Connecting Q & surveys: three methods to explore factor membership in large samples

Baker, Rachel M.; van Exel, Job; Mason, Helen; Stricklin, Michael

Connecting Q & surveys: three methods to explore factor membership in large respondent samples

Rachel Baker[1], Job van Exel[2], Helen Mason[3], , Michael Stricklin[4]

[1] Yunus Centre for Social Business and Health, Glasgow Caledonian University, Scotland, UK.

[2] Institute of Health Policy & Management, Erasmus University, Rotterdam, Netherlands.

[3] Institute of Health and Society, Newcastle University, UK.

[4] University of Nebraska Lincoln (Emeritus), Federal University of Piaui, Brazil.

Corresponding author
Rachel Baker PhD
Reader in Health Economics
Yunus Centre for Social Business and Health
Glasgow Caledonian University
Level 3 - Buchanan House
Glasgow Caledonian University
Cowcaddens Road
Glasgow G4 OBA
Scotland, UK.

Email: rachel.baker@gcu.ac.uk
Telephone: +44 (0)141 331 8246

"We are to propose that along Q-technique lines it is often possible to discover complex facts, of the kind usually regarded as inferences, by previous study of relatively few cases only. They can thereupon be counted, if need be, by using an appropriate questionnaire and large-sampling techniques" Stephenson (1953, p190)

1 Introduction

Q methodology as invented by Stephenson (Stephenson, 1953) was "designed to assist in the orderly examination of human subjectivity" (Brown, 1980, p. 5). Q incorporates quantitative factor analytic methods into a broadly qualitative, interpretive framework (Brown, 1996) and is perhaps one of few truly *mixed methodologies* (Stenner and Stainton Rogers, 2004). Q techniques are usually applied to research questions of a qualitative kind, typically: "What is the nature of attitudes and beliefs held by community (or individual) X on the subject of Y". Each participant expresses her point of view by sorting a set of items. Factor analyses based on the correlations between all participants' Q sorts are used to identify underlying *value sets* and shared beliefs. The capacity to 'tap into' underlying preference systems that may not otherwise have been articulated by respondents is a particular strength of Q methodology. Participants position each item relative to all other items in the set such that their Q sort as a whole conveys meaning, without necessarily having a readily constructed and coherent point of view on a topic.

The results of Q studies are rich descriptions of a small number of shared views (factors) in relation to a particular topic. Techniques more familiar to qualitative researchers are used in the interpretation of factors, including interviews or other open ended, qualitative methods. These supplementary data are used as a means of assessing the relevance and

appropriateness of a range of feasible factor solutions, and in the interpretation and description of factors. Typically the description of factors is the point at which the Q analyst concludes. In presenting this new knowledge - i.e. eliciting and describing viewpoints that exist around a subject - Q methodologists inform theory, policy or day-to-day practice. However, there are research questions that require not only explanation of the points of view that exist around a topic, but also information about how common those points of view are in a population, and how they are distributed over groups with different individual and socio-demographic characteristics. This is a different type of question. Rather than asking "what is the nature of …?" it presupposes that the nature of viewpoints has been established and asks "what is the extent of points of view A, B and C in population D?" Quantitative methods are required to address questions of prevalence and distribution.

Social scientists have, for some time, employed survey research methods to explore the demographic distribution of phenomena. Examples of well established, national surveys are the British Household Panel Survey (BHPS)[1], the Household, Income and Labour Dynamics in Australia (HILDA)[2], and the US Panel Study of Income Dynamics (PSID)[3]. Large scale surveys (often repeated over time) are an effective means to explore *facts* about populations such as the prevalence of an illness, number of children, or collecting information about housing or income. Such survey questions require careful framing and standardised, validated approaches to wording and categorising such that results can be compared. The

---

[1] The aim of the BHPS is "to further our understanding of social and economic change at the individual and household level in Britain and the UK" (see http://www.iser.essex.ac.uk/survey/bhps).

[2] See http://melbourneinstitute.com/hilda/.

[3] See http://psidonline.isr.umich.edu/.

relationship between selected *facts* and other variables can then be assessed using statistical procedures.

Attitudinal information is also collected using survey methods. The abovementioned panels collect information about peoples' social and political values, opinions and attitudes, on topics like work, consumption, environment, transportation, morality and religion. Attitudinal survey questions are, arguably, much more susceptible to issues of design, question choice and framing than 'factual' questions. There are questions of how to frame meaningful questions; whether they will be 'meaningful' (and what they will mean) to every respondent in a sample; how to scale or score responses to attitudinal questions and how to aggregate the resulting data. Designing effective questions about income and tax is concerned with specificity and clarity, around issues such as gross or net income and whether any other deductions are included; whether respondents should report their personal or their household income and the setting of bands allowing respondents to indicate their income level. Posing questions about individuals' *views* about the distribution of income, for example, or their attitudes to taxes is a very different art form. How should questions be worded or response categories defined when the survey researcher is dealing with questions of subjective opinion? Well designed questionnaires will often incorporate qualitative work to inform questionnaire design and establish appropriate questions and response categories. Versions of draft questionnaires are piloted and often cognitive interviews are used to generate data around the different understandings and linguistic nuances associated with different ways of posing questions. Factor analysis of items or Rasch models might then be used to identify underlying dimensions and dispose of

redundant items. A final questionnaire is then born! (although many emerge without several of the steps listed above).

Assuming that an attitudinal questionnaire has been carefully designed and thoroughly piloted, there remain issues that cannot be addressed within existing methods. The relationship between the items is unspecified and unknown except in relation to correlations between one item and another. In making a case for Q techniques preceding questionnaires, Stephenson remarks of the standard approach to questionnaire research: "The elegancies of the *discriminative function* can help out: but the end-product is still merely a catalogue of interesting, but disparate and unrelated, facts"! (Stephenson 1953 page 194)

We believe that Q has something to offer to the exploration of subjective opinion in larger populations that is not available using standard survey methods. Q analyses generate rich information sets including information about the way in which items relate to each other within different points of view. But, as (Brown, 2002)phrased it, Q is "a clumsy way to count noses". If, however, questionnaires can be devised from Q analyses, and if questionnaire respondents' association with factors can be estimated using methods derived from Q, then the extent of Q-factors in populations might be described. Furthermore, using quantitative techniques, the association between those viewpoints and other personal characteristics (such as age, gender, health, or socio economic status) could be subject to analyses.

In this paper we discuss the development and application of Q methods in questionnaire design.[4] In section 2, 3 and 4 we describe in brief three different approaches to the derivation of questionnaire items from Q studies and corresponding scoring systems to indicate 'factor membership' of questionnaire respondents, each followed by a worked example from our own work in the field of health research. In the final section we discuss the advantages and disadvantages of the different approaches and conclude by highlighting some issues that require further development and research.

2 Talbott's Q block

2.1 *Description of the approach*

In a conference paper in 1963 (and reproduced in this special issue), Albert Talbott described what he termed"The Q block method of indexing Q typologies" (Talbott, 1963). Building on Stephenson's suggestions in *The Study of Behaviour*, Talbott aimed to assign people to 'Q typologies' by designing questionnaire items that require the rank-ordering of a small number of specially selected statements from a previous Q study. Statements are selected on the basis of their salience and distinction in the original Q factor analysis and presented in 'blocks' such that each factor is represented by one statement in each block. For a 3 factor solution, therefore, respondents would be required to rank order 3 statements per question. Talbott referred to these ranking questions as "Q blocks".

---

[4] The starting point for this paper was an organised session -and the following discussions- during the proceedings of the Conference of the International Society for the Scientific Study of Subjectivity in Bethesda (US), 2007: Combining Q&R; with Rachel Baker (UK, University of Newcastle), James Good (UK, University of Durham), Michael Stricklin (Brazil, Federal University of Piauí) & Job van Exel (NL, Erasmus University).

Adopting Talbott's approach, a Q set can be whittled down to a smaller number of candidate statements by identifying those statements that have high z scores for one factor and substantially lower z scores for the other factors. A good place to start would, therefore, be the list of distinguishing statements for each factor. Consensus statements are of little use in the construction of Q blocks and would be discarded. Selected statements should also be salient to that factor, in other words, factor scores should be in the tails of the array – distinguishing statements in the middle of the array would not usually be included.

Once a sub-set of salient, distinguishing statements is identified for each factor, then statements can be grouped into blocks. There is little guidance on how to group statements into blocks or any indication on how many blocks should be constructed, although presumably the latter is determined by the size of the original Q set and the number of statements satisfying the criteria. Talbott does raise the issue of positive versus negative statements and points out that Q blocks should be constructed so as to contain one or other. Respondents are then directed to rank-order the statements in each Q block, through which they assign a score to each of the underlying factors. The scoring system demonstrated by Talbott involves the straightforward addition of scores for each factor across Q blocks. By totalling the scores for the statements chosen to represent each of the factors from the original Q study, any given respondent will be assigned a total score for each of the factors, from which we might infer that respondent X was most closely associated with Factor A and had least in common with factor D, for example.

2.2 Q-ing for health: a recent application of Talbott's Q block method

In a study of the views of the public regarding the principles underlying health care decision making (Baker et al., 2010) we identified three factors. The factors were distinguished by: egalitarian principles and equality of access (factor 1); a concern with health care outcomes, some concern for those with dependants and total rejection of socio-economic issues in prioritising health care (factor 2); the prioritisation of children, life saving and a belief that such decisions should be made by experts (factor 3). Based on this analysis we set about designing a set of survey questions using Talbott's approach.

From the original Q set of 46 statements, 20 statements were identified as both salient and distinguishing for any one of the three factors. In this study salience was interpreted as those statements with a factor score of +/- 3 or greater. Whilst 9 statements were distinguishing and salient for each of factors 1 and 2, only 6 statements satisfied these criteria for factor 3. Certain statements were distinguishing and salient for more than one factor, in different directions. Statements were categorised according to the factor they represented and the sign of the factor scores (such that positive statements would be blocked with other positive statements and negative with negative). This resulted in four blocks of 3 statements (Figure 1)[5]. Notice that questions 1 and 2 comprise positive statements for each factor and questions 3 and 4 negative statements.

---

[5] These were part of a longer questionnaire, details of which can be found in the project report Baker et al 2010.

Figure 1: The Q blocks from the SVQ questionnaire

Q block 1

#	Statement text	Factor scores/ z scores		
		F1	F2	F3
15	Everybody, no matter what you are, whether you are young or old, should get the same access to and choice of treatment.	5*	-1*	3*
		1.916	-0.333	1.081
29	The quality of life of patients and their life expectancy are the most important things. The characteristics of patients like whether they are employed, or whether they have dependants, or what gender they are shouldn't matter.	3	5*	2
		1.068	1.759	0.619
13	Age shouldn't come into it, unless you're talking about children. Children's health should be given priority over adults.	0	0	4*
		0.252	-0.129	1.291

Q block 2

#	Statement text	Factor scores/ z scores		
		F1	F2	F3
25	People with dependants should not be given priority over people without dependants. A human life is a human life, I think it should be irrelevant how many dependants they've got.	4*	-2*	2
		1.189	-0.449	0.579
20	People with dependants should be prioritised over people without dependants because their treatments would benefit others as well as the patient themselves.	-4	2*	-4
		-1.451	0.746	-1.507
35	The decisions about which services to fund, and how to spend NHS money should be made by a range of experts with a lot of information and experience of the issues, not the general public.	-3	0*	3*
		-0.929	-0.192	1.104

Q block 3

#	Statement text	Factor scores/ z scores		
		F1	F2	F3
16	You should prioritise the younger age group, because they are still able to have children.	-5*	-1	-3
		-1.579	-0.425	-0.827
28	Whether or not patients can contribute financially towards the cost of the treatment should be taken into account because it would allow you to treat more people who can't afford to 'go private'.	-3*	-5*	-1*
		-1.060	-1.883	-0.406
44	It's no good saving lives if the quality of those lives is really bad. Some treatments are keeping people alive for too long. You've got to have a decent quality of life otherwise what's the point of being alive.	1*	4*	-3*
		0.394	1.267	-0.982

Q block 4

#	Statement text	Factor scores/ z scores

		F1	F2	F3
14	The age of the patient is important; if you were treating children rather than older people then you would have a longer improved life.	-4* -1.328	2* 0.509	-1* -0.374
26	Poorer people should be given priority because they don't have the same opportunities to take care of their own health.	-1 -0.584	-5* -1.789	-3 -0.760
9	People who smoke and drink pay enough in extra taxes to pay for their own health care.	-2 -0.850	-4 -1.200	-5* -1.922

From each block, respondents were first asked to select the statement they most agreed with. Then from the remaining two statements to select the one they most agreed with. Finally they were asked to confirm that the one remaining statement was the one that they least agreed with from the block of 3[6]. This was repeated for 4 Q block questions[7].

2.2.1 Results

The survey was administered by a survey organisation (NatCen) to a random sample of 587 respondents aged 18 and over and living in England during February-April 2007. Of these 542 respondents gave complete responses to the Q block questions. Table 1 shows the result of scoring their responses according to Talbott's method (i.e. scores 1, 2 and 3 for

---

[6] The exact instruction to the Q block question was: "The next four questions are a little different. You will be shown 3 statements each time. These statements are things that members of the public have said about how health services should be prioritised. You will agree with some of these and disagree with others. There are no right or wrong answers. For each question, first read through the 3 statements on the show-card carefully. When asked, give the number of the statement that you agree with most. Then from the two statements that are left you will be asked to give the number of the next statement you agree with the most. Finally the one remaining statement should be the one you agree with least."

[7] Note that for blocks 3 and 4 this might have seemed somewhat counterintuitive since respondents are asked to indicate agreement with statements that distinguish factors at the negative pole. However, on balance it was judged to be less confusing for respondents than the alternative which is switching the Q block questions midway from selection according to agreement to selecting according to disagreement.

most, next and least agreed with statement, respectively) and also by applying a different set of scores (i.e. scores 1, 5 and 12) to reduce the possibility of ties. Those respondents whose scores revealed a preferred factor (i.e. no tie in first place) were 'assigned' their preferred one of the three factors (first 3 rows of Table 1). For a number of respondents, their rankings indicated that they had views in common with two factors: 78 respondents tied across two factors using scoring method 1 and 38 mixed across factors using method 2. It can be seen that no single factor appears to dominate in this population. This was a little surprising since factors 1 and 2 were familiar accounts from the literature whereas factor 3 was a less well-established account and a little more difficult to interpret. This is in itself an interesting finding which adds to the initial Q analysis, but the analysis of Q block data raised a number of questions that warrant discussion.

Table 1: Talbott Q block scoring method

Factor	Scoring method 1 [scores 1, 2, 3] Count (%)	Scoring method 2 [scores 1, 5, 12] Count (%)
F1	102 (19)	123 (22.7)
F2	157 (29)	178 (32.8)
F3	173 (32)	193 (35.6)
F1F2	21 (4)	10 (1.8)
F1F3	16 (3)	18 (3.3)
F2F3	41 (8)	20(3.7)
F1F2F3	32 (6)	0

2.2.2   Methodological issues

This was a first attempt to devise a Q block survey and, through it, we identified several issues of methodological interest. These relate to i) the selection of statements and framing

11

of the questions (the means by which we generate Q block data) and ii) the scoring and

analysis of Q block data.

Selection and framing

Designing Q blocks involves setting explicit criteria for the selection of a subset of

statements from a larger Q set. This is not inherently problematic, if those criteria are clear.

However the isolation of an item from the other items in the Q set (the Q set having been

carefully selected as the holistic representation of a concourse) and the choice to place it

with two or three other selected statements *might* have implications for the interpretation

and rank ordering of that statement. In addition there is often more than one way of

grouping statements into Q blocks and there is currently no methodological standard for

this. Different compositions of the same subset of statements into Q blocks could

conceivably influence responses and this is an issue requiring methodological and empirical

investigation if Q block methods are to be advanced.

The second framing issue we identified relates to the most appropriate treatment of

statements that have been selected as representing a factor in the positive sphere or the

negative. We have noted, following Talbott, that these should not be mixed within a single

Q block and in our study, there were 2 positive Q blocks and two negative. However, as in

Talbott's study, all of our Q block questions were phrased in terms of agreement. This was a

conscious decision made in an effort to enhance consistency and respondent

comprehension, especially since the Q block questions were posed at the end of a relatively

long, relatively complex questionnaire. However, it is possible that respondents would be

uncomfortable with stating they "agree most" with a statement that, in reality, they

"disagree with least" since one is not necessarily the inverse of the other.

Scoring and analysis

Although he notes that a more complex approach could be adopted, we followed Talbott's

simple scoring method, only adapting the scoring system to reduce the possibility of tied

scores. Analysis of Q block data in this way has several implications. If we hope to retain

the features of Q methodology that characterise the method, the notion of allocating

individuals to "Q types" and using this type of scoring system cannot accommodate the

intuition embedded in Q that individuals are very often associated, partially, with more than

one factor. Even respondents with one very high factor loading often have positive

(significant) factor loadings on another factor. Q block analysis also fails, in this simple

form, to make use of the information we have about the differential contribution of

different statements. The z scores included in Figure 1 indicate that some statements

contribute more to a factor than others. Incorporation of this information is an important

area for future methodological work.

3 Brown's standardized factor index score

3.1 Description of the approach

A second approach to investigate peoples' likely factor membership and the distribution of

factors in a (representative) sample of the population was proposed by Brown (2002). In this

approach participants in a survey questionnaire are asked to evaluate a sub-sample of a Q-

set, consisting of statements representing the factors found in the Q-study. Participants

score each statement using a common Likert-type scale (e.g., a 7-point scale ranging from 'least important' to 'most important'). The scores a participant gives are used to compute a standardized index score for each factor, which indicates how likely the participant is associated with each factor, and which factor is most in common with the participant's perspective.

3.2   A recent application: Attitudes of informal caregivers toward respite care

As part of a study on the demand for and use of respite care[8] among Dutch informal carers (Van Exel et al., 2006b), we conducted a Q study on their attitudes toward respite care. This Q study revealed three distinct attitudes, which for a health policy audience were labelled "need and ask for respite care", "need but will not ask for respite care", "do not need respite care". Full descriptions of the factors can be found in van Exel et al. (2007). Next, we designed a study aimed at 'nose counting' and exploring associations between attitudes towards respite care and characteristics of the care giving situation.

Much like the Talbott Q-block approach, Brown gives no clear guidance on the *number* of statements for use in the survey and how to select the *appropriate* ones from the Q-set. We used two intuitive criteria for the selection of statements: (1) the statement should be distinguishing; and (2) the statement should be salient for at least one factor.[9] The rationale

---

[8] Respite care is a generic term for different types of interventions aimed at providing support and relief to informal caregivers by (temporarily) easing the burden of their care giving task. Respite care has the objective to increase or restore the caregiver's ability to bear this load (i.e., the caregiver's supporting capacity), and may take many forms, such as in-home respite care, day care, short-break or short-stay (in-patient, for a week or weekend), special holiday arrangements, discussion/support groups, and training by professionals.

[9] A statement is distinguishing for a factor if its rank score on that factor is statistically significantly different from its score on all other factors. A statement is considered salient for a factor if it is ranked in one the two

for the first criterion was that the statements in the survey should highlight the differences

between the attitudes toward respite care. The second criterion was added because the

statements should be recognizable and appeal to the sentiment of groups of carers. In other

words, we tried to select statements from each factor in the Q study that carers with the

corresponding attitude would be able to relate to, and that carers with other attitudes

would vote down sufficiently. From the original sample of 39 opinion statements, 13 fulfilled

these two criteria. We used 12 of these in the survey, each factor represented by four

distinguishing statements (see Table 2) [10]. Respondents were asked to signal their

agreement with these statements on a 5-point Likert-type scale, which, in accordance with

the score sheet used in the Q study, ranged from 'totally disagree' to 'totally agree' (Van

Exel et al., 2008).

3.2.1 Results

Data were collected via postal questionnaires in the spring of 2005. We received useful

response from 249 informal carers, mostly non-employed females of 50 years and older,

caring for their partner or parent (in-law) for, on average, the past 9 years and 40 hours a

week (Van Exel et al., 2008).

Frequency distributions showed that participants had used all response categories (from 1,

'totally disagree', to 5, 'totally agree') for all 12 statements, indicating that the selected

---

outer columns of the composite sort of that factor (in the distribution used in this study, those statements
with a rank score +4, +3, -3 or -4+).

[10] For one factor five distinguishing statements were eligible and one statement was removed. This selection
was based on (1) the observation that two of these five statements related to the same underlying concept
from the theoretical structure that was used to develop the original Q set and, from these two, (2) the highest
rank score on this factor.

statements appealed very differently to the sentiments of carers. Reliability analysis showed that two statements (number 8 from factor 1 and number 24 from factor 2) were scored quite differently from the other three statements in that factor, and were consequently excluded from further analysis.[11] The remainder of the scoring procedure was conducted using 10 of the 12 statements.

Table 2 clarifies the scoring procedure. The first five columns show the factors, the selected statements and the factor scores of the statements in the original Q study. The next column shows the mean item score for each statement, representing mean agreement with the statements across participants. For instance, the statement "care giving makes me feel good" (39) had a mean score on the Likert-type scales of 4.1, indicating that carers in this sample overall tend to agree with this statement.[12] Then, for each participant individually, a statement index score and a factor index score were calculated. The statement index score is calculated as the product of the absolute value of the factor score (which is fixed across respondents, as it originates from the underlying Q study) and the item score (which varies between respondents, based on how they scored the statement on the Likert-type scale). The factor index score was calculated for each factor as the sum of the statement index scores of the respondent pertaining to that factor. Mean statement and factor index scores for the sample are shown in Table 2.

---

[11] In fact these two statements were only excluded after the scoring procedure using 12 statements led to ambiguous results, in particular when looking at associations between factor membership and characteristics of the care giving situation.

[12] Table 2 shows the reverse score (i.e. 1.9) because the factor score of this statement on factor 1 is negative (i.e. -1).

Table 2: Calculation of Brown's factor index score

Factor	Statement	Factor score [a]			Mean item score [b]	Statement index score			Factor index score			
		[1]	[2]	[3]		Mean	SD	Range		Mean	SD	Range
[1]	Lately, I've missed out on my social life because of my care giving tasks (8)	+4	-2	-1	-	-	-	-				
	Social workers recognise that because of my experience as informal caregiver I know what I am doing and I that I am well capable to indicate whether and when I need assistance (15)	-3	-1	+1	2.2	6.5	(3.6)	3-15	[1]	16.6	(7.5)	8-40
	I am satisfied with the assistance we get from care and welfare organisations (16)	-4	+1	+1	2.0	8.2	(4.7)	4-20				
	Care giving makes me feel good (39)	-1	+3	+4	1.9	1.9	(1.0)	1-5				
[2]	If possible, I'd like to reduce my informal care tasks (2)	0	-3	-1	2.9	8.6	(4.3)	3-15				
	The person I provide care to most wants to be cared for by me (6)	-1	+4	-2	3.6	14.6	(5.5)	4-20				
	I need information and advice about the best way to organise and carry out my care giving tasks (18)	+1	-3	-1	3.4	10.1	(4.2)	3-15	[2]	33.2	(7.7)	10-50
	I do not mind asking someone for assistance provided that I feel I need help (24)	-2	-4	+2	-	-	-	-				
[3]	I appreciate it when someone asks me how I am doing, how I am coping with my care giving tasks (23)	+1	+2	+4	4.3	17.1	(4.2)	4-20				
	Now and then I feel depressed and despondent because of my care giving tasks (25)	+1	+2	-3	1.6	4.9	(3.2)	3-15	[3]	44.2	(6.2)	26-58
	I see it as my duty to carry out this care giving task (28)	-1	0	+3	4.4	13.2	(3.0)	3-15				
	My circle of acquaintances leaves the care giving tasks entirely to me (36)	-1	+1	-4	2.3	9.0	(5.5)	4-20				

[a] factor scores from factor array (distribution -4, disagree most, to +4, agree most). [b] Item scores of statements with a negative factor score were reverse scored.

Finally, to correct for the effect of the selection of statements for each factor (i.e., the number of statements per factors and their factor scores) on the intra-caregiver comparability of factor index scores across factors, the scores were standardized (mean = 50; SD = 10). These standardized factor index scores were interpreted as likelihood of factor membership, with a higher score indicating higher agreement with the factor. The first columns of Table 3 show the standardized factor index scores for each factor, with considerable variation around the mean.

Table 3: Standardized factor index score

Factor	Total sample (n=249)				Most likely factor [1] (n=79; 32%)				Most likely factor [2] (n=93; 37%)				Most likely factor [3] (n77; 31%)			
	mean	SD	min	max	mean	difference			mean	difference			mean	difference		
						mean	min	max		mean	min	max		mean	min	max
[1]	50.0	10.0	18.6	61.5	56.9	-	-	-	46.0	-12.5	-0.1	-43.1	47.8	-11.9	-0.6	-47.0
[2]	50.0	10.0	27.1	68.9	44.7	-12.2	-0.1	34.4	58.5	-	-	-	45.2	-14.5	-0.1	-45.2
[3]	50.0	10.0	20.6	72.3	45.1	-11.8	1.6	36.9	46.1	-12.4	-0.1	-32.5	59.7	-	-	-

For each participant, the most likely factor membership was determined on the basis of the maximum standardized factor index score across the three factors. The top row of Table 3 shows the result of this nose counting, with a fairly equal distribution of our sample over the three attitudes toward respite care. The table also presents the mean standardized factor index score according to factor membership, and the difference between this score and the scores on the remaining factors. What stands out from

Table 3 is that, when looking at the minimum difference in standardized factor index score between the most likely factor and the second-best, for individual respondents this difference can be very small. This indicates that we should be cautious with interpreting most likely factor membership in terms of "person X has attitude Y".

3.2.2 Methodological issues

The approach proposed by Brown has a strong appeal. Evaluating a set of statements using a Likert-type scale is a very common measurement technique in questionnaires that is easy to administer in interviews and paper or web surveys. The burden for respondents is low.

There are however a few issues that deserve discussion. First of all, as in the Talbott Q block, this approach rests on the assumption that a factor can be well represented by a limited number of statements from the full Q set. It is however uncertain whether, taken out of their context, individual statements will have a similar meaning to participants. Agreeing or disagreeing with a single statement is simply a different exercise from positioning this same statement in a score sheet relative to 30 or 40 other statements. Second, it is not very clear how to select statements for use in a survey and how many statements are needed. More statements may be preferred to less for reasons of completeness, but may also make it more difficult to determine likely factor membership. Third, although the calculation and interpretation of the factor index scores is pretty straightforward, it is difficult to communicate this approach and its results to a lay audience. Finally, it is not a very good way to count noses. Irrespective of the size of the difference in standardized factor index score

between the most likely and the other factors, it remains questionable whether it is valid –theoretically or empirically- to assign people to a certain factor. It may also contradict some of the basic ideas behind Q.

Notwithstanding these issues, we observed some plausible associations between likely factor membership of informal carers and characteristics of their care giving situation. This is encouraging, but further tests of this approach are warranted.

4   Self-categorisation to abbreviated factor descriptions

4.1   *Description of the approach*

The experiences with the Talbott and Brown approaches, and in particular the concern about a limited sub-set of statements being representative for the meaning of a full factor, have motivated researchers to think of alternative approaches for investigating likely factor membership in larger survey samples.

An approach that has been under investigation in recent years is self-categorisation to abbreviated factor descriptions. Participants in a survey are presented with summary description of the Q factors and are asked to indicate the degree to which each one is similar to their own point of view on a subject. The rationale behind this approach is that respondents evaluate the factors that emerged from the Q study as a whole: that is, the statements remain in their factor context and not as separate stimuli.

4.2 Some recent applications

This approach is still in an early stage of development, and has evolved over the past few years. We briefly discuss preliminary findings from three ongoing studies.
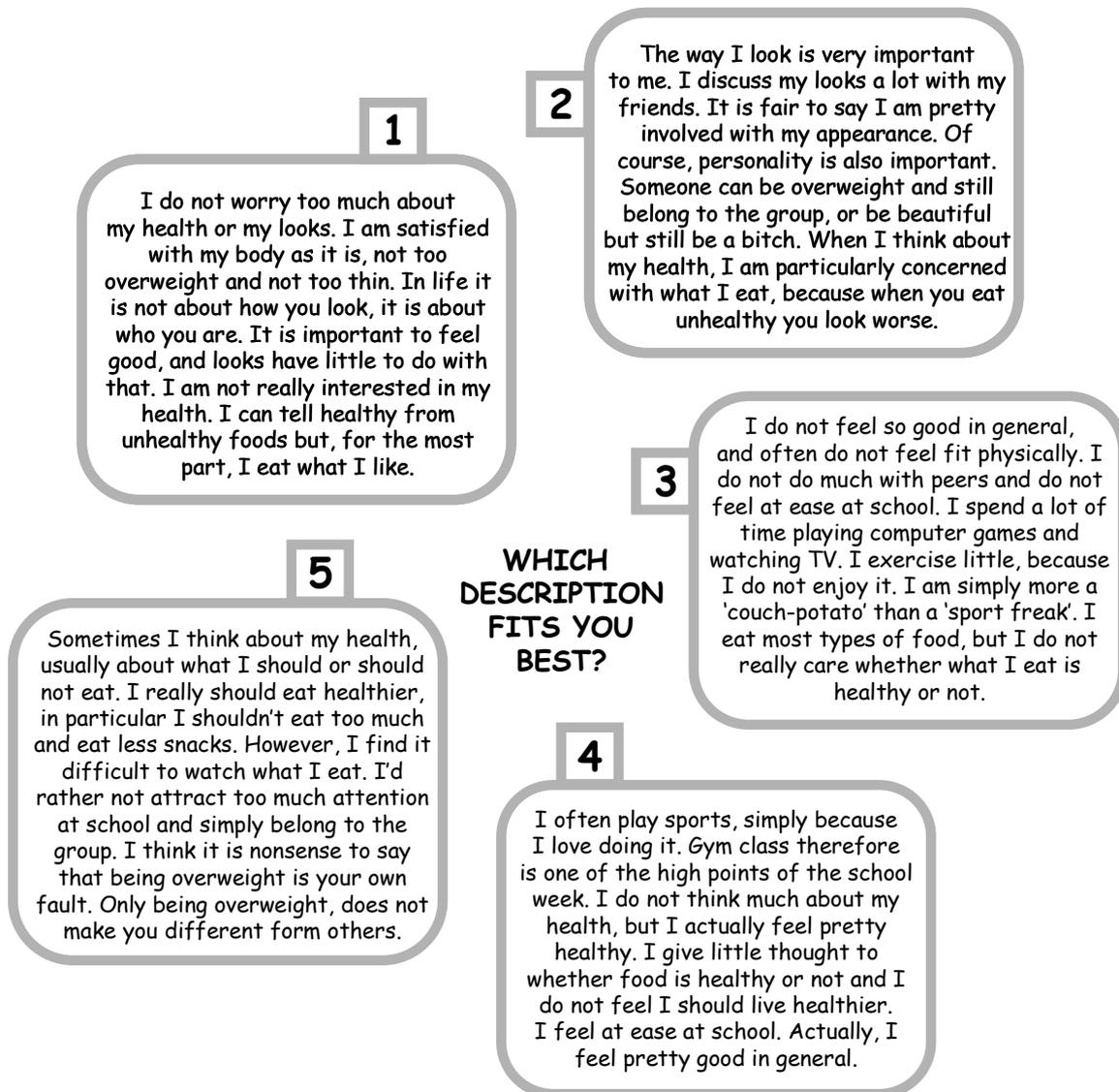
Attitudes about health and lifestyle

An early example is a survey conducted in 2005 in the Netherlands among young adolescents (Van Exel et al., 2006c). This survey contained abbreviated descriptions of five discourses about healthy lifestyle obtained from a Q study conducted in the same population: "carefree sporty", "worrying dependent", "contended independent", "looks over matter", and "indifferent solitary" (van Exel et al., 2006a). For this purpose, the full length factor descriptions were summarized into very short descriptions of less than 100 words each, which reflected the content of the original factors as closely as possible, although with a slight emphasis on the distinctive components. These short descriptions were compiled using fragments of the salient and distinguishing statements of a factor in combination with interview materials from the Q study, in particular the explanations young adolescents gave with their ranking of the statements. In the questionnaire, respondents were presented these five abbreviated factor descriptions (see Figure 2) and asked to indicate which of these five fitted best with their healthy lifestyle attitude.[13]

---

[13] Based on the expectation that respondents would not necessarily be of a 'type' but that there might be a bit of every 'type' in each respondent, we originally intended to have a Likert-type scale accompanying each summary description, so that the response would be the extent to which each description fitted with the healthy lifestyle attitude of the respondent. However, because we wanted respondents to read all descriptions before evaluating them, did not have a practical way for

Figure 2: Abbreviated factor descriptions

2
The way I look is very important to me. I discuss my looks a lot with my friends. It is fair to say I am pretty involved with my appearance. Of course, personality is also important. Someone can be overweight and still belong to the group, or be beautiful but still be a bitch. When I think about my health, I am particularly concerned with what I eat, because when you eat unhealthy you look worse.

1
I do not worry too much about my health or my looks. I am satisfied with my body as it is, not too overweight and not too thin. In life it is not about how you look, it is about who you are. It is important to feel good, and looks have little to do with that. I am not really interested in my health. I can tell healthy from unhealthy foods but, for the most part, I eat what I like.

3
I do not feel so good in general, and often do not feel fit physically. I do not do much with peers and do not feel at ease at school. I spend a lot of time playing computer games and watching TV. I exercise little, because I do not enjoy it. I am simply more a 'couch-potato' than a 'sport freak'. I eat most types of food, but I do not really care whether what I eat is healthy or not.

5
Sometimes I think about my health, usually about what I should or should not eat. I really should eat healthier, in particular I shouldn't eat too much and eat less snacks. However, I find it difficult to watch what I eat. I'd rather not attract too much attention at school and simply belong to the group. I think it is nonsense to say that being overweight is your own fault. Only being overweight, does not make you different form others.

WHICH DESCRIPTION FITS YOU BEST?

4
I often play sports, simply because I love doing it. Gym class therefore is one of the high points of the school week. I do not think much about my health, but I actually feel pretty healthy. I give little thought to whether food is healthy or not and I do not feel I should live healthier. I feel at ease at school. Actually, I feel pretty good in general.

randomizing the order in which the descriptions would be presented to respondents (to prevent question order bias), had no clear ideas about how to decide on factor membership in case of tied scores, and external pressure to deliver a distribution over a short time period, this second-best alternative approach was chosen.

Approximately 2,000 adolescents aged 12-14 years participated in the study. The resulting distribution of respondents over the five factors was plausible ("carefree sporty" 39%, "worrying dependent" 21%, "contended independent" 14%, "looks over matter" 24%, and "indifferent solitary" 2%), and some interesting associations were found with background characteristics (e.g., exercising, body image, personality dimensions, loneliness, the probability of being overweight), providing some support for the validity of this approach.[14]

Attitudes toward health care delivery and self-management

In a second study conducted late 2006, we asked adolescents with chronic conditions to evaluate abbreviated descriptions of four attitudes toward health care delivery and self-management (Figure 3) (Jedeloo et al., 2009). About 1,000 adolescents completed a web survey and indicated how well these attitudes fitted them, using a Likert-type scale (1= 'not at all'; 5 = 'very well'). It was reassuring that hardly any participant rated all four descriptions with a very low score, indicating that no significant view in the larger population was missed in the Q study, and the correlations between the ratings of factor descriptions were all weak, meaning that the factors were fairly distinct. More than half of the participants (56%) indicated that the attitude "Conscious & Compliant" fitted them (very) well (i.e., a score of 4 or 5), 16% viewed themselves as a "Backseat Patient", 26% as "Self-confident & Autonomous" and 25% as "Worried &

---

[14] These results were presented at the 2005 ISSSS Conference in Vancouver, Canada.

Insecure".[15] We were able to match 58% of the participants to a factor based on them having a unique highest score on that factor of 4 or higher. Of the remaining respondents, half showed a tie on the highest score (i.e., giving two or more attitudes a score of 4 or 5) while the other half scored all profiles with a 3 or lower.

Figure 3: Abbreviated factor descriptions [16]

How well does this profile fit you?

A = Conscious & Compliant

If I just do what the doctors say, I don't have to worry about my disease or health. After all, my disease permits living a normal life. I think I am pretty independent now and would like to be treated like an adult. That is why I want the doctors to talk to me and not to my parents. Health professionals should point out my own responsibility to me, for I don't want to regret it or be confronted with my disease later, for not following the rules now. I know enough about my disease, but I would like to know the consequences of a treatment for my daily life. Also, I don't need any support from the hospital in coping with my disease, I will take care of that myself, or my parents will help me.

B = Backseat Patient

I am not too bothered about my disease; my parents will take care of this. They are helping me with my treatment and see to it that everything is in control. I think it will be pretty difficult to take care of this myself. I don't need to know everything in detail, if my parents do, it's okay. When I'm at the doctor's, it's convenient to have my parents with me; they can do the talking for me. I find this difficult and actually, I'm not too interested. But the doctors shouldn't treat me like a child, nor tell me what I'm doing wrong. That's my parents' job. I'm not ready yet to leave the children's hospital and go to adult care.

C = Self-confident & Autonomous

Living easy now is more important to me than being completely compliant. It's me who decides how to live; I won't have my disease limit me in daily life. I'll change my own treatment if this suits me better. I am capable of doing this myself, for I know enough about it. I do want to be treated like an adult and do my own talking. I don't need my parents for that. No one needs to hold my hand, but I do find it convenient to get information. Though this does not mean I will always use it.

D = Worried & Insecure

I'm pretty worried about my disease. Suppose things will turn for the worse later. Sometimes, through my illness, I don't feel at ease with myself, but I try not to think about it too much. I'd rather pretend there's nothing wrong with me. I do not always take advice too seriously, that way I don't have to think about my disease. But I do need my parents and health professionals to help me discuss difficult subjects that are important for my future. I feel more secure about myself if they tell me how to live best.

Using the Likert-type scores, various anticipated associations were found between agreement with the attitudes and characteristics of the young patients, their disease, and its impact on their life.

Views about the prioritisation of health care

In a recent study (European value of a QALY (EuroVaQ) project, 2007) we asked participants to indicate how much five abbreviated Q factor descriptions were like their point of view on how health care services should be prioritized across members of the population[17]. The question in this study differed from the two previous examples in that we used a wider Likert-type scale (1 = 'very unlike my point of view'; 7 = 'very much like my point of view') to give respondents more opportunity to differentiate and, if respondents still tied across two or more points of view, they were asked to break the tie.

Using this procedure, 37% of respondents were matched to a factor directly based on a unique highest score on that factor of 4 or higher, 2% were not matched to a factor because the maximum score on any factor was 3 or lower, and 61% were matched to a factor based on them breaking the tie (on a highest score of 4 or higher).

---

[17] This work is part of a wider European Commission funded project which will report late in 2010. As yet unpublished it has been presented to the European Conference on Health Economics in Helsinki in July 2010.

Based on our experience from these three studies looking at self categorisation to brief descriptions, the best approach to scoring appears to be a combination of Likert-type scale and ranking to break the ties. The Likert-type scores provide an indication of how strongly the descriptions appeal to the sentiment of respondents and how well the factors represent the different points of view in the wider population. If an important point of view was overlooked, a considerable proportion of the population may not feel represented in any of the descriptions and may be expected to return low scores on all. This strength of preference information is also useful for investigating associations with, for instance, socio-demographic characteristics of respondents. Unbreaking ties is important when one wishes to provide an indication of the distribution of factors across a population, as respondents then need to be matched to a single description. The usefulness of such a proportional distribution is however debatable and depends on whether one believes people can be seen as being of a single 'factor type' and to what extent the relative score on other factors also matters. The fact that two or more descriptions were given the same score on a likert scale means that people were fairly ambiguous between at least two factors.[18] However, if one believes for instance that there is a bit of every 'factor type' in everyone, such a distribution is of much less interest (as may be unbreaking any ties).

---

[18] Suppose there are three factors, scored on a 7-point Likert-type scale. Does it make a difference whether the relative strengths of preference were 6 / 6 / 1 (and tie unbroken), 6 / 5 / 1 or 6 / 1 / 1? Does it make a difference whether the tie was unbroken at 6 / 6 / 1 or at 3 / 3 / 1? These

5 Discussion and conclusion

Q methodology was first conceived of as a method to study, in depth, the nature of subjectivities. It provides us with a powerful set of techniques, drawing on qualitative and quantitative traditions, to elicit the range of views and values that exist around a given topic, and to generate rich descriptions of those shared accounts. Typically Q analysis ceases at this point of factor description but there are times when, for research or policy reasons, the prevalence and distribution of such accounts, and their connection with other attributes may be of interest. For this reason we engaged in the use of Q techniques in the design and analysis of surveys.

In this paper we have described three approaches to the application of Q techniques in survey research. Basing our work on a limited methodological literature, we have followed Talbott (1963) and Brown (2002) and adapted these approaches. In doing so we identify a number of methodological issues relating to:

i. the selection (and grouping together) of statements from a Q set to represent the salient features of factors for questionnaire design;

ii. the application of different scoring techniques to rank-ordered and Likert-scale questionnaire responses and the impact of scoring systems on results;

iii. the interpretation of (probable) 'factor membership' derived from these approaches, with particular reference to the observation that all respondents are likely to have *some* degree of association with all factors;

iv. the significance of the distribution of factors in a larger population, and the relevance of scores on other factors relative to the highest absolute score used to match a respondent to a single factor.

v.   the lack of a methodological standard for questionnaire research using Q

techniques and some suggestions for future methodological research in this

area.

The studies referenced here represent a programme of work and the evolution of

ideas and practices through exploratory, methodological studies.   We have made

some inroads into issues of questionnaire design and the advantages and

disadvantages of different methods to assign factor membership to large samples of

survey respondents. In the future we see this research agenda progressing further and

incorporating more sophisticated analytic methods that would attempt to deal with

some of the issues above. By way of example, we are currently exploring the potential

for modelling respondent choices using a 'discrete choice approach' to analysis and the

random utility model as a framework for the choices made. In this approach z-scores

would be treated as choice characteristics in the analysis and represent the

contribution of each statement to each factor (with, arguably, greater precision than

the factor scores). Another approach which has potential and warrants further

exploration, recently proposed by Kroesen and colleagues[19], is the application of latent

class analysis to match respondents to factors. Finally, emphasizing that this is a field

[19] This study (unpublished), in which they used a selection of statements from their Q study on
discourses of aircraft noise annoyance (Kroesen, M & Bröer, C (2009). Policy discourse, people's internal
frames, and declared aircraft noise annoyance: An application of Q-methodology. Journal of the
Acoustical Society of America, 126(1), 195-207.), was presented during the 'Q & Coffee' meeting at
Erasmus University Rotterdam, 12 March 2010.

in evolution, we are currently pursuing an integrated approach by means of Q study in

a large sample[20] followed by a survey questionnaire.

By progressing both questionnaire design and analytic methods, we hope to highlight

the advantages of Q methodology both as a set of techniques for in-depth, small

sample methodology *and* as a valid, reliable method for investigating the distribution

of attitudes and values in large population samples. As such Q methodology has a

unique capacity to connect qualitative and quantitative projects within a single

methodological approach.

---

[20] Given current limitations in PQMethod, this can be a maximum of 299.

References

BAKER, R., BATEMAN, I., DONALDSON, C., JONES-LEE, M., LANCSAR, E., LOOMES, G., MASON, H., ODEJAR, M., PINTO-PRADES, J.-L., ROBINSON, A., RYAN, M., SHACKLEY, P., SMITH, R., SUGDEN, R. & WILDMAN, J. 2010. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. Health Technology Assessment

BROWN, S. R. 1980. *Political subjectivity: applications of Q methodology in political science,* London, Yale University Press.

BROWN, S. R. 1996. Q methodology and qualitative research. *Qualitative Health Research,* 6**,** 561-567.

BROWN, S. R. 2002. Q-technique and questionnaires. *Operant Subjectivity,* 25**,** 117-26.

EUROPEAN VALUE OF A QALY (EUROVAQ) PROJECT. 2007. Available: http://research.ncl.ac.uk/eurovaq/ [Accessed 22/09/08].

JEDELOO, S., VAN STAA, A., LATOUR, J. & NJA., V. E. 2009. Preferences for Health Care and Self-Management among Dutch Adolescents with Chronic Conditions: A Q-Methodological Investigation. . *International Journal of Nursing Studies* 47**,** 593-603.

STENNER, P. & STAINTON ROGERS, R. 2004. Q methodology and qualiquantology: the example of discriminating between emotions. *In:* TODD, Z., NERLICH, B., MCKEOWN, S. & CLARK, D. (eds.) *Mixing methods in psychology.* London: Routledge.

STEPHENSON, W. 1953. *The study of behavior: Q-Technique and its methodology,* Chicago, The University of Chicago Press.

TALBOTT, A. 1963. The Q block method of indexing Q typologies. *AEJ Convention.* Lincoln, Nebraska.

VAN EXEL, J., DE GRAAF, G. & BROUWER, W. 2007. Care for a break? An investigation of informal caregivers' attitudes toward respite care using Q-methodology. *Health Policy,* 83**,** 332-342.

VAN EXEL, J., DE GRAAF, G. & BROUWER, W. B. F. 2006a. "Everyone dies, so you might as well have fun!" Attitudes of Dutch youths about their health lifestyle. *Social Science and Medicine,* 63**,** 2628-2639.

VAN EXEL, J., MOREE, M., KOOPMANSCHAP, M., SCHREUDER-GOEDHEIJT, T. & BROUWER, W. B. F. 2006b. Respite care - An explorative study of demand and use in Dutch informal caregivers. *Health Policy,* 78**,** 194-208.

VAN EXEL, J. A., DE GRAAF, G. & BROUWER, W. B. F. 2008. Give me a break! Informal caregiver attitudes towards respite care. *Health Policy,* 88**,** 73-87.

VAN EXEL, N., KOOLMAN, X., DE GRAAF, G. & BROUWER, W. 2006c. Overweight and obesity in Dutch adolescents: Associations with health lifestyle, personality, social context and future consequences: methods & tables. Rotterdam: Erasmus MC, iMTA.