Qual Quant (2014) 48:3109–3126
DOI 10.1007/s11135-013-9944-z

Overly ambitious: contributions and current status of Q methodology

Jarl K. Kampen · **Peter Tamás**

Published online: 9 November 2013

Abstract This essay offers a small description of recent contributions and status of Q methodology by means of a review of suggested best practices, a systematic review of practice, and a methodological audit. Both theoretical and empirical study suggest that Q methodology neither delivers its promised insight into human subjectivity nor accounts adequately for threats to the validity of the claims it can legitimately make. These concerns in turn, render the method inappropriate for its declared purpose, the scientific study of subjectivity, and suspect for the full range of ontological perspectives, from (neo) positivist to constructivist.

Keywords Q methodology · Sampling · Validity · Systematic review

> SOCRATES: Great or small, never mind about that: we must first enquire whether what you are saying is true or not. Now we are both agreed that justice is advantage of some sort, but you go on to say "of the stronger." I am not sure about this, and must therefore consider further.
> TRASYMACHUS: Proceed.
> SOCRATES: I will. First tell me, do you admit that it is just of subjects to obey their rulers?
> TRASYMACHUS: I do.
>
> From Plato, *The Republic*, Book I, 339.

J. K. Kampen (✉)
Onderzoeksmethodologiegroep/Research Methodology Group, Wageningen University and Research Center, Hollandseweg 1, Room 5036, 6706 KN Wageningen, The Netherlands
e-mail: jarl.kampen@wur.nl
URL: www.rme.wur.nl

P. Tamás
Onderzoeksmethodologiegroep/Research Methodology Group, Wageningen University and Research Center, Hollandseweg 1, Room 5040, 6706 KN Wageningen, The Netherlands
e-mail: peter.tamas@wur.nl

1 Introduction

In 1961, Wittenborn published an update of the contributions and current status of Q methodology in *Psychological Bulletin*. Since that time the use of Q methodology (QM) has expanded in manners that invite an update to his succinct assessment. The clinical uses of QM found useful by Wittenborn are no longer dominant. Today, QM is presented as providing "a foundation for the systematic study of subjectivity" (Brown 1993: 93). In a typical application a researcher will ask respondents to sort items into an equal number of slots arranged in a quasi-normal distribution. The items sorted can be any set of prompts that can meaningfully be distributed along a continuum. The items are understood to be representative of the full range of items (the *concourse*) about which a researcher wishes to learn. The method assumes that analysis of the rank data produced when a sample of respondents (P sample) sort a sample of statements (Q sample) reveals representations of different points of view. QM is receiving increasing interest in the social sciences because the method promises the strength of both qualitative and quantitative methods (Dennis and Goldberg 1996; Ahktar-Danesh et al. 2008: 760), offers a scientific approach to the examination and understanding of subjectivity (Barker 2008: 918; Robbins and Krueger 2000: 637; Brown 1980), and requires small non-random samples to accomplish these goals (Dennis 1986; Danielson 2009). The designation of the method as Q "is intended to differentiate it from R methodology, which comprises methods used for 'objective' or 'scientific' research in the social sciences" (Brown et al. 1999: 599).

"Rather than worrying too much about epistemological questions," say (Brown et al. 1999: 603), "researchers and practitioners may be satisfied knowing that Q method can supplement R methods, providing additional and different insights into issues and policies of interest." However, concern about epistemological foundations of QM becomes legitimate when, for example, advocates argue that "methodological triangulation" is possible between positivist and post-positivist perspectives (*ibid*; see also, Sell and Brown 1984). Concern is deepened when even sympathetic discussions of QM assert that "claims by the founders of QM that hold that the procedure distances and removes the bias of the researcher are shown to be unfounded and epistemologically naïve" (Robbins and Krueger 2000: 636). Such excursions invite methodological audit otherwise "one risks not really knowing what one knows" (Weimer 1999: 426). In the case of QM the authors found reason to doubt QM's promise as a "method to study subjectivity" (Robbins and Krueger 2000: 642).

This paper is a methodological audit of current QM practice based on a study of suggested best practices, practice observed in all QM studies reported in Web of Science in 2010 and review of the methods used. Our study is limited to sampling, measurement validity, and to some extent, internal validity of the method. We do not discuss external validity of QM, on the ground that its founders and users do not claim this type of validity. Nor do we consider the claim by proponents of QM that the method is consistent with Quantum Theory (see e.g., Brown 1971, 2009). The following sections discuss the theoretical background and best practices suggested for QM by its advocates, followed by a systematic review of reported QM practice. Some conclusions and direction for further research are given in the final section.

2 Theory of Q methodology

2.1 The research objective of Q methodology

The design of empirical social science proceeds through a series of steps (see e.g., Kampen and Tobi 2011). Q methodology, being an empirical social science, proceeds through these steps, roughly delineated as (1) theory design, (2) instrument design, (3) sampling design, (4) analysis design. At the outset of designing the research, the research objective has to be stated. QM focuses on the so-called *concourse*, that is "the flow of communicability surrounding a topic" (Brown 1993: 94). The most frequently cited purpose for Q methodology studies is to produce systematic studies of subjectivity (*ibid*: 93). In QM, the a priori idea prevails that "although social issues are subject to debate and differing values, when individuals are asked about specific elements of a topic *only a limited number of culturally available alternatives are possible*" (see Barker 2008: 920; italics added). The research objective of Q methodology then, is to identify patterns in this concourse: the limited number of subjective representations of views (SRVs), and in order to achieve this aim QM proposes an instrument, and in addition makes assertions with respect to sampling design.

2.2 Constructing the concourse

The task in concourse construction is to identify its components for all analytically relevant subjects at all possible analytically relevant moments in all analytically relevant contexts. It comprises of words, paintings, pieces of art, photographs, musical selections, and so on (Brown 1993: 95). This concourse then is the population from which a representative sample of statements is to be drawn. The concourse, according to Farrimond et al. (2010), "can never be fully known, of course, but the sample of items (usually written statements) should give a workable estimate of it." Needless to say, constructing a concourse is difficult, and practitioners would benefit from clear prescriptions from the literature as to how to compose a sound concourse. However, the QM literature remains uncomfortably silent with respect to how to assemble and verify completeness of a concourse, and how to verify or falsify the representativeness of a sample drawn therefrom. Wittenborn observed that many of the samples of statements to be sorted in QM appeared "to have been somewhat informally assembled" (1961: 138). In order to continue our discussion of QM, for the remainder of this paper we inappropriately assume the concourse as known and the sample of items sorted as representative of that concourse.

2.3 Establishing the subjective representations of views (SRVs)

The process by which Q methodologists produce SRVs proceeds through a sequence of four steps. After the so-called concourse is specified, as a first step a representative subset of statements from the concourse is sampled (so-called Q sample; Wolf 2010). Second, these statements are ranked on a given dimension (e.g., disagree–agree) in a specific format and within a specific context by a suitable number of appropriately diverse respondents (so-called P sample). Third, the patterns of rankings are aggregated by statistical means (factor analysis, cluster analysis). And fourth, the resulting clusters of rankings are interpreted as SRVs.

Consider a dataset produced by making a sample of respondents (P sample) sort a number of stimuli (Q sample) in a pre-specified format (Q sort). In Q methodology, cases (respondents) become variables and variables become cases. Thus, where usually, data are organised in a $P \times Q$ matrix $\mathbf{X}$ where each row corresponds to observed values in a Q-array of variables for the pth respondent, $p = (1, \ldots, P)$, the point of focus in QM is the transposed matrix $\mathbf{X}^{\mathrm{T}}$ where each row corresponds to a sorted stimulus and each column to a respondent. The first approach (focussing on $\mathbf{X}$) is referred to in the QM literature as the "R method approach" (e.g., Brown et al. 1999: 599). In QM, the variance-covariance matrix $\mathbf{Q} = \mathbf{X}\mathbf{X}^{\mathrm{T}}/Q$ is subjected to factor analysis (think of a principal components analysis with VARIMAX rotation, or an obsolete method known as the centroid method of factor analysis by Thurstone; very occasionally, a conventional cluster analysis is applied), where the resulting factors represent clusters of respondents with similar views (e.g., Brown et al. 1999: 601; Ahktar-Danesh et al. 2008: 766; Brown 1993: 111).

Subsequent to establishing clusters, the SRVs are interpreted. First, the factor loadings resulting from PCA are inspected. These loadings correspond at an analytic level to memberships of clusters, and the clusters in turn correspond at a conceptual level to the SRVs. In other words, each SRV is represented by a (VARIMAX rotated) principal component, and the subscription of an individual to a particular SRV is deduced from his/her factor loading. Second, the meaning of each SRV is derived. How this meaning is derived is somewhat unclear. Some authors derive meaning from so-called "exemplar cases", that is cases that have relatively high loadings on a single factor (e.g., Farrimond et al. 2010). Other authors arrive at interpretations by correlating rankings and other participant characteristics with factor loadings (see e.g., Vermaire et al. 2010). Sometimes, the number or proportion of respondents belonging to a SRV is expressed in terms of the explained variance of the corresponding factor (see e.g., Killam et al. 2010; Brown 1980); however, others state that "as opposed to R methodology, QM traditionally intends to give a picture of the perspectives that exist (the variety of perspectives) among the population, rather than analysing the level of support for those perspectives among the population" (Cuppen et al. 2010). The lack of consistency with respect to appropriate interpretation is not unusual between (and sometimes even within) research reports discussing QM.

3 Measurement validity in Q methodology

3.1 QM as a measurement instrument

As said, the objective of Q methodology is identification of SRVs. The assessment of the validity of any research proceeds through three steps: assessment of measurement validity (does the instrument measure what it is supposed to measure), assessment of internal validity (can the findings within a sample be trusted), and assessment of external validity (can the findings within a sample be generalized to the larger population). There is a strict hierarchy in these different types of validity, because lack of measurement validity immediately implies lack of internal validity, and lack of internal validity immediately implies lack of external validity. This section does not consider external validity of QM with respect to the population of respondents as this is rejected already by its founders (see e.g., Amin 2000). Instead, we focus on measurement validity insofar as surrendering this would surrender the usefulness of QM for science altogether.

Brown et al. (1999: 601) claim that QM can provide more in-depth analysis into how respondents think. Ahktar-Danesh et al. (2008: 766) state that "the Q-sorting operation is subjective and represents an individual's view" and "each individual's set of rank ordered statements is deemed a valid expression of his or her opinion." When keeping in mind Kagan's (2009: 130) diagnosis of reasons for the present malaise in social science in which a major one was "the fact that despite many efforts social scientists have failed to measure human psychological states," QM represents an incredible achievement. But exactly how does a Q sort correspond to a view, that is, how does it accurately reflect a respondent's subjectivity? The measurement process by which views are established consists of an unusual mixture of sampling issues (Q sample, P sample), measurement issues at micro-level (the Q sorts), as well as issues of analysis and interpretation at aggregate level. Each step in this process requires detailed discussion before the issue of measurement validity can be fully addressed.

3.2 The impact of clustering method: SRVs as analytic artefacts

QM, say Brown et al. (1999), requires "a small number of well selected subjects." Regarding the sample size, some claim that the number of statements should be at least twice the number of subjects (Nikolaus 2010), while others propose that "it is desirable to have at least four Q statements that represent each expected viewpoint and four subjects per statement" (Schwartz et al. 2010). Apparently, in QM "a larger P sample increases costs for no purpose, as a well selected small sample is adequate to identify the relevant factors" (Danielson 2009: 4). The researcher should of course define "well selected" and then defend which of the mutually inconsistent guidelines they have chosen.

Some basic mathematics can reveal the consequences of choosing large or small P samples. Denote the size of the P sample by P and the size of the Q sample by Q. Using the rules rank $\mathbf{X} \leq \min(P, Q)$ and rank $\mathbf{Q} \leq \min(\text{rank } \mathbf{X}, \text{rank } \mathbf{X}^{\mathrm{T}})$ from elementary matrix algebra, and further applying the suggestions from the literature (see e.g., Dziopa and Athern 2011: 41) that in QM the P sample can be much smaller than the Q sample or $P < Q$, we have rank $\mathbf{Q} \leq P$. Because the number of non-zero eigenvalues of $\mathbf{Q}$ cannot exceed rank $\mathbf{Q}$ (for a proof, see e.g., Schott 1997: 99), the number of extracted factors (groups, clusters) cannot exceed P. This is not problematic, because we would expect no more than P different "views" from P different respondents. QM defies logic however, when $P > Q$, because factor analysis then no longer allows for P different views—it only allows for a maximum of Q different views (and this only in the case that rank $\mathbf{Q} = Q$). Concretely this means that if for instance, 7 billion people sort 100 statements, these 7 billion people are allowed no more than 100 different SRVs. (Of course, if each item was to be scored on a binary scale by each respondent, no less than $2^{100} = 1.3 \times 10^{30}$ different response patterns are possible).

The graveness of this problem follows from the objective of a Q sort, which is to distribute statements in accordance with respondent's subjectively determined preference. The object of the analysis is to cluster these distributions in a manner that represents to the greatest extent possible, the full diversity of viewpoints that can be found, but the model fails to achieve that aim. Even when we accept that "a well-chosen small P sample will provide enough information to define factors that will be present in any additional group drawn from the same population" (Danielson 2009: 6), the mathematics dictate that we are impotent to find *all* the factors (groups, clusters). These mathematical constraints destroy the measurement validity of the method. It is curious that Q methodology studies continue to use factor analysis given that this problem is easily overcome by using conventional cluster analysis, as is done for instance, by Massetti and Bracken (2010).

3.3 The impact of forcing Q sorts: SRVs as measurement artefacts

Taking one step back from the phase of analytically establishing SRVs, a first foundational criterion for measurement validity would be that QM respect respondents' expressions of views. That is, in order to meet the claim that the sorts represent views, the way sorting is structured, minimally must allow respondents to communicate the content and the strength of their view. This, however, is immediately violated by the structure of most Q sorts. The QM literature suggests a forced sort is to be preferred over an "unstructured sort" because it reduces the extremity of views (Block 1961: 56). A forced sort is accomplished through apportioning stimuli into a quasi-normal distribution, that is, respondents are forced to rank only few statements at each extreme and to rank the greatest fraction as (near) neutral (e.g., Ahktar-Danesh et al. 2008: 764; Brown 1993: 102). So, in QM by default, strength is not part of a view.

The forced sort has another analytic consequence, in that it reduces variance of the rankings which increases correlations of rankings across respondents. Higher correlations between respondents increase the likelihood of respondents being clustered together in a factor (group, cluster). While the choice for forced sorts may be analytically productive (i.e. produce better clusters) the method used does not appear to be able to detect if respondents would have been clustered if the Q sort was unstructured. This means that the method is unable to separate clustering by the design of the research instrument from clustering by truly shared views. In other words, it may silently produce SRVs as measurement artefacts.

However, the QM literature is not consistent regarding the issue of preference for forced or unforced Q sorts. Some researchers maintain that "the range and shape of the Q-sort distribution have no effect on the statistical results, as has been repeatedly demonstrated" (Giannoulis et al. 2010; see also Brown 1993, 2009; Cottle and McKeown 1981). That is, "the shape of the grid will not affect the factor structure, and free and forced distribution are therefore both just as valid for factor analysis" (Cottle and McKeown 1981). Since both statements (forcing matters versus forcing does not matter) cannot simultaneously be true, practioners are hard pressed to decide which method to select.

3.4 The impact of reflexivity: SRVs as researcher's artefacts

Finally, the claim of measurement validity of QM further requires that the rankings given in the interview context correspond to rankings that would be given outside of the interview context. In short, QM should be able to distinguish between a respondent's unique subjective (view-driven?) sorting behaviour and behaviours informed by (un)consciously and conscientiously carrying out the researcher's instructions in the interview context. There are a number of phases in QM at which the researcher is active in ways that may introduce bias. A few of these are as follows:

1. Identifying the Q sample, that is

(a) selecting the population from which the *concourse* of statements is derived
(b) constructing the *concourse* of statements,
(c) selecting from the concourse the statements that constitute the Q sample,

2. Selecting the P sample,
3. Guiding the Q sort, that is

(a) Determining the context in which the Q sort is conducted,
(b) Providing instructions for the (forced) Q sort,
(c) Engaging in a discussion with the sorter,

4. Deciding thresholds for "distinguishing items" and "consensus items" i.e. membership and non-membership in groups on the basis of factor loadings,
5. Interpreting the research findings,
6. Deciding how to report the findings.

In the interview context QM researchers provide the respondents with both explicit and implicit instructions. Explicitly, a detailed set of instructions is given to the respondents "depending on the number of statements, distribution of the Q-sort table, and whether the Q-sorting is done using cards or electronically" (Ahktar-Danesh et al. 2008: 765). Also, the researcher can engage in a dialogue with the sorter as this will "aid the investigator to understand more fully the perspectives" of the respondents carrying out the sorts (Brown et al. 1999: 601). In fact, it is claimed that (Duenckmann 2010: 287):

> The interview accompanying the sorting is at least as important as the actual sorting for several reasons. On the one hand, the Q-Sort represents the 'skeleton' of subjectivity, which only becomes interpretable through the comments and reflections about the underlying motives at work. On the other hand, the act of sorting the statements fulfils the function of a guide for a qualitative interview. One particular advantage of this procedure is the fact that during the interview the test persons permanently overview their Q-Sort that represents the outcome of the conversation up to now. This makes it possible for the researcher or the test person to perceive interrelations and inconsistencies and to refer to them directly.

Exactly how the information contained in the open interview accompanying the Q sorting is used is not methodologically prescribed in the QM literature, and a systematic review of QM applications, such as conducted in the next section, should shed light on this issue. Besides this explicit interference of the researcher in the measurement process, implicitly, he or she selects from the universe of possible statements the subset, the Q sample, that is to be sorted (e.g., Robbins and Krueger 2000: 638–639), and then presents this subset in a context where respondents are encouraged to recognise statements as both legitimate and sortable.

The issues of reflexivity must be addressed before it is possible to assess the measurement validity of the method. If reflexivity is not adequately considered, Q sorting has the inherent risk of turning into a Socratic dialogue, wherein Socrates (the researcher) with great certainty obtains the correct responses from Trasymachus (the respondent).

4 A systematic review of QM in practice

4.1 Materials and method

In this section we report on a short examination of how practitioners actually use QM. This review looked at the units of analysis referenced, how items were sorted, how factors were identified, how identified factors were interpreted, and any claims of validity. This review is based on all of the articles found through Scopus that reported using Q-methodology in the arbitrarily chosen year 2010 that could be returned by the library at our university at no cost. In 2010 there were 109 articles in Scopus that explicitly made mention of Q-method (identified using the search string TITLE-ABS-KEY(Q-method OR Q-sort OR "Q methodology" OR Q-methodology OR "Q sort") AND PUBYEAR = 2010). Of those 109 articles, 71 were readily accessible through our library. The 71 returned articles were loaded into and analysed in the software package Atlas.ti (version 7). The entire set of articles and the Atlas.ti data file is available for inspection from the corresponding author. Of the 71 articles, 39 were found to be empirical studies that used QM (see Appendix). Those remaining were either non-empirical studies that promoted QM or empirical studies that used Q sorts and then submitted the results to analytical treatments that did not involve by-person clustering. The analysed articles were published in journals of clinical psychology, public health, education, nursing, ecology, economy, public administration, logistics, dentistry, and management.

4.2 The unit of analysis (P sample)

The size of the P sample (mean = 52.6, SD = 62.0) had a range between 7 and 388. The review of 2010 applied QM literature revealed some confusion regarding the actual unit of analysis, which can be defined as individuals (those doing the Q sorts i.e. the P sample), the SRVs (the product of aggregation of the Q sorts), or 'discourses'. In the articles reviewed it

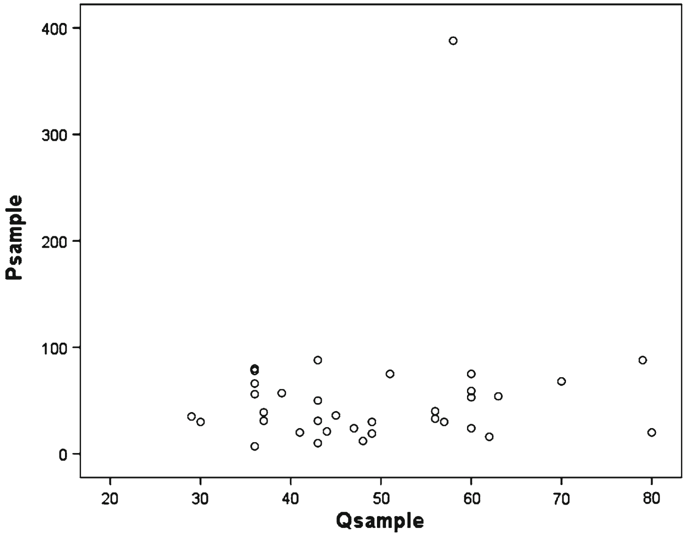

Fig. 1 Scatterplot of Q sample size and P sample size (N = 35, missing = 4)

appears that the dominant unit of analysis is the individual (32 of 39). That is, individuals are most commonly associated with the factors that are derived from QM. In all of the cases reviewed, an individual is associated with a factor when that individual's loading on a factor is statistically significant. Twenty-six of the 32 articles that referenced individuals also referenced either perspectives or discourses. The ambiguity here is important as there is no logical reason to expect a clean relationship between the factors generated by by-person factoring of Q sorts (aggregate level) and the individuals who did the sorting (micro level). In the six articles that did not reference individuals in the rationale, the results or conclusions restrained their interest either to 'perspectives' or to 'discourses.' In this minority perspective, individuals were not forced to fit into factors. Instead, factors were seen to partially explain subjective views and, as such, an individual could be present in more than one factor.

4.3 Validity issues regarding the concourse and Q sample

As said (Sect. 2.2), in QM conclusions are supposed to be (externally) valid with respect to a concourse. The items (Q sample) sorted by respondents are supposed to be representative of that concourse. The size of the Q sample (mean = 48.0, SD = 13.6) ranged between 19 and 80. The ratio of Q sample size and P sample size (mean = 1.62, SD = 1.25) was between .15 and 5.14, with most articles having much larger Q samples than P samples. See Fig. 1.

Every one of the 39 articles provided some discussion of the manner in which the concourse was assembled. Taking an analogy from R research, (external) validity of the Q sample depends on full knowledge of the sampling frame, and in QM the sampling frame is the concourse. There is, as discussed in Sect. 2.2, no way to fully identify this sampling frame. As such, and despite authors' at times exhausting efforts to identify a concourse in the sample of 2010 applied QM literature, it seems that claims of validity with respect to a concourse are not supportable. Taking the next step, from sampling frame to sample, while every article offered some discussion of how sorted items (the Q sample) were assembled, only half (20/39) provided a (however weak) claim that the items sorted were representative of the concourse from which they derived.

Looking at this question from another perspective, QM does not impose requirements of representativeness on subjects (P sample). Rather, the task is to purposively select subjects whose sorts will encompass the maximum diversity of possible views. As with the development of items, every one of the studies reviewed provided some discussion of the mechanisms by which respondents were selected. None, however, provided a test by which they could ascertain whether there was adequate diversity among respondents to support inference from the factors derived from sorts to the concourse from which sort items derive. Curiously, and despite the very clear contrary guidance in the theoretical QM literature, eight of the studies reviewed argued that sorters were representative of a larger population and that this supported validity claims with respect to their respective populations (see also, Sect. 4.7).

4.4 Forced or unforced sorting?

Of the 39 papers reviewed, 33 declared that respondents were forced to sort items into a quasi-normal distribution. Of those 33, one permitted respondents to deviate from the quasi-normal distribution (producing a positive skew; see Pim et al. 2010) and eight used a two stage process identified in the 1960s (Gaito 1962) in which respondents initially sorted statements into three piles and then sorted those piles into the quasi-normal distribution. Outside of this dominant trend, one paper specified a free sort, one paper used a two stage rectangular sort, and three papers did not declare what form of sort was used.

Every one of the articles used Q sort as a survey (cross-sectional inter-respondent comparison) rather than a clinical tool (longitudinal intra-respondent comparison) and, as such, all required comparison or aggregation across respondents. The universal requirement for aggregation combined with the typical use of a forced normal distribution implies the unnecessary (Shill 1966) and undesirable (Gaito 1962) assumption that respondents' views are consistently distributed. Furthermore, the typical use of quasi-normal distributions is inconsistent with early empirical work which found that respondents had no consistent preference for the shape of a forced sort and that a rectangle (only used once in the articles reviewed) produced the most reliable outcomes (Livson and Nichols 1956).

4.5 Methods of aggregation

The papers reviewed followed a variety of analytic strategies. Of the articles reviewed 17 did not declare what form of factor analysis they used. Centroid factor analysis and PCA were used, respectively, both in 10 papers. Once factored, 19 of the papers declared use of VARIMAX rotation[1] while five mentioned that factors were hand rotated. The single exception to this basic pattern is the lonely paper that used cluster analysis followed by visual inspection of a dendrogram to identify "factors". Taking a step back from the mechanics, factor analysis was for many years identified as part of qualitative analysis insofar as factor identification inevitably involves considerable subjective judgement. That said, only ten of the papers reviewed explicitly noted subjective judgement in their analysis.

Statistically significant loadings for an individual on a factor were noted as low as .35 and very few were higher than .5. This indicates that an individual can be assigned fully to a factor when a third of their 'subjective perspective' is reflected by that factor. No paper explained how it was logically defensible or practically possible to assign an individual fully to a factor on the basis of small loadings, nor did they explain what happened to the left-over bits and

[1] The prevalence of VARIMAX rotation is interesting given that the determinism of VARIMAX apparently renders it inappropriate for Q-methodology (Stephenson 1952: 239).

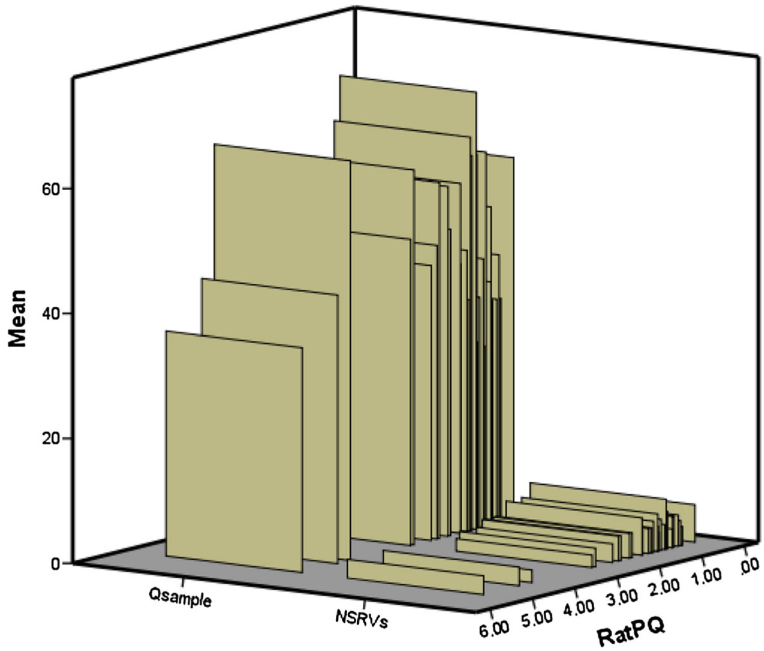

Fig. 2 Relationship between the ratio of Q sample size and P sample size with the number of identified SRVs and the Q sample size (N = 35, missing = 4)

pieces that, when total variance was reported, most often amounted to a greater fraction of the variance than that explained through the identified factors.

Twenty of the papers reviewed referenced the software package PQMethod. Given its frequency of use, we decided to briefly examine the software. Documentation about this program is available at the Internet. We referred to the PQMethod Manual (Schmolck 2012) which is distributed with the PQMethod package (version 2.33). The manual states that PQMethod performs two ways of aggregation of the Q sorts: centroid factor analysis and PCA (with options of orthogonal VARIMAX rotation and oblique "hand rotation" of factors). For both methods the maximal number of factors PQMethod can handle is eight. The size of the P sample is limited to 299 subjects and the Q sample is limited to 200 statements.

None of the reviewed papers using PQMethod refers to the limitation of the maximum number of SRVs they could detect.

4.6 Identification and interpretation of SRVs

The number of SRVs identified (mean = 4.1, SD = 1.43) ranges between 1 and 8, with 3 or 4 SRVs being a typical number (corresponding to respectively 9 and 11 papers). The data show a negative relationship ($r = -.51$, $p < .01$) between the ratio of the Q and P sample, and the number of SRVs. See Fig. 2, which shows that when this ratio is low, the number of identified SRVs is higher. Applied QM researchers used a variety of rationales for identifying the number of SRVs. A first approach was to point to the factor solution, and claim that the reported number of SRVs equalled the number of factors returned in statistical analysis referring to scree plots, eigenvalues, factor loadings, etc. (4 articles), and a related approach that sought to maximize loadings on factors (5 articles). For instance, referring to Brown (1980) and Watts and Stenner (2005), Takshe et al. (2010), state that

> The factors were selected based on two criteria. The first one was that their eigenvalues had to be greater than 1.0 (this is needed to be statistically significant) and the second factor was that a minimum of two Q-sorts had to load significantly on that factor.

A second approach is to claim that no elaborate justification is required because the groups or clusters were self-evident, emerged from the data, were retained on unspecified grounds, etc. (8 articles). And a third approach was to point to the maximal number of "interpretable" factors (9 articles), for instance, the solution "provided the most clear and intelligible portrayal of attitudes of parents with 6-year-olds toward oral health" (Vermaire et al. 2010). Some articles mixed two or more of these approaches, however all of these remaining papers included 'interpretability' as an argument (7 articles). And a number of papers (4) revealed no information about the number of SRVs extracted.

Every well-referenced methodological guide for QM consulted by the authors of this article has argued that the (qualitative) interviews that accompany sorting items are important (see Sect. 3.4). This strong recommendation does not seem to be reflected in QM practice. Of the articles reviewed, only 17 noted that researchers solicited explanatory narratives from respondents. When clearly stated, 14 out of these 17 narratives focused either on extreme placements, or on the rationale for placement generally. Of the remaining three papers, one solicited a narrative in order to be able to identify researcher bias and two solicited narratives for reasons that were not clearly explained. Of the papers that solicited explanatory narratives, one identified the method ("constant comparative", Hurley and Horn 2010) used in analysis. The remaining papers, when it was mentioned at all, suggested that the narratives were used in interpreting factors (SRVs). With the exception of the single case in which constant comparative was used, the lack of specification of an analysis method for the explanatory narratives opens the possibility that researchers, perhaps unwittingly, cherry-pick narratives that bolstered emerging interpretations of the factors.

4.7 Claims of external validity

There seem limited grounds even within the QM literature on which it is possible for research done in QM to make any claims of external validity. It was, therefore, a bit of a surprise to find that half of the papers read (19) made some claim of external validity either to a concourse (e.g., "This Q-methodological study reveals four preference profiles among adolescents with different chronic conditions…"; Jedeloo et al. 2010), or to a population (e.g., "Our aim was to include a diverse sample of patients, giving an adequate reflection of the usual RA outpatient population with all levels of fatigue"; Nikolaus et al. 2010). Eleven of the papers read went so far as to make prescriptions (e.g., "For those seeking to engage with these young people there are three key messages from this research…"; Bradley and Miller 2010). All of these seem somewhat of a stretch even within the framework of the QM literature and would be looked at with even greater scepticism among R researchers.

5 Conclusions and outlook

5.1 Discussion

In the previous sections a number of issues were addressed that raised concern regarding measurement validity in Q methodology. In order to inquire to what extent these concerns are justified, a mini review of applied Q methodology in the peer-reviewed literature was executed.

The first issue regarding QM addressed above regarded the constructing of the concourse (Sect. 2.2). Neither a targeted search in theoretical QM literature nor a systematic review of 2010 applied QM literature uncovered any systematic approach or protocol to construct a concourse, let alone a formal procedure to test its completeness. It cannot even be ruled out that a concourse, as a countable set of statements, does not exist, or that the set is not finite. In any case (not existing, not countable, not finite), interpreting the Q sample as representative for "a concourse" is meaningless just as inferences to a population in R research are meaningless absent both a representative sample and a valid frame.

Our next concern (Sect. 3.2) regarded the size of the P sample in relation to the size of the Q sample. We found that the claim that QM requires only small non-random samples (Dennis 1986) is correct, though surely not, as understood by practitioners, "because there are a limited number of independent views that can be held about almost anything" (Stephenson 1952: 239). The literature on QM suggests that the P sample can be smaller than the Q sample. A simple examination of the dominant aggregation methods used in QM to obtain SRVs revealed that the maximal number of different SRVs in a population of individuals is constrained to the size of the Q sample, so that adding individuals to the P sample will not produce more SRVs. Arguably, a method designed to establish subjective views of respondents that cannot allow as many views as there are respondents is flawed. Conceptually, this flaw returns in QM when it becomes evident that QM users *must* accept the analytically determined postulate that only a limited number of alternative views are possible.[2] The postulate is cemented in the mathematics of the model two ways. First, the factor analysis used limits the number of different views to the size of the Q sample (see Sect. 3.2). Second, the quasi-normal distribution required in the forced sort decreases variance (Sect. 3.3), thereby increasing co-variances (and correlations), and thus increasing the amenability of the data to clustering (e.g., by factor analysis). The postulate is further cemented in the popular software program used for aggregation of the Q sorts (PQMethod, see Sect. 4.5), because it limits the number of factors to a maximum of 8. Where R methods are claimed to be "concerned with individual differences by postulation" (Stephenson 1953: 484), QM appears to be concerned with individual similarities by postulation. While such a priori notions are common in scientific inquiry, QM is incredible in that it is analytically incapable of producing a contradictory conclusion (meaning that QM results cannot be falsified).

Finally, it was demonstrated on a number of grounds that QM is hard pressed to fulfil its promise of measuring the internal structure of subjectivity (Sect. 3.4). For instance, in order to claim that the rankings given in the interview context correspond to views held by respondents in their daily lives, QM must establish the identity of the process of forming a view during the course of daily life and during the course of sorting a stack of cards (while having a conversation with a researcher). At minimum, the interview context forces a respondent to be consciously aware of his or her own decisions, when the same level of conscious awareness may not be found at those moments when the views (thought implicit in the sorting of statements) become relevant in daily life. Insofar as there is no discussion in QM of the possible discrepancies between the consciously produced sorts in an assisted interview context, and the perhaps tacit views that inform conduct in the real world, it is

[2] Briefly stated, the defect is that QM can establish "typical representations of different points of view" if and only if we accept the presupposition that these typical representations exist. This problem has been identified more generally by Kagan (2009: 130) as the fatal flaw that, contrary to natural scientists, social scientists presuppose validity of an a priori idea. Social scientists start investigation with the defining of a concept and then look for evidence for that concept in empirical settings, without specifying the nature of empirical evidence that would lead them to reject the a priori idea.

difficult to accept the argument that the sorts produced by respondents correspond to views held in the real world.

The empirical studies contained in the body of 2010 scientific literature on applied QM, show that nothing is done that allows for distinguishing between the effects of explicit and implicit aspects of the research environment and respondents' unique subjective sorting behaviour. A range of additional questions concerning the non-trivial role of the researcher in shaping the "research findings" at several phases in QM inquiry further compromises its internal validity. Conventional "R type" research is, of course, like QM susceptible to researcher bias. However, in most empirical "R" research, meticulous effort is undertaken to identify, minimize and, when that is not possible, account for these sources of bias. QM, by contrast, is dismissive of such concerns. The composition of the items for the Q sample is, for example, "not necessarily a concern for QM since different Q samples taken from the same concourse tend to produce the same conclusions" (Dziopa and Athern 2011: 40), and in the case of the P sample, "... the key issue is not the number of participants but the representation of different points of view about the theme of the study" (Dennis 1986). There are, from start to end, no less than six phases where the researcher has decisive influence on the research findings (see Sect. 3.4). If these are of no concern, then Q methodology would appear to be incredibly robust with respect to threats that trouble conventional researchers.

5.2 Conclusion

We arrive at a somewhat different conclusion than Wittenborn (1961). We find no reason to reject his argument that Q methodology is useful in clinical settings. But the evidence of the last half century of its use does not support the hopes he had that Q methodology would contribute to a broad study of personality. Use of QM would trouble neo-positivist researchers insofar as the analytical method used imposes artificial constraints on the findings it can generate, it is incapable of producing falsifying findings, it over-reaches in making claims with regard to views in the real world, and it fails to recognise and control for researcher subjectivity. Going to the other end of the ontological spectrum found in the social sciences, constructivist researchers will reject the claims that QM makes with respect to objects in the real world, a reach to which constructivists do not pretend, and, granting that, QM does not transparently account for its dependence on the researcher (see e.g., Fereday 2006; Meyrick 2006; Tobin and Begley 2004). That said, we do find promise in the possibility that proper incorporation of advances made over the last half century in the area of qualitative data analysis will support systematic analysis of justifications for placements of items.

5.3 Directions for future research

What then, is the road ahead? The easy part at this moment in time is to acknowledge that present state-of-the-art science does not permit direct measurement of mental states or subjective representations. The hard part, being attempted in part by those scientists who have achieved crude successes through the use of fMRI devices to objectively study subjective conditions (Owen et al. 2009), will be to produce methods that can. As such, the perhaps humbling decision at this time would be to recognise QM for what it is: an analysis technique that in attempting to combine the strengths of quantitative and qualitative inquiry inherits weaknesses from both (respectively reductionism and subjectivity).

Current uses of QM appear to be an artefact of the marriage of a research strategy that was useful in affording some insight into subjectivity within clinical settings in a manner that supported intra-subject comparisons, an analysis method that was novel at the time of its inception (centroid factor analysis), and the laudable aspiration to produce insights into subjectivity across a set of respondents (inter-rater comparison). In a clinical setting, where a physician has a long-term interest in tracking changes in psychiatric patients' deliberations with regards to extreme placements of sort items, a quasi-normal distribution which focuses attention on extreme cases makes sense. In that context, the absolute placement of an item is perhaps of less interest than is the modes of deliberation used by the patient. Change in these modes of deliberation may indicate patients' progress or regress with regards to a treatment protocol.

While, as this paper has shown, there may be little to validate links between aggregations of sorts and subjective perspectives, the insight that prompted and continues to support the use of Q sorts remains interesting. In clinical contexts the data of interest in intra-subject (by-person) analysis pertains to the deliberative mechanisms employed while sorting. These deliberations, which can be captured via a carefully supported talk-as-you-go interview structure, can be associated both with items and their placement. These deliberations can be then aggregated and analysed through a combination of qualitative and quantitative methods in hypothesis generating research, producing insight into the modes of deliberation espoused by subjects in the interview context. While perhaps not as compelling as a quick and easy direct connect to the 'subjective perspectives' present in a population, these modes of deliberation may be of interest to researchers who are interested, precisely, in understanding better the patterns through which a set of respondents under specific conditions consciously deliberate.

Acknowledgments The authors are thankful to numerous scholars, both inside and outside the QM community, for their helpful comments on earlier drafts of the manuscript. Special thanks go to Hilde Tobi (Wageningen University, NL) and Jennifer B. Barrett (University of South Wales, UK)

Appendix

The following articles were included in the mini-review (Sect. 4):

Baker, M. (2010). Men in a female-majority profession: Q-sort narratives of male clinical psychology trainees. *Clinical Psychology Forum*(214), 25.

Bang, H., & Montgomery, D. (2010). Exploring Korean and American teachers' preferred emotional types. *Roeper Review, 32*(3), 188.

Bayles, B. (2010). Perceptions of childhood obesity on the Texas-Mexico border. *Public Health Nursing, 27*(4), 328.

Berry, R. A. W. (2010). Preservice and early career teachers' attitudes toward inclusion, instructional accommodations, and fairness: Three profiles. *Teacher Educator, 45*(2), 95.

Bischof, B. G. (2010). Negotiating uncertainty: Framing attitudes, prioritizing issues, and finding consensus in the coral reef environment management "crisis". *Ocean and Coastal Management, 53*(10), 614.

Bradley, J., & Miller, A. (2010). Widening participation in higher education: Constructions of "going to university". *Educational Psychology in Practice, 26*(4), 413.

Chen, L., Chen, T. L., & Liu, H. K. J. b. (2010). Perception of young adults on online games: Implications for higher education. *Turkish Online Journal of Educational Technology, 9*(3), 84.

Crasovana, M., & Ungureanu, D. (2010). *Teachers' convictions and beliefs regarding the professional debut*.

Cuppen, E., Breukers, S., Hisschemöller, M., & Bergsma, E. (2010). Q methodology to select participants for a stakeholder dialogue on energy options from biomass in the Netherlands. *Ecological Economics, 69*(3), 591.

Danielson, S., Webler, T., & Tuler, S. P. (2010). Using Q method for the formative evaluation of public participation processes. *Society and Natural Resources, 23*(1), 96.

Duenckmann, F. (2010). The village in the mind: Applying Q-methodology to re-constructing constructions of rurality. *Journal of Rural Studies, 26*(3), 295.

Fairweather, J. (2010). Farmer models of socio-ecologic systems: Application of causal mapping across multiple locations. *Ecological Modelling, 221*(3), 562.

Fard, H. D., & Nikraftar, T. (2010). An inductive model of the main barriers faced by the collaborative government in iran. *European Journal of Economics, Finance and Administrative Sciences*(27), 43.

Farrimond, H., Joffe, H. b., & Stenner, P. (2010). A Q-methodological study of smoking identities. *Psychology and Health, 25*(8), 998.

Giannoulis, C. b., Botetzagias, I., & Skanavis, C. (2010). Newspaper reporters' priorities and beliefs about environmental journalism: An application of Q-methodology. *Science Communication, 32*(4), 466.

Hurley, J. J., & Horn, E. M. (2010). Family and professional priorities for inclusive early childhood settings. *Journal of Early Intervention, 32*(5), 350.

Hurley, J. J., Wehby, J. H., & Feurer, I. D. (2010). The social validity assessment of social competence intervention behavior goals. *Topics in Early Childhood Special Education, 30*(2), 124.

Jedeloo, S., van Staa, A., Latour, J. M., & van Exel, N. J. A. (2010). Preferences for health care and self-management among Dutch adolescents with chronic conditions: A Q-methodological investigation. *International Journal of Nursing Studies, 47*(5), 603.

Kangas, A., Saarinen, N., Saarikoski, H., Leskinen, L. A., Hujala, T., & Tikkanen, J. (2010). Stakeholder perspectives about proper participation for Regional Forest Programmes in Finland. *Forest Policy and Economics, 12*(3), 222.

Kaufmann, L., Carter, C. R., & Buhrmann, C. (2010). Debiasing the supplier selection decision: A taxonomy and conceptualization. *International Journal of Physical Distribution and Logistics Management, 40*(10), 821.

Killam, L. A., Montgomery, P., Luhanga, F. L., Adamic, P., & Carter, L. M. (2010). Views on unsafe nursing students in clinical learning. *International Journal of Nursing Education Scholarship, 7*(1).

Kindsvatter, A., Osborn, C., Bubenzer, D., & Duba, J. (2010). Client perceptions of pretreatment change. *Journal of Counseling and Development, 88*(4), 456.

Klaus, T., Wingreen, S. C., & Blanton, J. E. (2010). Resistant groups in enterprise system implementations: A Q-methodology examination. *Journal of Information Technology, 25*(1), 106.

Lim, C. (2010). Understanding Singaporean preschool teachers' beliefs about literacy development: Four different perspectives. *Teaching and Teacher Education, 26*(2), 224.

Massetti, G. M., & Bracken, S. S. (2010). Classroom academic and social context: Relationships among emergent literacy, behavioural functioning and teacher curriculum goals in kindergarten. *Early Child Development and Care, 180*(3), 375.

Miharia, M., & Woltier, J. (2010). Interlocal government collaboration and perceived transaction costs in Indonesian metropolitan transport planning. *International Development Planning Review, 32*(2), 189.

Militello, M., & Benham, M. K. P. (2010). "Sorting Out" collective leadership: How Q-methodology can be used to evaluate leadership development. *Leadership Quarterly, 21*(4), 632.

Nijnik, M., Nijnik, A., Lundin, L., Staszewski, T., & Postolache, C. (2010). A study of stakeholders' perspectives on multi-functional forests in Europe. *Forests Trees and Livelihoods, 19*(4), 358.

Nikolaus, S., Bode, C., Taal, E., & van de Laar, M. A. F. J. (2010). Four different patterns of fatigue in rheumatoid arthritis patients: Results of a Q-sort study. *Rheumatology, 49*(11), 2199.

Pim, V., Lenders, R. H. J., Devilee, J. L. A., Leuven, R. S. E. W., van der Veeren, R. J. H. M., Wiering, M. A., & Hendriks, A. J. (2010). Stakeholder value orientations in water management. *Society and Natural Resources, 23*(9), 821.

Sawang, S. (2010). Sex appeal in advertising: What consumers think. *Journal of Promotion Management, 16*(1-2), 187.

Schulze, J. L. (2010). Estonia caught between East and West: EU conditionality, Russia's activism and minority integration. *Nationalities Papers, 38*(3), 392.

Takshe, A. A. d., Huby, M., Frantzi, S., & Lovett, J. C. (2010). Dealing with pollution from conflict: Analysis of discourses around the 2006 Lebanon oil spill. *Journal of Environmental Management, 91*(4), 896.

Valenta, A. L., Browning, M. M., Weddle, T. E., Stevenson, G. W. P., Boyd, A. D., & Hynes, D. M. (2010). *Physician perceptions of clinical reminders*. In Proceedings of the 1st ACM International Health Informatics Symposium; 2010 Nov 11 – 12; Washington, DC. Arlington, VA: Association for Computing Machinery

Vermaire, J. H., Hoogstraten, J., Van Loveren, C., Poorterman, J. H. G., & Van Exel, N. J. A. (2010). Attitudes towards oral health among parents of 6-year-old children at risk of developing caries. *Community Dentistry and Oral Epidemiology, 38*(6), 520.

Westwood, D., & Griffiths, M. D. (2010). The role of structural characteristics in video-game play motivation: A Q-methodology study. *Cyberpsychology, Behavior, and Social Networking, 13*(5), 585.

Wolf, A. (2010). Orientations to academic workloads at department level. *Educational Management Administration and Leadership, 38*(2), 262.

Wolsink, M., & Breukers, S. (2010). Contrasting the core beliefs regarding the effective implementation of wind power. An international study of stakeholder perspectives. *Journal of Environmental Planning and Management, 53*(5), 558.

Zagata, L. (2010). How organic farmers view their own practice: Results from the Czech Republic. *Agriculture and Human Values, 27*(3), 290.

References

Ahktar-Danesh, N., Baumann, A., Cordingly, L.: Q-methodology in nursing research: a promising method for the study of subjectivity. West. J. Nurs. Res. **30**(6), 759–773 (2008)

Amin, Z.: Q methodology: a journey into the subjectivity of the mind. Singap. Med. J. **41**, 410–414 (2000)

Barker, J.H.: Q-methodology: an alternative approach to research in nurse education. Nurse Educ. Today **28**, 917–925 (2008)

Block, J.: The Q-sort method in personality assessment and psychiatric research. Am. Lect. Ser. **457**, 56–69 (1961)

Bradley, J., Miller, A.: Widening participation in higher education: constructions of "going to university". Educ. Psychol. Pract. **26**(4), 401–413 (2010)

Brown, S.R.: The forced-free distinction in Q technique. J. Educ. Meas. **8**(4), 283–287 (1971)

Brown, S.R.: Political subjectivity: application of Q methodology in political science. Yale University Press, New Haven, CT (1980)

Brown, S.R.: A primer on Q methodology. Operant Subj. **16**(3,4), 91–138 (1993)

Brown, S.R., Durning, D.W., Selden, S.: Q methodology. In: Miller, G.J., Whicker, M.L. (eds.) Handbook of Research Methods in Public Administration. Marcel Dekker Inc, New York (1999)

Brown, S.R.: Q technique, method, and methodology: comments on Stentor Danielson's article. Field Methods **21**(3), 238–241 (2009)

Cottle, C.E., McKeown, B.F.: The forced free distinction in Q-technique: A note on unused categories in the Q-sort continuum. Operant Subj. **3**, 58–63 (1981)

Cuppen, E., Breukers, S., Hisschemöller, M., Bergsma, E.: Q methodology to select participants for a stakeholder dialogue on energy options from biomass in the Netherlands. Ecol. Econ. **69**, 579–591 (2010)

Danielson, S.: Q method and surveys: three ways to combine Q and R. Field Methods **21**(3), 219–237 (2009)

Dennis, K.E.: Q methodology: relevance and application to nursing research. Adv. Nurs. Sci. **8**, 6–17 (1986)

Dennis, K.E., Goldberg, A.P.: Weight control self-efficacy types and transitions affect weight-loss outcomes in obese women. Addict. Behav, **21**, 103–116 (1996)

Duenckmann, F.: The village in the mind: applying Q-methodology to re-constructing constructions of rurality. J. Rural Stud. **26**, 284–295 (2010)

Dziopa, F., Athern, K.: A systematic review of the applications of Q technique and its methodology. Methodology **7**(2), 39–55 (2011)

Farrimond, H., Joffe, H., Stenner, P.: A Q-methodological study of smoking identities. Psychol. Health **25**, 979–998 (2010)

Fereday, J.: Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int. J. Qual. Method. **5**(1), 80–92 (2006)

Gaito, J.: Forced and free Q sorts. Psychol. Rep. **10**, 251–254 (1962)

Giannoulis, Cb, Botetzagias, I., Skanavis, C.: Newspaper reporters' priorities and beliefs about environmental journalism: an application of q-methodology. Sci. Commun. **32**(4), 425–466 (2010)

Hurley, J., Horn, E.: Family and professional priorities for inclusive early childhood settings. J. Early Interv. **32**(5), 335–350 (2010)

Jedeloo, S., van Staa, A., Latour, J.M., van Exel, N.: Preferences for health care and self-management among dutch adolescents with chronic conditions: a Q-methodological investigation. Int. J. Nurs. Stud. **47**(5), 593–603 (2010)

Kagan, J.: The Three Cultures. Natural Sciences, Social Sciences, and the Humanities in the 21st Century. Cambridge University Press, Cambridge (2009)

Kampen, J.K., Tobi, H.: Social scientific metrology as a mediator between socionomy and sociology: a cri de coeur for the systemizing of social indicators. In: Baird, C.M. (ed.) Social Indicators: Statistics, Trends and Policy Development, pp. 1–26. Nova Science Publishers, Inc., New York (2011)

Killam, L., Montgomery, P., Luhanga, F., Adamic, P., Carter, L.: Views on unsafe nursing students in clinical learning. Int. J. Nurs. Educ. Scholarsh. **7**(1) (2010)

Livson, N.H., Nichols, T.F.: Discrimination and reliability in Q-sort personality descriptions. J. Abnorm. Psychol. **52**(2), 65–159 (1956)

Massetti, G., Bracken, S.: Classroom academic and social context: Relationships among emergent literacy, behavioural functioning and teacher curriculum goals in kindergarten. Early Child Dev. Care **180**, 359–375 (2010)

Meyrick, J.: What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. J. Health Psychol. **11**, 799–808 (2006)

Nikolaus, S., Bode, C., Taal, E., van de Laar, Mb: Four different patterns of fatigue in rheumatoid arthritis patients: results of a q-sort study. Rheumatology **49**, 2191–2199 (2010)

Owen, A.M., Schiff, N.D., Laureys, S.: A new era of coma and consciousness science. Prog. Brain Res. **177**, 399–411 (2009)

Pim, Vb, Lenders, Rb, Devilee, J., Leuven, Rb, van der Veeren, R., Wiering, M., Hendriks, Ab: Stakeholder value orientations in water management. Soc. Nat. Resour. **23**, 805–821 (2010)

Robbins, P., Krueger, R.: Beyond bias? The promise and limits of Q method in human geography. Prof. Geogr. **52**(4), 636–648 (2000)

Shill, W.J.: Unforced and group response to a Q-sort. J. Exp. Educ. **34**(4), 19–20 (1966)

Schmolck, P.: *PQMethod manual*. Retrieved from the internet 11.12.2012 (2012). URL:http://schmolck.org/qmethod/pqmanual.htm

Schwartz, A., Peacock, N., Mcrae, K., Seymour, R., Gilliam, M.: Defining new categories of pregnancy intention in African-American women. Women's Health Issues **20**(6), 371–379 (2010)

Schott, J.R.: Matrix Analysis for Statistics. Wiley, New York (1997)

Sell, D.K., Brown, S.R.: Q methodology as a bridge between qualitative and quantitative research: application to the analysis of attitude change in foreign study program participants. In: Vacca, J.L., Johnson, H.A. (eds.) Qualitative Research in Education, pp. 79–87. Kent University, Kent, OH (1984)
Stephenson, W.: Some observations on Q technique. Psychol. Bull. **49**(5), 98–483 (1952)
Stephenson, W.: The Study of Behavior: Q Technique and its Methodology. Chicago University Press, Chicago (1953)
Tobin, G.A., Begley, C.M.: Methodological rigour within a qualitative framework. J. Adv. Nurs. **48**(4), 388–396 (2004)
Vermaire, J., Hoogstraten, J., Poorterman, J.: Attitudes towards oral health among parents of 6-year-old children at risk of developing caries. Community Dent. Oral Epidemiol. **38**, 507–520 (2010)
Watts, S., Stenner, P.: Doing Q methodology: Theory, method and interpretation. Qualit. Res. Psychol. **2**(1), 67–91 (2005)
Weimer, D.L.: Comment: Q-method and the Isms. J. Policy Anal. Manag. **18**(3), 426–429 (1999)
Wittenborn, J.R.: Contributions and current status of Q methodology. Psychol. Bull. **58**(2), 132–142 (1961)
Wolf, A.: Orientations to academic workloads at department level. Educ. Manag. Adm. Leadersh. **38**, 246–262 (2010)