applications

As mentioned, Q analysis has found many applications in qualitative social science research on respondents' perceptions and preferences. More recently, also the Generalized Q method has an application in some empirical research studies. This section presents four examples to better understand how this method has been applied and which results it can generate.

In a project on "Comprehensive climate hazard mapping and risk assessment and development of risk model for Timor-Leste. UNDP/TLS/PS/2021/016F" (D'Haeyer et al., 2022), small focus groups of stakeholders in each one of the six priority municipalities (Aileu, Baucau, Ermera, Liquica, Lautem, Viqueque) were asked to rank 6 x 8 x 6 ranked statements through a brief questionnaire with 6 + 8 + 6 + 6 = 22 statements, summarized into two main components that can be related upwards to prevailing socio-economic conditions and downwards to adaptive capacity and vulnerability, a more autonomous urban reaction and a rural point of view more dependent on the reaction of the authorities. The results show that beyond the socio-economic conditions of population and of each place, assumed ex ante and based on secondary data, it is the people's attitude that influences the adaptive capacity and the vulnerability regarding climate-related hazards.

In their study on "Regional science knowledge needs for the recovery of the Ukrainian spatial economy: A Q-analysis" (Pascariu et al., 2023), each questionnaire applied to 35 respondents uses a Likert scale (1–5) on 25 relevant regional science topics and on 19 knowledge interest activities, leading to (25+19=44) simple statements and (25*19=475) combined pairwise choices that are more adequate to perform the Q analysis, leading to the identification of 6 components or types of attitudes: Component 1, common to all respondents, focuses on knowledge exchange, Component 2 highlights the need for twinning publication projects, Component 3 reflects topics of research. The empirical results show that a regional science methodology can play an important role in the geopolitical conflict in Ukraine, during and after the war.

In the same Special Issue of Regional Science Policy and Practice on "Ukraine: Geopolitical Realities and Regional Development Perspectives", Simionov (2023) uses an Generalized Q analysis in order to handle 50 online questionnaires in the Republic of Moldova (including people living in close proximity with Romania/at the borders or that have close links to Romania) which reach 2048 = 2^11 composed statements, where 6 sentences outlined on identity (European, East European, Romanian, Moldavian, Transnistrian and other) and 5 sentences on Moldova's geopolitical prospects ('Moldova stays as it is', 'Moldova becomes part of Romania', 'Moldova becomes part of Russia's EEU', 'Moldova integrates into the EU', 'Moldova expands to the Black Sea'). The questionnaire also included a socio-demographic data collection of the respondents' specific features (age, gender, education, language, occupation, and city). Component 1 defends the European integration of the Republic of Moldova and represents 82% of the variance. Interestingly, some native Russian-speaking people also adopt this perspective. Nevertheless, most of the respondents that support Moldova's status quo in Component 2 live in Chisinau and many of them are native Russian speakers. Component 3, which promotes a Greater Moldova (until the Black Sea) gathers Moldovans with higher levels of education, many of whom work in academia. Finally, Component 4, which highlights a desire to unite Moldova and Romania, is mostly represented by older men and whose native language is Moldovan/Romanian.

Finally, during a workshop in Portugal on 'University and Sustainable Regional Development' (Dentinho, 2022), 85 respondents were asked to answer 'yes' or 'no' to the following questions: 1) Do people want the freedom to live or the security of common sense and infrastructures? 2) Do people want to love creating with risk and delivery or to be loved without risk and delivery? 3) Do people want a university that spreads common sense or creates and spreads the vocation of people and places? 4) Does the University want to understand the relationship between Faith and Reasoning, or does it prefer to see them separately? 5) Does the Church want to understand the relationship between Faith and Reasoning, or does it prefer to see them separately? 6) Does the State want to understand the University as the diffuser of common sense or as the creator of new common senses? 7) What is the preference in sustainable development: socio-economic; socio-environmental; economic-environment? The average responses defined that 'We want to be safely free, to love and be loved, to discover new meanings more than spreading consensus, a University with more Reasoning than one with Faith and Reason, as well as research and teaching, a Church with Faith and Reason more than just Faith, and the State to foster new meanings rather than to promote the dissemination of consensus'. Nevertheless, when 2x2x2x2x2x2x3 = 192 combined responses were considered, results become richer and more specific, with a clear distinction between regions: Picking up just the first two components, the Northeast of Portugal opposes against the South and supports the social aspects of development, against the absolute position of reasoning and pro the state of common senses (C1); the youngest are pro the economic dimension of development and against the absolute of reasoning and the new senses (C2).

Summing up, the Generalized Q analysis not only allows the usual search of respondents' average attitudes, but also provides suitable re-combined responses to find common and different attitudes and relate them with the features of people and places.

Conclusions

The Generalized Q analysis effectively addresses the primary limitations associated with the traditional Q Method, which typically requires the ranking of an extensive number of statements, imposes constraints on the quantity of non-redundant respondents, and permits subjective interpretations of the components.

Being based on a larger number of responses, the Generalized Q Analysis not only allows the usual search of average profiles but also facilitates the identification of common and different attitudes and relates them with the features of people and places.

The Generalized Q analysis operates with ranked combined statements, derived from combinations of basic statements ranked by smaller groups. It operates under the assumption that respondents exhibit consistency in their sequencing of choices. Consequently, the data collection time per respondent is significantly reduced. Moreover, by enabling the involvement of a larger number of respondents, the regressions supporting interpretation can attain greater robustness.

The new Generalized Q methodology has a great potential for further applications in the social sciences. For example, we may use a Generalized Q method to analyse the consistency of public spending by relating the combined distribution of public spending per sector (Observations) to regional units (Variables) and by then estimating the Principal Components that can be related to the features of the regions, permitting therefore to assess the consistencies of the allocation of public funds. Generalized Q analysis may be particularly useful in case of citizen-oriented planning issues. A great potential may exist for applications in the fields of public participation and citizen science, in particular since the value statements of respondents or stakeholders can easily be collected and understand. Clearly, many more applications in urban, regional and public policy may be anticipated.

References

Brown, S. R. (1993). A primer on Q methodology. *Operant Subjectivity*, *16*(3/4), 91-138.

Churruca, K., Ludlow, K., Wu, W., Gibbons, K., Nguyen, H. M., Ellis, L. A., & Braithwaite, J. (2021). A scoping review of Q-methodology in healthcare research. *BMC Medical Research Methodology*, *21*(1), 125. https://doi.org/10.1186/s12874-021-01309-7

D'Haeyer, T., van de Moortel, I., de Ruijter, A., Thant, S., Issler Biesmer, A., Krishna Krishnamurthy, K., Dentinho, T.P., & Sulistiawaty, E. (2022). *Comprehensive climate hazard mapping and risk assessment and development of risk model for Timor-Leste.* UNDP/TLS/PS/2021/016.

Dentinho, T. P. (2022, February 16). *Universidade e Desenvolvimento Regional Sustentável, Tatuados pela Geografia* [University and Sustainable Regional Development, Tattooed by Geography]. Workshop on University and Regional Development. University of Évora.

Ellis, C. (2004). *The Ethnographic I: A Methodological Novel about Autoethnography* (Vol. 13). Rowman Altamira.

Gawron, V. J. (2016). Summary of fatigue research for civilian and military pilots. *IIE Transactions on Occupational Ergonomics and Human Factors*, *4*(1), 1-18. http://dx.doi.org/10.1080/21577323.2015.1046093

Guastello, S. J., Correro, A. N. II, Marra, D. E., & Peressini, A. F. (2019). Physiological synchronization and subjective workload in a competitive emergency response task. *Nonlinear Dynamics, Psychology, and Life Sciences*, *23*(3), 347-376.

Hair, J., Anderson, R., Tatham, R., & Black, W. (1998). *Multivariate Data Analysis*, Prentice-Hall, Englewood Cliffs.

McKeown, B., & Thomas, D. B. (2013). *Q Methodology*, Sage, London.

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, *59*, 210-220. https://doi.org/10.1016/j.pragma.2013.07.012

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81. https://psycnet.apa.org/doi/10.1037/h0043158

Nijkamp, P. (Eds.) (1985), *Measuring the Unmeasurable*, Martinus Nijhoff, Dordrecht.

Pascariu, G. C., Nijkamp, P., & Kourtit, K. (2023). Regional science knowledge needs for the recovery of the Ukrainian spatial economy: A Q-analysis. *Regional Science Policy and Practice*, *15*(1), 75-94. https://doi.org/10.1111/rsp3.12638

Pepeka, M. E., de Fátima Ferreiro, M., & Dentinho, T. P. (2022). Acesso à terra e segurança alimentar. As perspetivas das partes rurais interessadas de Huambo, África. *Revista portuguesa de estudos regionais*, 61, 69-84. https://doi.org/10.59072/rper.vi61.532

Sánchez, A. F., Dentinho, T. P., Arroz, A. M., & Gabriel, R. (2021). Sustentabilidade Urbana: Aplicação da Metodologia Q em Cinco Cidades do Arquipélago dos Açores [Urban Sustainability: Applying the Q Methodology to Five Cities in the Azores Archipelago]. *Revista portuguesa de estudos regionais, 57*, 33-56.

Simionov, L. M. (2023). Shifting attitudes towards identity, borders and geopolitical choices: The case of Moldova. *Regional Science Policy & Practice*, *15*(1), 200–221. https://doi.org/10.1111/rsp3.12613

Stephenson, W. (1953). *The Study of Behaviour; Q-technique and its Methodology*, University of Chicago Press, Chicago.

Thomas, D. B., & Baas, L. R. (1992). The issue of generalization in Q methodology: “reliable schematics” revisited, *Operant Subjectivity*, *16*(1/2), 18-36 http://dx.doi.org/10.22488/okstate.92.100599