Using Q methodology for (comparative) policy analysis

Astrid Molenveld

Introduction

Objective policy measures, free from preferences, do they exist? Numerous scholars and practitioners would argue that they do not (Brown, 1974; Howlett, 2009; Peters & Hoornbeek, 2005; Scharpf, 2000; Wlezien, 2004). Authors say that ideas about the nature of knowledge, the approach, envisioned outcomes, spending and politics associated with policy solutions differ among stakeholders, politicians, civil servants and citizens (Peters, 1998). In other words, both the problem and solutions are often viewed by multiple stakeholders, who have different interests and values, from a different angle. This is typically how wicked and complex policy problems are described (Rittel & Webber, 1973; Termeer, Dewulf, & Breeman, 2013).

The basis of these differences is fundamental, people view and perceive complex policy issues fundamentally different, and therefore policy solutions are inherently biased by different problem-frames and value-laden (Brown, 1974; Rittel & Webber, 1973). This becomes clear if you bring stakeholders and public servants with different expertise and knowledge together in the debate about possible policy solutions and instruments (Ansell & Gash, 2007). A policy solution that satisfies one, might be objectionable or create a problem for another individual or organization. These perceptions and views are hard to comprehend, and form ‘perceptual lenses’ (Allison, 1971), through which ideas, frames or policy stories are created around specific issues.
As a consequence, many policy scientists are interested in what frames can be identified and how they influence the debate, collaboration and public policy decision-making.

Next to policy preferences, participation of non-governmental stakeholders in policy decision-making processes often urges the public sector to take into account minority voices, and give stakeholders room to be involved in the debate. Basically, public policy is expected to be legitimate, inclusive and open to an array of stakeholders with diverse opinions. These two fundamental issues: policy preferences and inclusiveness might explain why Q methodology is gaining ground.

Q methodology was already developed in 1935 by William Stephenson to study subjectivity. Nowadays, it is a method that gets increasing attention (see annex) in comparative policy analysis. The method is comparative by nature, as it looks to unravel significantly different patterns in people’s views (Van Exel & De Graaf, 2005; Watts & Stenner, 2012; Wolsink, 2004, p. 2682). At the same time, the approaches, goals and research questions in Q-methodological applications differ, but also show clear similarities. In this book chapter, the methods’ state of play in comparative policy analysis is clarified. On the basis of a review of forty articles, and the most prominent Q methodology textbooks and articles, this chapter will distinguish different elements on which Q methodology applications in comparative policy research vary. The chapter will show different conceptual ideas, approaches and the lessons learned from these applications, which might be of interest to researchers planning a Q-methodological study.

Philosophy

Q methodology is a methodology to study perceptions and prevalent discourses that exist within a given population (Brown, 1993; Van Exel & De Graaf, 2005), e.g. a policy community,
a governance network trying to govern a common pool resource, or civil servants in the center of
government. It is a distinct methodology, it entails a specific ontological stance, a philosophy and
an method, which are quite distinct from other methodologies (Ramlo & Newman, 2011, 2015;
Stenner, Watts, & Worrell, 2007). It has many qualitative aspects and uses statistical analysis to
unravel peoples’ preferences (Brown, 2008), and has a different ontological basis according to
some (Ramlo & Newman, 2015), compared to other methods. Susan Ramlo and Isadore Newman
(2015) explain that some call it a constructivist (or a qualitative) method, and some call it a
positivist method (or a quantitative method). Moreover, they state that Stephenson just said that is
it ‘the objective study of subjectivity’ (p. 174). Q methodology is, in essence, a qualitative method,
which assesses whether the opinion of one person differs qualitatively compared to another person
(Stenner et al., 2007), which uses quantitative factor analysis to indicate possible areas of
(dis)agreement. It is studying individual viewpoints – which is qualitative - through the use of a
systematic (factor) analysis, which is the quantitative element. This systematic process of assessing
qualitative data in a quantitative way might lead some researchers to think that it is a positivistic
approach. It does, however, has an interpretative nature as well, which will be clear once two
important steps are explained: how the concourse is build ex-ante (which will be explained in the
next paragraph), and how factors are interpreted ex-post (later in this chapter).

was a psychologist and used his Q-technique it in clinical settings to assess individuals (Van Exel
& De Graaf, 2005). He was a student of Spearman, and convinced that individuals should be
studied as a whole, and not in individual items or characteristics of the individual (Watts & Stenner,
2012). Stephenson advocated and developed a method for the systematic study of subjectivity
(Stenner et al., 2007, p. 212). Spearman developed his R-correlation, which is an item analysis,
Stephenson developed an inverted analysis, which enabled him to compare people and their viewpoints (Watts & Stenner, 2012). This by-person analysis shows which respondents express their viewpoint in a similar fashion, and have similar propositions about the debate. In other words, it is useful to describe the population of viewpoints (Van Exel & De Graaf, 2005). This by-person factor analysis – leading to a cluster of respondents sharing the same discourse – makes the methodology an attractive method for policy analysts who try to uncover why conflicts emerge, which frames exists and how policy specificities are perceived. While there are still discussions about whether Q methodology is a quantitative or a qualitative method, the focus on people as variables instead of single items, makes the method distinct (see table 1).

Many of the applications are geared towards large groups of respondents, this is particularly true in public policy analysis (none of the articles use an intensive individual study). However, when Q methodology was applied in its early days it was use to explore the standpoints of an individual (Brown, 1993; Stenner et al., 2007), as a holistic subject.

Thus, Q methodology can be applied to conduct:

- **Intensive individual studies**: to reveal how a person thinks about different constructs. In such a study the researcher can use the same set of statements, however (s)he varies the object, e.g. a civil servant has to sort the same set of statements about how (s)he sees certain users of a particular public service. For instance, a low-income single mom and a migrant middle-aged man.

- **Population studies**, such studies focus on a larger group, i.e. a given population and studies significant differences between respondents in a sample of the population. All articles, under review in this book chapter, apply this approach. Therefore, the remainder of this chapter will deal with population studies.
A Q-methodological study often starts with exploring and develop insights in the viewpoints in a particular population about a certain topic. Built on the viewpoints within this population, the researcher develops knowledge about the breadth of the debate – i.e. typical called ‘concourse’ in Q methodology. The researcher tries to grasp the concourse with a representative set of statements (a subset of the whole debate – Webler, Tuler, & Krueger, 2001) which the same, or different respondents in a later stage have to sort on a scale from ‘most agree’ to ‘least agree’. Once the statements are chosen the researcher presents this sample – called the Q-sample, with the Q referring to Questions – to respondents.

By sorting the statements (i.e. the Q-sorting process) respondents express their opinion about the debate. They have to rank statements, according to their individual opinion. Rarer, especially in policy analysis literature, are studies on the basis of photos, pictograms, or single words. Based on the place where respondents place particular statements, the researcher can recognize patterns. In other words, the researcher looks for commonalities in the way respondents have sorted the statements, instead of their individual characteristics (like gender, profession, age etc.). After the factor analysis, it is up to the researcher to interpret the patterns of argumentation which are revealed by the respondents.

After this short introduction, a more thorough discussion on the method will be provided in the next paragraphs – step per step – based upon the applications in comparative public policy. These applications were selected based on a systematic search for scholarly literature applying Q methodology. How this selection was executed is explained in the annex.
How to execute and analyze a Q methodology study

Without going into much detail, a short overview of the methodology will be given in the following paragraphs. However, the main focus of the text underneath will be on how the method is applied, and its potential for comparative policy analysis. More information on the method can be found in the excellent work of (at least) the following scholars: McKeown and Thomas, (1988), Brown (1993), Watts and Stenner (2005), Ramlo and Newman (2011), Jeffares and Skelcher (2011), Van Exel and De Graaf (2005) and many others. A list of prominent Q methodology books and informative articles which can be used to get acquainted with the methodology is included in the last part of this chapter, just before the bibliography. Q methodology roughly exist of six steps, via these six steps the method and the range of applications is explained: gaining insight in the concourse, selection of statements, selection of participants, ranking by participants, analysis, defining the discourses, and drawing conclusions based upon the discourses.

Gaining insight in the concourse

The first step in Q methodology is defining the concourse, i.e. the debate about a certain issue. In popular speech, the word concourse means something like ‘the crossroads’ or ‘the place where different streets meet’ or from the Latin word concurus, which means something like running together (Davies, Blackstock, & Rauschmayer, 2005). Both the idea of the crossroads and the notion of ‘a place’ are important to understand what Q methodology is about. A debate about a particular issues often exists of multiple patterns of thought (cfr. the streets). In debates that are fought hard – like climate change policy – different opinions are often already formed, which leads to ideas about how these particular patterns might look. Other debates might be less well known or clear to the researcher, for instance about a topic in communities that do not necessarily meet or know each other. These communities sometimes have a common background, like professional,
organizational or in terms of experience, which researchers expect to inform their patterns of thought. The insight that Q methodology can offer in a debate in a policy community or program makes it suitable to use for evaluation or to grasp the ideas behind a policy program (Dickinson, Jeffares, Nicholds, & Glasby, 2014a, p. 837).

Scholars apply different approaches to gain insight in the concourse. In terms of the sources which are used to cover the concourse, researchers often turn to existing concepts from the literature, policy documents, interviews, focus groups, mass media, ‘grey literature’, etc. Studying the population a priori ensures that the statements which are shown to the respondents, later on, are representative for the concourse or discourses existing in the population. Representativeness refers in Q methodology to the representativeness of the statements for the concourse in the population, and not to the participants. However, the statements have to represent the debate and are not exhaustive, it is rather a subset of the whole discourse (Webler et al., 2001). Understanding the real debate is of key importance, because the factors that are derived latter are based on the ordering of the statements. If the statements do not represent both conflicting and overlapping areas of thought the analysis will become fuzzier, and factors will be less clear and pronounced. However, most important: a real understanding ensures that the points are salient for the participants, rather than that they are biased towards the categories of interest of the researcher.

From scholarly literature, most scholars do sketch the dominant ideas that re-appear. Some researchers take those as the basis is to define the breath of the concourse. In this regard, one can clearly distinguish two types of approaches in comparative public policy, grasping this concourse: inductively or deductively. In the forty articles studied for this chapter scholars use Q methodology to further theory on a particular topic (qualitative), as well as to confirm theory or theoretical concepts (quantitative). When applying Stephensons’ ideas rigorously, the researcher has to base
his/her statements on documents, articles, interviews, focus groups, newspapers or any account of the possible concourse and define the concourse *inductively*, instead of pre-defining categories (Curry, Barry, & McClenaghan, 2013; Dimitrova & Kortenska, 2017; Ockwell, 2008). After the analysis one reveals the viewpoints and tries to understand them, rather than check the overlap between pre-existing concepts and ideas. Some researchers go even further, and that they say Q methodology is made to *avoid* ex-ante description of theoretical concepts and ‘psychometric’ tests, that entail de facto: data reduction (Dickinson, Jeffares, Nichols, & Glasby, 2014b; Watts & Stenner, 2005, p. 68). Typically, Q researchers do not formulate hypotheses because of what is written above.

Abovementioned points clearly show the everlasting difference in viewpoints of what knowledge is and how to conduct research. Basically, this is the difference between constructivism and positivism, but it might well be the difference between applied science and theory testing. Instead of wide generalization, Q methodology is especially geared towards theory-building for policy design, and conceptual generalization. In table 1 the distinct features of Q methodology, as explained above are shown.

<table>
<thead>
<tr>
<th>Q methodology</th>
<th>R-oriented methods</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual/viewpoint generalization</td>
<td>Relationships between variables</td>
<td>(Watts &amp; Stenner, 2005)</td>
</tr>
<tr>
<td>Participant led</td>
<td>Researcher led</td>
<td>(Curry et al., 2013; Ockwell, 2008)</td>
</tr>
<tr>
<td>Subjective viewpoints</td>
<td>Pre-conceived hypotheses</td>
<td>(Curry et al., 2013; Ockwell, 2008)</td>
</tr>
</tbody>
</table>
Representing the debate: Q-sample

As the researcher progresses in his/her set-up of the study, the debate about the issue becomes clearer. Subsequently, the information gained about the viewpoints within the population has to be narrowed down by the researcher to statements. Sorting statements is the way in which the participants, that take part in the study express their thoughts about a topic. These statements should represent opinions and not facts, as the method tries to reveal discourse and subjectivity. The set of statements – which is called the Q-sample – is prominent and important, as it lies the basis for the discourses that one finds in the analysis.

Most policy analyists try to come up with different methods to ensure that they select statements from the discourse which are salient to the respondents: pilot studies (a.o. Cuppen, 2012; Curry et al., 2013; Nekola, 2012; D. L. Pelletier, Kraak, McCullum, & Uusitalo, 2000), a Delphi study (Nikraftar, 2013), coding with multiple researchers, asking practitioners for feedback etc. Two approaches stand out, these researchers go even a little bit further to make sure that they understand the debate and account for the most salient topics of the debate. First of all, Peritore (Peritore & Peritore, 1990) includes a researcher to sort the statements to control and understand its position and discourse. Second, de Wijs et al. (2016) include two empty cards to give respondents the opportunity to include their salient points. This latter approach, however leads to a situation in which one cannot conduct factor analysis anymore.

To warrant a systematic selection of statements most researchers use some kind of thematic clustering (or sometimes even a matrix in which statements are placed). One of the more well-cited articles in political science is that of Dryzek & Berejikian (1993). They cluster inductively-gathered statements in a deductively informed matrix, which consists of substance elements (rows) and types of arguments (columns). To come up with the final set of statements, statements were
placed within each cell. This is in line with what McKeown and Thomas (1988) call ‘a factorial design’, which is de facto a list with thematic clusters that shows how statements are derived. Although most researchers account for the number and the salience of the statements, Van Exel and De Graaf (2005) do stress that the selection of the statements is partially based on the impression of the researcher.

Researchers select, on average 25 up to 50 statements from the concourse. In the studies in comparative policy the minimum of statements is 24 (Nijnik, Nijnik, & Bizikova, 2009) and the maximum 80 (Eeten, 2001). While in R-oriented methods researchers have to explain in depth how many respondents they included, and what kind of sample they represent, Q-methodologists have to account the selection of the statements. Once the statements are roughly chosen the researcher decides upon the number if statements, that could fit a normal (or quasi-normal) distribution (e.g. like figure 1). Although almost all articles use a forced distribution, this is not necessary. Brown (1980, pp. 288–289) even shows that the effect of the forced distribution can be neglected. Nevertheless, a forced distribution can make it easier for a respondent to choose extremes, and afterwards fill the grid up with more neutral statements, rather than sorting the whole set in a rank order which is not set yet (Watts & Stenner, 2012). The main issue is not how steep or how many statements there are under each point of the scale, the analysis is about the pattern of the statements. On the steepness of the normal (or quasi-normal) distribution there is in general one advice: if the respondents are very knowledgeable, the researcher can make the distribution more flat (Watts & Stenner, 2012). A more steep distribution, allows respondents to place more statements in the middle rows, which represent the statements on which he/she is neutral. The most
used software PQmethod² – specifically geared towards to analysis of ‘forced’ distributions – necessitates this distribution.

Figure 1: example of a forced normal distribution

Selection of participants: the p-sample

Often comparison is the basic reason why policy analysis researchers engage in a Q methodology study. This, in order to discover (and to try to understand) the multifaceted and conflicting concepts groups of participants have. Often the studies describe wicked issues, in which many stakeholders have conflicting interests, problem frames and strategies. The lion share of the articles studies governance networks or deliberative fora, with a focus on decision-making in the energy, water, or environmental sector. More than 75% of the articles address conflicting or difficulties in understanding each other’s frames of references. This starting point informs the way in which researcher analyses the data, draw conclusions from the data, and reflects back to the pre-defined groups.

Although not necessary to overly account for the selection of participants, most researchers consciously select respondents, this is called a purposive sample. As the factors will be most

² Current release is PQMethod 2.35 with PQROT 2.0 (10-Nov-2014)
strongly demarcated by people who are very opinionated, around eighty per cent of the articles start with selecting respondents based upon their professional, organizational or personal background. Other studies select on the basis of the participants’ different counties, their involvement at different times in a certain policy process, etc. In essence, one only needs to select enough respondents who are relevant in light of the study, people how are able to assess the debate and who are opinionated (Van Exel & De Graaf, 2005). In terms of generalizability, the extent to which a researcher can explain the results based on the respondents’ characteristics is limited. The conclusions only apply to those who took part. Researchers do refer back to the background or characteristics of the respondents, by ‘eye-balling’ and checking whether certain characteristics are more strongly present in a certain factor, or not.

Q methodology allows analysis of a small non-representative sample of respondents, called the P-sample (referring to person). Q methodology can be considered in both the number of participants and the participants necessary to conduct a study, a small N-method. Half of the articles in comparative policy analysis engage around 20 to 40 participants, with 15 being the minimum (Nijnik et al., 2009) and 197 being the maximum (Morinière & Hamza, 2012). Brown (1996)) writes that many Q-sort studies have around 30-50 respondents, and this is considered adequate. Stephenson recommended a person-statement sample ratio of ≤ 2 (Stephenson 1953, cited in Mazur & Asah, 2013, p. 82), which is a balance between a 2:1 and 3:1 statements to respondents (Cotton & Mahroos-Alsaiari, 2015, p. 97; Webler, Danielson, & Tuler, 2009).

**Ranking by participants**

After purposefully selecting participants, the Q sorting process can start. Either online or live, the Q-sorting process is a ranking of statements, often presented on small cards, which are presented to the respondents in a random order (Van Exel & De Graaf, 2005). Nowadays, there
are many options to conduct the Q-sorting procedure via a variety of programs and tools. In policy analysis, the most often used approach is however, face to face, in networks, focus groups or individual interviews with respondents. For obvious reasons these approaches are very different. Via the online line tools (e.g. PoetQ, FlashQ and Q-assessor), on the one hand the researcher cannot influence the respondent, on the other hand the researcher can also not ask for why a certain respondent (dis)agrees with a particular statement. However, doing interviews and the face-to-face dialogue is intensive and less time-efficient. It does give the opportunity to study why certain discourses pop up within the group of respondents. Mechanisms and explanations and can be substantiated with actual quotes and illustrations. Watts and Stenner (2012) call the Q-sorting process a dynamic medium in which subjectivity can be expressed (Stephenson also considered this an essential element of the method).

In interviews or in an introductory page of a software tool respondents can be instructed on how to rank-order the set of Q statements. Most researchers use a ‘structuring principle’, of a condition of instruction as Van Exel and the Graaf (2005) call this. As respondents have to rank order statements – relative to each other – the respondent is ought to know on which topic he/she should rank the statements, and on which scale. As respondents have rank statements in relation to other statements, present in one concourse, and this principle is the leading guideline for the respondent to rank-order the statements. This can be a topic or a central question, e.g. perceptions about joint commissioning in health and social care (Dickinson et al., 2014a), or what is a good participation process? (Webler et al., 2001).

Ranking follows a scale, e.g. from totally agree to totally disagree. Most researchers instruct the interviewees first to sort the statements in three piles: agree, disagree and neutral (see Van Exel & De Graaf, 2005 for an excellent guideline on how to proceed with the interviews). This first step
helps the respondent to get order into the many cards, and proceed more easily with the following step. Thereafter, the respondent fills in the grid. Ideally, one starts from the extremes, as it is easier to pick the one or two statements on which you most (dis)agree, compared to others. If all the statements are sorted the interviewer can discuss the sorting with the interviewer to make sure the grid really represents her/his meaning.

The researcher has to be aware that with Q methodology one takes a ‘photo’ at that particular time about a certain topic. Q methodology does not deal with time, and researcher rarely execute time-series. Interviewees often develop meaning on the topic while they sort the items, and therefore live interviews can be very insightful. While sorting the participants develops his/her own frame. Watts and Stenner (2012) explain that the researcher should be cautious not to interfere too much with the sorting process. Whenever a researcher explains a statements, he/she already attaches meaning to the card.

**Analysis: defining the discourses**

When the researcher has *enough* and *relevant* respondents, one can start the analysis. PQmethod\(^3\) is most often used as analysis program. However, this is a dos-based program. So, more and more researchers use SPSS or other statistical tool for a principal component analysis. However, PQmethod still is the software which is most used. In other programs – other than PQmethod – researchers carry out a Principal Component Analysis on an inverted data matrix, with the Q-sort grids as variables, i.e. the respondents (Watts & Stenner, 2012, p. 12). Some researchers use SPSS or other statistical tools for a principal component analysis (PCA), though Q methodologists recommend and only consider a factor analysis like the one in PQmethod (i.e. a

---

\(^3\) Current release is PQMethod 2.35 with PQROT 2.0 (10-Nov-2014)
Centroid factor analysis) suitable (Watts & Stenner, 2012). This because PCA is not in line with the philosophy of Q methodology. PCA entails de facto: data reduction, and because the richness of the data is valued in Q methodology (i.e. Q-methodologists do not engage so much in data-reduction), a centroid factor analysis is preferred.

The analysis looks for clusters of respondents and for saturation, by placing as many respondents under the number of factors, i.e. every sort has subsequently a high degree of association with only one factor (McKeown & Thomas, 1988, p. 52). Watts and Stenner (2012) explain this by means of a cake. The first part that is taken from the whole cake (all the q-sorts) is the biggest commonality. Afterwards, the piece that is still left is under scrutiny, and the largest degree of commonality is search in this ‘leftover’. That is why the first factor displays often the largest share of respondents, and shows the highest explained variance. To select the factors, most Q researchers will probably advise not to refer to statistical arguments only, but there are a few. A few ‘statistical’ rules of thumb (adapted from McKeown & Thomas, 1988; Webler et al., 2001, p. 438; Webler, Tuler, Shockey, Stern, & Beattie, 2003, p. 112):

- Select only the factors that have **eigen values > 1** (Kaiser-Guttman criterion, Watts and Stenner, 2012). This ensures that the researcher includes factors which explain a large proportion of the respondents.

- **Humphrey’s rule:** select a factor if the cross-product of the two highest loadings make up more than twice the standard error of the study (Brown, 1980, p. 223, quoted in Watts and Stenner, 2012)

- Select factors which at least have **two pure loadings**. A pure loading is a respondent which load significantly on a factor, and only on that factor. This will again ensure that the factors are relevant and helps to establish an explanation of why a certain factor appears
- Look for high variance. Choose individual factors that account for >8 %, and a cumulative variance of >30%. This will highlight how much of the variance the factors explain together.

- Factors should **not be highly correlated**, as Q methodology seeks to understand the different viewpoints in the debate. The interfactor correlations should be < 0.33.

- Others, outlined in Watts and Stenner (2012), are the scree test (p. 108), which looks into a drop in the eigen values, if you plot the eigen values in a graph. Besides that, the parallel analysis (p. 109), which assesses whether the factors and eigenvalues really present distinct factors or whether they appear to be random.

Most Q-methodological researchers explain that it is not enough to base the selection only on statistical rules. As the methods’ final goal is to reveal subjectivity, one needs to go beyond a technical correct analysis (McKeown & Thomas, 1988, p. 54). A few ‘non-statistical’ rules of thumb (adapted from McKeown & Thomas, 1988; Webler et al., 2001, p. 438, 2003, p. 112) are whether the factors are understandable, and relate to theory and practice, i.e. if they ‘make sense’.

After one has included all individual Q sorts, studied the correlations and took out the largest communalities, PQ method asks you to indicate the Q sorts which load significantly on a particular factor. This is called by Watts and Stenner (2012) the ‘eye-ball’ analysis (p. 198). To calculate the threshold significance level of an individual Q sort, one can use:

- \(1.96 \times \sqrt{(# \text{ of statements})} = \text{significance level of } p < 0.05\)
- \(2.58 \times \sqrt{(# \text{ of statements})} = \text{significance level of } p < 0.01\)
- \(3.29 \times \sqrt{(# \text{ of statements})} = \text{significance level of } p < 0.001\)

(Van Exel & De Graaf, 2005, p. 19)
If the researcher has – based on the significance levels above – established which Q sorts load on a particular factor one can ‘flag’ these, and commence. Afterwards, PQmethod produces a text document, if ‘analyze’ is pushed. It gives information about: firstly, ‘pure loadings’: the respondents which load on a particular factor; the respondents which load most highly on a particular factor is called an ‘exemplar’, as he or she is most clearly correlated with the factor (Willis & Jeffares, 2012, p. 547). This informs the researcher about which specific respondents to look into for instance for an illustration or a quote. It also helps to understand why this particular respondent is associated with a particular factor, the researcher can go back to its participant selection criteria and to the interview material to understand the discourse and mechanisms.

Second, the analysis produces ‘characterizing statements’ (Van Exel and De Graaf, 2005), these are statements which are highly scored in a certain factor. The highly scored statements help to reconstruct the discourse following from the factor analysis, and are together with the distinguishing statements key to understand the discourse. Based on this, the researcher can grasp the most important points within this discourse. The distinguishing statements show the basic statements on which the discourses differ, they help as ‘sign-post’, to highlight the significant differences.

‘Distinguishing statements’ are a helpful third. These statements are significantly different compared to other factors. Although they are not always on the extreme of the scale (like the characterizing statements), these are important to understand a certain perspective. Lastly, ‘consensus statements’. These are the statements which are not significantly different, when compared with other factors. In other words, the statements on which the respondents agree.

Most researchers start to get an understanding by studying the extreme statements and distinguishing statements. As these represent the statements on which the respondents are most
outspoken. Factor interpretation can be endless because there are different readings and meaning attached to the factors (Watts & Stenner, 2005), quotes and interviews can help to clarify the factors more clearly.

**Drawing conclusions based upon the discourses**

After defining and describing the statements that belong to a certain factor, the researcher tries to label the discourse, and understand it. Most researchers try to interpret the clusters: what do these respondents think about this debate, and how can we understand their viewpoints? There are different ways to proceed here.

Some researchers delve back into their *respondent selection* and start analyzing whether the factor configuration relates to the pre-defined groups. As is clear from most articles, the discourses often turn up irrespectively of the pre-defined groups. So, what does one conclude after a Q methodological study? Explaining why certain discourses are present is not the key finality of the method. Researchers often they try to sketch areas of consensus and conflict when *comparing* the factors. Others analyze a single factor more thorough, and see it as an ‘advocacy coalition’ (Wolsink, 2004, p. 2675), as the group holds a shared belief in certain aspects of the discourse. If one want to explain *why* a discourse turns up, the analysis can be enriched with focus groups, or by diving back into the interviews.

Q methodology strives for *conceptual generalization* (Watts & Stenner, 2012). In terms of what it delivers, it is clear that it enriches our understanding of policy realities, as it not only shows the majority perspective, but also the minority one. Q methodology does not reduce the data, it does not look into single linear explanation for policy design struggles, but for a rich one: *by*
losing generalization for complexity, this qualitative approach eventually leads to another way of theory building and testing’ (Peters, Fontaine, & Mendez, 2018, p. 137).

NPG and public value scholars urge scholars to think, next to facts and figures, also about values and perceptions. Furthermore, some researchers explain that values need to be give thought when designing a new policy (tool) around a specific theme (Pelletier et al., 2000). When policy options acknowledge different viewpoints – whenever policy analysts or civil servants propose new policy – this might help to overcome a tardy process later on. The factors often show that what will work for one group of respondents (established by a certain factor), will be counterintuitive for others. By enlighten these significant clusters, Q-methodology can provide an evidence-base for (new) policy programs and for the customization of old ones (Curry, 2003). For instance, Vugteveen et al. (2010) show how a Q-analysis can be very beneficial for developing planning scenarios, and Dickinson et al (2014) show how government can approach joint commissioning.

Conclusions

Policy and its design and change is a multidimensional matter which raises many questions that policy analysts would like to unravel (Peters et al., 2018). What are the underlying viewpoints that informed a policy program? What kind of views are there, and where do they overlap and conflict? In this chapter Q methodology is presented as one of the methods that can be included in the policy analysis toolbox to conduct a comparative analysis. This small-N method allows studying a small non-representative sample of respondents, and a small sample of statements (a subset) stemming from the debate about a particular issue. Subsequently, it unravels significant differences in peoples’ viewpoints. The method is not so much geared towards understanding causal mechanisms, and does not deal well with time. However, the methods’ strong points are giving insight in dominant and minority viewpoints, and points of conflict and overlap
in particular policy design issues. These assets can be fruitful to provide an evidence-base for new policy programs, or to re-configure existing ones.

Annex 1

Although Q-methodology is not developed recently, it seems that the added value is more and more recognized. Also in Public administration, and more specifically in studying public policy, Q methodology is a method to apply. Considering the ideas put forward in the first paragraph, this method appears to be a way to comprehend wicked policy problems and stakeholder frames.

To get insight in the Q-methodological applications in (comparative) policy analysis a literature study was carried out. Based upon a systematic search via search engine Scopus, forty articles were studied to grasp how they apply Q methodology. In the search engine two keywords were entered: Q-metho* (because some authors refer to Q-method or Q methodology) AND policy. This led to a long list of 195 articles and 5 book chapters. First of all, book chapters were excluded. The biggest group of articles is from the social sciences (n= 119), followed by environmental science (n= 91). Some articles are earmarked in multiple categories. The main journal is Ecological Economics (n= 10), and afterwards Policy Sciences (n= 8), followed by the Journal of policy analysis and management (n= 5). The oldest article is from 1974, the second oldest is from 1988 (see also figure 2). In the period from 1999 to 2005 the number of articles applying Q methodology in public policy was on average 3, in 2016 this is tenfold (see also figure 2). The abstracts of those 200 articles were scanned for eligibility for the next round, with the following criteria:

- Only articles which have a **comparative element** are selected. Q methodology is in itself comparative. However, only pieces with a comparative focus, in other words with a
specifically mentioned comparative design were selected. This can be in terms of respondents, policy issues, or background, or cases.

- Articles which deal with **public matters** were selected; those about accessible goods, services and decision-making, belonging to the public realm. Moreover, public organizations or civil servants have to be part of the study to some extent.

- If the article deals with policy (making) as **object of study**, it was selected. Articles mentioning ideas that could be possibly used for policy improvement - in general terms - were left out.

- Conceptual and theoretical papers were not taken into account. Only empirical articles/book chapters were selected.

After a first round of reading abstracts and deciding on whether the articles and chapters are eligible for analysis, 108 studies were left. These 108 were more thoroughly read – and 40\(^4\) articles were selected for this chapter.

**Figure 2**: applications of Q methodology in policy studies

Output from scopus, derived in June 2017

---

\(^4\) At this moment there are 61 articles selected. Depending on the feedback at ICPP, and focus of the book this database can be altered
Bibliography


nested model of policy instrument choice and policy design. *Policy Sciences*, 42(1), 73–89.


