Food Quality and Preference 19 (2008) 511–518

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Comparing two image research instruments: The Q-sort method versus the Likert attitude questionnaire

Peter M. ten Klooster, Martijn Visser, Menno D.T. de Jong *

University of Twente, Institute for Behavioral Research, Faculty of Behavioral Sciences, P.O. Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 3 April 2007
Received in revised form 21 February 2008
Accepted 21 February 2008
Available online 29 February 2008

Keywords:
Image research
Q-sort method
Product image
Applied research
Methodology

ABSTRACT

Despite the attention for corporate, brand and product images, only few studies focus on methodological comparisons of image research methods. This article presents a comparison of two current instruments: the Q-sort method and a Likert attitude questionnaire. The study applies both methods to measure the image of beef, using the same assertions in similar samples of consumers. The two methods produce consistent results, but differ in the possibilities of data analysis and interpretation. An advantage of the Q-sort method is that it offers straightforward insights in the underlying structure of image within audience segments. On the other hand, the Q-sort method does not give overall indications of an image, and limits occur for analyzing the relationships between image and other variables.

1. Introduction

With the rise of image as an object of communication and marketing research, many different instruments have been developed to measure corporate, brand and product image among stakeholder groups. Most of these instruments stem from psychological research traditions. The diversity of research approaches available reflects the ambiguity of the image concept itself, which has been defined in many different ways (Christensen & Askegaard, 2001; DeFleur & Westie, 1963; Poiesz, 1989). There is no universally accepted image definition, and the more recent introduction of the equally intangible and strongly related concept of reputation only seems to add to the confusion.

Poiesz (1989) categorizes the various image definitions by placing them on a consumer elaboration continuum. He thus distinguished three views on image, which can be easily connected with possible research approaches. In the case of *high elaboration*, an image represents a complex network of meanings stored in memory. This implies that the measurement of an image must aim at revealing and exploring associations people have with the image object, which calls for qualitative and open methods like the Kelly repertory grid or laddering. In the case of *medium elaboration*, an image is a theoretical and operational equivalent of an attitude: it is the overall evaluation of an artifact based on salient beliefs held by consumers. Consequently, the measurement of image resembles attitude scales, involving the evaluation and weighting of beliefs. In the case of *low elaboration*, an image is merely a holistic impression of the relative position of an object among its perceived competitors. Measurement must then focus on differences and similarities between image objects, using multi-dimensional scaling.

This classification offers a fruitful basis for comparing image research instruments. Within each particular image definition, researchers have several methods at their disposal to measure image. A relevant research question is whether or not competing research methods are equally suitable for measuring image and whether they produce similar results. So far, only few researchers respond to this challenge (Van Riel, Stroeker, & Maathuis, 1998).

This article focuses on the measurement of image in the case of medium elaboration. Image is supposed to be the equivalent of an attitude: image is an overall evaluation that is a consequence of a series of beliefs and importance weightings. Two competing methods that are potentially useful in this context are Likert attitude questionnaires and the Q-sort method. Both methods can be used to measure attitudes and both present the respondents with a predetermined set of items (beliefs), which must be judged on an *x*-point scale, so that respondents must make trade-offs between scale positions.

Our study investigates the suitability of the two methods for product image research. First, the rationale of the two methods is discussed. After that, the possibilities of the methods are analyzed

* Corresponding author. Tel.: +31 53 489 3313; fax: + 31 53 489 4259.
E-mail address: m.d.t.dejong@utwente.nl (M.D.T. de Jong).

0950-3293/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodqual.2008.02.007

in an empirical study into the image of beef. Finally, the article addresses the strengths and weaknesses of both methods for image research.

2. The Q-sort method

The Q-sort method is rooted in Q-methodology, an inverted technique of factor analysis. Developed by Stephenson (1953) as an alternative measurement technique to existing scales and tests in psychology, the method can be used in any situation in which subjectivity is at issue, including attitude measurement (Stephenson, 1965, 1968). Mainly due to the work of Brown (1980) and McKeown and Thomas (1988), who further developed the principles and procedures of Q-sorting, the Q-sort method has more recently found its way to a wide variety of research areas.

Today, Q-methodology has become a popular method of investigating attitudes (Cross, 2005). Particularly within medical and nursing research, the Q-sort method is often used to measure patients' or professionals' attitudes toward health-related issues, such as chronic pain (Risdon, Eccleston, Crombez, & McCracken, 2003), the adoption of information technology (Valenta & Wigger, 1997), and patients' needs and concerns (Staley-Gane, Flynn, Neitzel, Cronister, & Hagerman, 1996). The method has also become quite common as a research approach for public opinion (Webler, Tuler, & Krueger, 2001), communication (Carlson & Trichtinger, 2001), policy analysis (Durning, 1999), landscape planning and rural research (Swaffield & Fairweather, 1996; Previte, Pini, & Haslam-McKenzie, 2007), environmental issues (Barry & Proops, 1999), and education (Lecouteur & Delfabbro, 2001).

Surprisingly, the academic business and marketing literature has so far paid little attention to the Q-sort method. The potential value of the Q-sort method in public relation, advertising, and image research was already demonstrated by Stephenson (1963, 1969, 1979) and confirmed by Schlinger (1969) and current marketing handbooks suggest Q-sorting as a suitable technique for corporate image research (e.g., Smith & Albaum, 2004). However, its use in academic image studies is limited to a small number of studies into the relationship between self-image and brand or product image (e.g., Hamm & Cundiff, 1969; Martin & Reynolds, 1976; Sommers, 1963). Despite the promising results of these studies, the method has been rarely used in this context since then.

2.1. Design

The Q-sort method typically involves the rank-ordering of a set of statements in a near-normal distribution, ranging from agree to disagree (see Fig. 1). Through the years, however, researchers apply the method in a much broader sense. Any set of items that can be meaningfully evaluated may be used (Schlinger, 1969). Respondents may, for example, be asked to sort photographs (Fairweather & Swaffield, 2002), product names (Fyock et al., 2001) or colored papers (Gelineau, 1981). The corresponding distribution scales may range from like to dislike, from best label to worst label, or between opposite mood adjectives.

The Q-sort method is a forced-choice research approach: all items must be ranked and each position can only be used once. An important difference between the Q-sort method and more conventional attitude research approaches involves the data analysis: the data matrix is inverted, so that the respondents are the variables and the items are the cases. Respondents are correlated, instead of items.

When designing a Q-sort study, researchers must decide on the number of items and the shape of the (forced) distribution. The number of items corresponds with the complexity of the research topic (Amin, 2000). All possible aspects of the research topic must

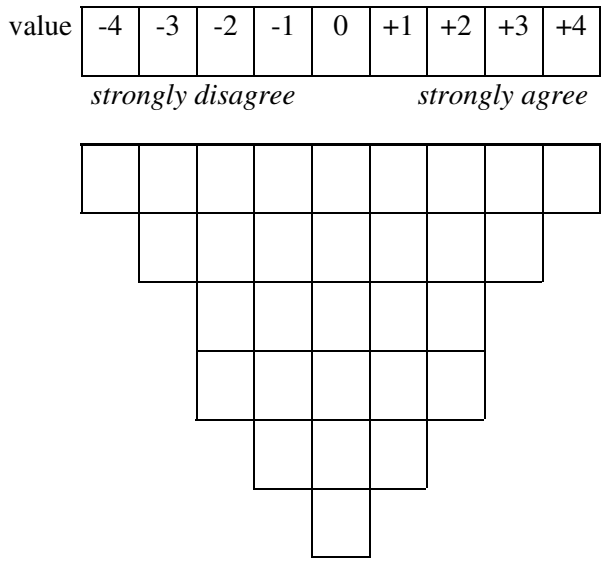

Fig. 1. Q-sort distribution.

be represented. In practice, samples of 60 items or more are seldom necessary (Dennis, 1988). Although considerable research has examined the properties of different distribution shapes (e.g., Brown, 1971), no specific guidelines in this respect are available. Both the range and the distribution shape are usually arbitrarily designed to accommodate the number of items used in the study (Addams, 2000). Q-samples typically consist of 20–50 statements, which must be ranked using 7–11 piles (Mrtek, Tafesse, & Wigger, 1996). The number of items per pile varies, but usually assumes a near-normal distribution, with one or two items in the extremes, gradually increasing to four to nine items in the middle pile(s).

2.2. Procedure

The Q-sort procedure consists of four steps. The first step is the collection of relevant ideas, beliefs and opinions concerning the research object. Such a collection is referred to as the *concourse*, and can be based on various sources, such as interviews, content analysis or previous research. The second step involves the selection and formulation of a set of meaningful statements, which results in the *Q-sample*. If the items are statements about a product or brand, the final selection ideally consists of an equal number of positive and negative statements (Schlinger, 1969). Up to here, the procedure resembles the development of attitude questionnaire items (Stainton Rogers, 1995). The resulting statements are randomly numbered and printed on separate cards. In the third step, respondents have to express their views on the research topic by placing all cards in the pre-structured Q-sort distribution. A completed sorting task is called a *Q-sort*.

The fourth step is the *data analysis*. A correlation matrix is made of all Q-sorts, indicating the degree of correspondence between respondents. The correlation matrix is submitted to a by-person factor analysis to explore attitudinal groupings. Factor scores are computed for each of the items in the resulting clusters of respondents, which leads to one representative Q-sort per group. Also, confounding respondents (loading significantly on more than one factor) and non-significant respondents (not loading on any of the factors) are identified. The final task is to interpret and explain similarities and differences among the factors. Two established computer software packages are available that standardize and automate data analysis and thus enhance the method's feasibility

for all researchers: PQMethod (Schmolck & Atkinson, 2002), which is freely available on the web, and PCQ for Windows (Stricklin & Almeida, 2004).

2.3. Characteristics

Several researchers commend the Q-sort method for its potential to combine the strengths of qualitative and quantitative research (Amin, 2000; Brown, 1996; Dennis, 1988; Valenta & Wigger, 1997). The method is specifically designed to identify groups of respondents whose overall attitudes are similar, and to closely examine the differences with respondents who have other views (Brown, 1996; Lecouteur & Delfabbro, 2001; Schlinger, 1969; Van Riel et al., 1998). Although segmentation techniques such as cluster analysis and discriminant analysis can be used for the same purposes with survey data, these analyses tend to be more complex and their interpretation is less straightforward.

A more practical advantage of the Q-sort method is that it requires relatively small samples of respondents (Dennis, 1988). Since Q-sort studies are designed to sample from a universe of perspectives, rather than from a population of people (Anderson, Avery, Pederson, Smith, & Sullivan, 1997), representativeness does not depend on large samples of respondents. As such, the diversity of the respondents is more important than the number of respondents. All the method requires is that there are enough respondents to identify these perspectives as factors for purposes of comparison (Brown, 1980). Most Q-sort studies result in fewer than seven factors, and often not more than two or three. Since only four or five respondents are required to produce stable factors for purposes of identification, a Q-sort study usually requires only 20–50 respondents (Brown, 1986; Dennis, 1988; McKeown & Thomas, 1988; Stainton Rogers, 1995).

Finally, performing a Q-sort is often a pleasant task for respondents, who seem to enjoy (re)arranging the attitude descriptors until they are fully satisfied (Prasad, 2001). Kitzinger (1986), Van Riel et al. (1998), Wood, Griffiths, Derevensky, and Gupta (2002), however, also emphasize that the method may be time-consuming and demanding to respondents.

3. The Likert attitude questionnaire

Attitude rating scales are the most widely used research approach within marketing and opinion research (McDonald & Vangelder, 1998; Zikmund, 1997). One of the best known rating scales is the typical five-point agreement scale developed by Likert (1932). Although its principles are very different from Q-sorting and both procedures were developed to serve different ends (Stephenson, 1953; McKeown, 2001), Likert scaling is a well-accepted technique for attitude measurement. Mainly due to its simplicity and reliability, Likert attitude scales have gained in popularity over similar scales, such as the Thurstone scale.

3.1. Design

In its typical form, the Likert attitude questionnaire presents respondents with a series of statements, to which they must react on a five-point scale ranging from *strongly disagree* to *strongly agree*. Likert (1967) assumed that every statement is equally important to the overall attitude. The original scale did not consider the weight respondents attached to individual items. Later, Fishbein and Ajzen (1975) argued for an attitude research instrument that would also include weight measures of the importance or desirability of each statement. The overall attitude can then be computed by multiplying the weights and the 5-point scores of the various items.

3.2. Procedure

The first step in developing a Likert attitude questionnaire is the construction of an extensive item pool. In preliminary research, a large number of opinions and statements are collected. Statements are selected and edited, in accordance with criteria for "good" research statements, like the ones used by Likert (1967). The next step is scale construction, which also includes the elimination of items that do not clearly represent the attitude construct. This can be done by pre-testing the statements among potential respondents. Weak statements are identified and excluded from the scale using item analysis. The remaining statements constitute the final scale (Edwards & Kenney, 1967; Fishbein & Ajzen, 1975). Half of the statements must reflect a positive attitude toward the research object, the other half a negative attitude (Hogg & Vaughan, 1995; Likert, 1967). The scale can be administered in several ways. The most common way is the mailed questionnaire. When analyzing data and reporting results, factor analysis is often used to identify underlying dimensions of the research object. In general, the analyses result in an average attitude score toward the object and overviews of the average scores on all items or dimensions. Calculating an overall response per group of participants based on demographic similarities is also possible.

3.3. Characteristics

The literature documents well the advantages and disadvantages of written questionnaires in general and Likert attitude questionnaires (e.g., Chadwick, Bahr, & Albrecht, 1984; Spector, 1992). Written questionnaires are cost-effective compared to other types of research, in particular for studies with large sample sizes and geographic dispersion. Besides, written questionnaires are free of bias caused by the presence of a researcher. Likert attitude scales in particular are relatively easy to construct and analyze, and are generally expected to yield reliable results (Barclay & Weaver, 1962; Edwards & Kenney, 1967; Roberts, Laughlin, & Wedell, 1999; Seiler & Hough, 1970). The reliability of Likert scales can be easily assessed by split-half, equivalent-form, or test–retest correlations. Another advantage of Likert scales is that they are familiar to most respondents, which may reduce errors. Since the validity of traditional survey techniques such as Likert attitude questionnaires strongly depends on the representativeness of the respondent sample, large sample sizes are usually required to obtain useful results. A major disadvantage of written questionnaires, however, is the generally low response rate, which may correspond to non-response bias. Another disadvantage is the impossibility to avoid confounding errors in the answering and response sets.

4. Applying the two methods: investigating the image of beef

4.1. Image object and target population

To explore the characteristics of both instruments in an image research setting, the Q-sort method and a Likert attitude questionnaire were used in a study into the product image of beef. The study focuses on beef for two reasons. Beef is a well-known product category and recent food scares concerned the safety of beef products, such as the BSE-crisis in 2000, and the 2001 outbreak of foot and mouth disease among cattle in The Netherlands. Moreover, the ambivalent thoughts and feelings of people towards food in general, and beef in particular (De Jonge et al., 2004; Smith, Young, & Gibson, 1999; Verbeke & Viaene, 1999), make beef an appropriate attitude object. The two instruments were used in exactly the same time frame, and they were used in the same target population: students at a Dutch university.

4.2. Item collection and selection

The same list of statements about beef were used for both methods. The statements originated from 10 semi-structured interviews among members of the target population. Together, these interviews led to 145 assertions about beef. The statements were not limited to product attributes, since a product image interacts with and may be affected by other image aspects. The 30 statements selected covered the entire range of views that resulted from the preliminary interviews. Half of the statements represented a negative attitude toward beef, the other half a positive attitude. The 30 statements comprised topics like the price of beef, product quality, safety, the welfare of cows, and the economical and environmental contribution of the beef sector (see Appendix).

4.3. Design of the Q-sort study

The Q-sort data collection involved one-to-one research sessions. The materials were 30 small cards with statements about beef, and a large board with an empty Q-sort distribution (see Fig. 1). The cards were presented to the respondents in random order. At the start of a session, the facilitator explained the sorting task to the respondents. The respondents had to read the cards and divide them into three piles: one pile for the statements they agreed with, one for the ones they disagreed with, and one they felt uncertain or neutral about. After that, the respondents placed the cards from the agree pile on the positions on the board, followed by, respectively, the disagree and the neutral pile. They were free to move the cards, until all positions were filled and they were fully satisfied with the resulting distribution.

51 Respondents were recruited in the canteens of several faculty buildings of the university. They formed an adequate representation of the various educational programs at the university. 31 Respondents (61%) were male. The mean age was 21 years. Four respondents (8%) indicated to be vegetarian.

4.4. Design of the Likert attitude questionnaire study

The Likert questionnaire includes the same 30 statements appearing in the Q-sort method. To avoid order effects, two questionnaire versions were used, both with a different random ordering of statements. Respondents had to rate the statements on a five-point scale, ranging from strongly disagree to strongly agree. In the final part of the questionnaire, respondents also had to weigh the importance of various aspects of beef, again on a five-point scale. To limit the number of items in the questionnaire, these weight questions did not apply to each separate item, but to nine clusters of items.

The questionnaire was distributed among 1000 randomly selected students, who again formed an adequate representation of the various educational programs. In total, 160 completed questionnaires were returned: 88 of questionnaire version 1, 72 of version 2. One 112 Respondents (70%) were male. The mean age was 21.5 years. Eight respondents (5%) were vegetarian.

5. Results Q-sort method

The Q-sort data were analyzed using PCQ for Windows (academic edition, version 1.41). The data were analyzed using centroid factor analysis with varimax rotation. Of the various results the program produces, the factor analysis—including representative sorts per factor and differentiating items—and consensus statements are the most important. Due to the relatively large sample of respondents used in this study, the factor analysis revealed up to nine factors with eigenvalues greater than 1, accounting for 70% of the variance in the Q-sorts. A three-factor solution was chosen, which made both theoretical and statistical sense. The average Q-sort factor arrays for each factor are reported in the Appendix. This three-factor solution accounted for 49% of the variance. Of the 51 respondents, 14 respondents could not be placed in one of the factors: four because they were confounded (i.e., significantly loading on more than one factor), and ten because they were non-significant (i.e., not loading significantly on any of the factors). For each factor group of respondents a representative sort was made. Particularly statements in the extreme positions of a sort were of interest, as these strong agreements and disagreements highlighted the unique viewpoints of the respondent group concerned.

5.1. Factor A: overall enthusiasts

The first respondent group were the overall enthusiasts about beef (eigenvalue 9.6, accounting for 19% of the variance). In this group, beef appeared to have an outright positive image. Especially the intrinsic properties of beef—including its taste, ease of preparation, and variation—received favorable judgments. Besides, the respondents in this group shared the opinion that cows in the meadows are an essential part of the Dutch scenery. Other aspects of the image of beef were judged moderately positively (e.g., safety) or neutrally (e.g., price). In general, the respondents in this group disagreed with negatively worded statements, and agreed with the positive ones. The group consisted of ten male and four female students, who were all regular beef consumers (3–4 times a week).

5.2. Factor B: idealistic critics

The second group were idealistic critics (eigenvalue 9.8, also accounting for 19% of the variance). The respondents in this group, too, were very positive about the intrinsic properties of beef, but were also concerned about the welfare of cows and the mass production in the beef sector: according to them, the keeping, transportation and slaughtering of cattle was not as animal-friendly as it should be. The group consisted of sixteen students: nine female and seven male. As could be expected, all four vegetarian participants formed part of this group, and the average beef consumption among the non-vegetarians was lower than that in the first respondent group (2–3 times a week).

5.3. Factor C: pragmatic non-enthusiasts

The third group were the pragmatic non-enthusiasts (eigenvalue 5.6, accounting for 11% of the variance). The respondents in this group were essentially negative about beef, and as such almost the opposite of the respondents in group A. They felt negative about the intrinsic properties of beef, in particular its taste and the difficulty of preparing beef. They did not think of cows as an essential part of the Dutch scenery. And they did not think of beef as something healthy, but rather as a potential health risk. In general, these respondents had a pragmatic view on beef; they simply did not like it, and were not interested in matters such as the welfare of cows or the agribusiness. The respondent group consisted of five female and two male students, with an average beef consumption similar to the non-vegetarian respondents in group B (2–3 times a week).

5.4. Consensus statements

Apart from the factor analysis, PCQ also identifies consensus statements (i.e., statements that do not distinguish between any of the three factors). The three-factor solution resulted in only one consensus statement. In each of the factor arrays, the state-

ment "beef is too expensive" (statement 1) was placed in the same pile. Apparently, the price of beef was generally considered quite reasonable or was simply not a major concern to any of the groups, possibly for different reasons within each group.

6. Results Likert attitude questionnaire

All statements were coded on a scale from −2 (*strongly disagree*) to +2 (*strongly agree*). The respondents' importance weightings of the nine clusters of variables were also coded on a scale from −2 (*very unimportant*) to +2 (*very important*). By multiplying each statement with the importance weight of the cluster of items it belongs to, weighted average scores were computed, which could range from −4 (*extremely negative*) to 4 (*extremely positive*).

6.1. Item reduction

Item-analysis identified six weighted statements with negative or low item-total correlation (2, 9, 10, 24, 25, and 29). Exploratory factor analysis (principal components analysis with varimax rotation) of the remaining statements yielded seven factors with eigenvalues >1.0. A scree test resulted in a three-factor solution that accounted for 41% of the variance of the data. Twenty-one statements loaded >.35 on one of these factors. Factor 1 consists of ten statements about animal welfare, environment, and agribusiness (18–23, 26–28, and 30; Cronbach's alpha = .85); factor 2 consists of seven statements about the price and safety of beef (1 and 11–16, Cronbach's alpha = .68); factor 3 contains four statements about the physical properties of beef (3 and 6–8, Cronbach's alpha = .73).

6.2. Attitude scores

The summated weighted average score over the 21 remaining statements was −0.06, suggesting a neutral to slightly negative overall image. Nine statements contributed positively to the image of beef, 12 statements negatively. The dimension concerning animal-wellbeing, environment, and agribusiness (factor 1) was judged most negatively ($M = -0.54$), with the environmental threat due to the manure surplus as the absolute negative outlier (statement 22, $M = -1.04$). All other statements within this dimension also received negative judgments, the only exception being that farmers were believed to treat their cattle well (statement 18, $M = 0.26$). Price and safety of beef (factor 2) were mostly judged neutral ($M = 0.03$), although the respondents had some concerns about the risks of beef consumption due to cattle diseases (statement 16, $M = -0.69$). The physical properties of beef (factor 3) were rather highly appreciated ($M = 0.99$). Respondents were especially positive about the taste of beef (statement 7, $M = 1.51$) and its ease of preparation (statement 3, $M = 1.01$).

The general conclusion is that the properties of beef as a product received positive judgments, but that these positive sides were largely counteracted by sector issues such as animal welfare, environment, and agribusiness.

6.3. Demographic differences

Differences between male and female respondents were analyzed using independent *t*-tests. The summated average attitude score differed between the genders. Male respondents were considerably more positive about beef than female respondents ($M = 0.06$ vs. −0.32, $t(158) = 3.69$, $p = .001$). Male respondents judged more favorably about animal welfare, environment, and agribusiness ($M = -0.37$ vs. −0.94, $t(158) = 3.25$, $p = .002$) and the price and safety of beef ($M = 0.11$ vs. −0.15, $t(158) = 2.36$, $p = .020$). The physical properties of beef were judged similarly between males and females ($M = 1.03$ vs. 0.92, $t(158)] = 0.73$, $p = .466$).

Since the attitudes of vegetarians were not normally distributed, differences between vegetarians and non-vegetarians were analyzed using Mann–Whitney *U* tests. The vegetarians were considerably less positive about beef in general than non-vegetarians ($Mdn = -1.50$ vs. 0.07, $U = 14.50$, $p < .001$). Vegetarians judged significantly differently about 15 of the 21 statements (p's < .05). Vegetarians were more negative about animal welfare, environment, and agribusiness ($Mdn = -2.85$ vs. −0.25, $U = 15.50$, $p < .001$) and the physical properties of beef ($Mdn = 0.00$ vs. 1.00, $U = 382$, $p < .005$). Vegetarians and non-vegetarians, however, held similar opinions on issues regarding the price and safety of beef ($Mdn = -0.21$ vs. 0.00, $U = 213$, $p = .076$).

6.4. Relationship between image and beef consumption

A final analysis focuses on the factors associated with the actual consumption of beef. For this purpose, weekly consumption of beef (never, sometimes, regularly, or often) was dichotomized into low (never or sometimes, $n = 80$) and high consumption (regularly or often, $n = 78$). A forced-entry logistic regression analysis was performed, with the three image dimensions as potential predictors of high consumption. The only factor that was significantly associated with high consumption, was the physical properties of beef (Wald[1] = 11.25, odds ratio: 2.16, $p < .001$). A more positive opinion on the physical proprieties of beef strongly increased the likelihood that a respondent was a regular consumer. However, the total predictive power of image appeared to be limited (Nagelkerke $R^2 = .16$).

7. Comparison of the two methods

The use of a Likert attitude questionnaire and the Q-sort method to investigate the same image object with exactly the same items and highly similar respondent groups gave the opportunity to compare and illustrate the use of the two methods for image research.

7.1. Feasibility

Likert scales are high in feasibility, both for the researcher and for respondents. Respondents are usually familiar with the response categories and do not need any supervision or instruction to fill out a questionnaire. An important problem of written questionnaires is the resistance of people to participate, which often results in high non-response rates. In this study, too, the non-response was high (84%), which may have affected the results of the questionnaire. This implies that many questionnaires must be sent out, to actually collect the number of questionnaires required to obtain reliable data and explore relationships as the ones mentioned above. The Q-sort method is often assumed to require more efforts of the researcher and the respondents. However, in the current study, the efforts of respondents appeared to be comparable to those of filling out a questionnaire: it took the respondents an average of 10–15 min to sort the 30 statements. The respondents did not experience any problem in performing the sorting task and many of them indicated that they enjoyed the method, which was, in their view, a welcome change to the usual research practices. The extra efforts of the researcher concern the data collection sessions: since respondents are unfamiliar with the procedure and the task involves higher elaboration, supervision of the sorting sessions is recommended. On the other hand, the method requires a relatively small sample size and there may be contexts in which it is considerably easier to recruit respondents for a brief face-to-face session than to find them willing to fill out a questionnaire.

7.2. Similarity of results

To investigate the degree of similarity between the questionnaire and Q-sort results, Pearson correlation coefficients were computed between the mean scores the thirty items were given under both methods and between the mean Likert questionnaire item scores and the factor scores resulting from the Q-sort analysis. For the Likert questionnaire items, the unweighted scores were included in the analyses. A very high correlation was found between the mean Likert item scores and overall Q-sort item scores ($r = .93$, $p < .001$), indicating that the mode of administration (filling out the questionnaire versus performing the sorting task) did not substantially affect the way respondents rated the 30 items overall.

When the mean Likert item scores were correlated with the actual factor scores from the Q-sort analysis (see Appendix), a different picture emerged. Mean Likert item scores were not correlated with factor A ($r = .005$, $p = .979$) or factor C ($r = .029$, $p = .881$) and negatively correlated with the factor B ($r = -.506$, $p = .029$), providing support for the assumption that the Q-sort method reveals functional information that is not captured by standard Likert scale analysis (Brown, 2002).

7.3. Usefulness of outcomes

Despite their similarity in item ranking, the two methods offered entirely different possibilities of analyzing and implementing the results. The Q-sort method identified three groups of people with different attitudes toward beef. This may be helpful, since products, brands, and organizations are usually not expected to have one overall image, but multiple images in different stakeholder groups. A distinction of audience segments based on their own perspectives on the image object may be an important step toward targeted interventions. An important limitation of the Q-sort results was that they did not generate an overall indication of the image of beef. Factor score arrays and the representative sorts are not exact figures, but more or less ideal or theoretical constructions. The scores of the three factor groups cannot be summarized into one overall image of beef. Besides, Q-sort data do not give substantial insights in demographic differences within the research population. The composition of factor groups provides some global indications, but demographic conclusions are always tentative.

The outcomes of the attitude questionnaires, on the other hand, did not provide the same level of detail as the Q-sort data, but instead focused on an overall impression of the image and the dimensions within this image. Moreover, the questionnaire produced clear ratings of individual image attributes, making the outcomes suitable for longitudinal or benchmark studies. The relationship between image and demographics can be easily explored, as well as the relationship between image and behavioral intentions.

8. General discussion

The comparison of the Likert attitude scale and the Q-sort method this article describes leads to mixed results. The two image research methods appear to produce both similar and different outcomes. At the level of specific ratings per statement, the results indicate a high degree of similarity between both methods. This finding provides concurrent validity of the methods: the tasks of ordering statements in a normal distribution and judging statements on a 5-point scale do not lead to substantially different rankings of the statements. Despite this similarity in results, however, the conclusions that may be drawn on the basis of the two methods differ considerably. A choice for either of the two methods may depend on a large number of considerations. The current study into the image of beef sheds light on one of the crucial considerations for the selection of a research method: the kinds of questions that can be answered with it.

The use of Likert attitude questionnaires results in an overall impression of the image, expressed in one composite average score, and to specific ratings of individual items or image dimensions. Compared with the Q-sort method, attitude questionnaires have several clear advantages. First, they offer the possibility to compare an image with that of competing brands, products or organizations, or to track developments across time. Second, questionnaire data may pinpoint differences in image between audience segments, distinguished on the basis of, for instance, demographic, geographic or lifestyle characteristics. Third, they offer the possibility of investigating the relationship between image dimensions and consumer behavior. By doing so, it is possible to identify those image dimensions that are most relevant for the viability of a brand, product or organization. And fourth, the data collected allow analyses with a wide range of statistical techniques, which may, among other things, shed light on the reliability and construct validity of the results and the items used.

The Q-sort method, on the other hand, provides a unique opportunity to distinguish salient groupings within the population with similarly structured attitudes toward an image object. So, instead of segmentation on the basis of the target audience's background characteristics, the Q-sort method results in segmentation on the basis of functional, content-specific criteria. Apart from that, the method offers the possibility to conduct an in-depth analysis of the structure in the opinions and attitudes of respondents, and as such may combine unique strengths of qualitative research (the exploratory nature and a profound analysis of the data representing the respondents' perspective) and quantitative research (formal and well-defined data collection that can be analyzed statistically). A practical advantage of the Q-sort method is that it appears to be an attractive way of data collection in the eyes of the respondents. In a time when people's willingness to participate in questionnaire-based research is definitely decreasing, the Q-sort method may be a welcome alternative data collection approach. To conclude, an additional possibility of the Q-sort method not used in the study reported here is the combination with an in-depth interview afterwards focusing on the reasons why the respondents put certain cards in the extreme positions. In other studies conducted with the Q-sort method, such a combination also proved to be highly useful: the interviews may serve as additional qualitative information that helps uncover the thoughts respondents had about the image object, and thus provide explanations for differences in Q-sorts between factor groups. The results of the sorting task form an excellent opportunity to discuss specific image aspects with respondents.

When conducting a Q-sort study, researchers have no standard means of testing the reliability of the specific instrument used. Reliability of a Q-sort instrument can only be determined by applying the same instrument at several occasions, or by simulation research analyzing the results of various sub samples. This study does not address the reliability of the Q-sort method. But several researchers have investigated the consistent ordering of statements on different occasions and they all reported acceptable to high correlations (Brown, 1980; Fairweather, 1981; Gelineau, 1981; Pease, Boger, Melby, Pfaff, & Wolins, 1989; Steller & Meurer, 1974; Treadwell & Harrison, 1994). A recent study on the test–retest reliability of the factors found in a Q-sort instrument for attitude measurement led to the conclusion that the factors found were reliable over time (Prasad, 2001). The results of all these studies provide evidence for the reliability of Q-sort constructs over time and across subjects.

The Q-sort methodology is especially suitable for exploratory image research. The method provides a cost- and time-effective means for detailed exploratory or targeted image research. For longitudinal or benchmark research, large group surveys methods like the Likert attitude questionnaire are probably still the best option. The Q-sort method, however, could provide valuable input for the construction of such an instrument. As suggested by Anderson et al. (1997), Likert attitude scales could be structured around the factors revealed by a Q-sort study, so that demographic characteristics can be examined in relation to content-specific groups.

Appendix A

Statements used

	Q-sort factor scores		
	A	B	C
1. Beef is too expensive*	−1	−1	−1
2. Beef is cheaper than other types of meat	−1	0	3
3. Beef is difficult to prepare*	−3	−3	3
4. Beef goes with almost any dish	0	1	−2
5. The quality of beef from the supermarket is insufficient*	−3	−2	1
6. Hallmarked beef from the butcher is of high quality	3	1	−1
7. Beef is not tasty*	−4	−4	4
8. The supply of beef is varied	3	0	0
9. Beef contains less fat than other types of meat	0	0	−2
10. Beef contains essential nutrients to the human body	2	−1	−3
11. The current hallmarks guarantee the safety of beef	2	0	0
12. Beef contains hormones which are hazardous to my health*	−2	−1	2
13. Dutch beef is safer than imported beef	1	−1	1
14. The government does not sufficiently oversee the observing of safety regulations*	−2	−1	1
15. The hygiene in butchers' shops is good	2	0	0
16. Diseases like the foot and mouth disease involve risks for the consumption of beef*	0	2	−1
17. The culling of cattle during the foot and mouth disease crisis was needless*	−2	1	−3
18. Farmers treat their cattle well	1	−2	−2
19. The transportation of cattle is not animal-friendly*	1	3	2
20. Abattoirs work as animal-friendly as possible	1	−3	0
21. The welfare of cattle used to be better*	−1	1	1
22. The manure surplus is a threat to the environment*	1	2	−2
23. The live stock in the Netherlands needs to reduce to protect the environment*	−2	1	2
24. The beef sector is vital to the Dutch economy	0	−2	0
25. Cows belong with the Dutch scenery	4	2	−4
26. Outbursts of BSE en FMD are consequences of the mass production of beef*	0	2	−1
27. The government sufficiently stimulates biological production of beef	0	−2	2
28. The mass production of beef is encouraged by the consumers*	2	4	0
29. Farmers themselves would like to make the change to biological cattle breeding	−1	0	1
30. In the Netherlands, too many cows are housed in mass production cow houses*	−1	3	−1

* Negative statements.

References

Addams, H. (2000). Q methodology. In H. Addams & J. Proops (Eds.), *Social discourse and environmental policy: An application of Q methodology* (pp. 14–40). Cheltenham, UK: Edward Elgar.

Amin, Z. (2000). Q methodology – A journey into the subjectivity of human mind. *Singapore Medical Journal, 41*(8), 410–414.

Anderson, C., Avery, P. G., Pederson, P. V., Smith, E. S., & Sullivan, J. L. (1997). Divergent perspectives on citizenship education: A Q-method study and survey of social studies teachers. *American Educational Research Journal, 34*(2), 333–364.

Barclay, J. E., & Weaver, H. B. (1962). Comparative reliabilities and ease of construction of Thurstone and Likert attitude scales. *Journal of Social Psychology, 58*, 109–120.

Barry, J., & Proops, J. (1999). Seeking sustainability discourses with Q methodology. *Ecological Economics, 28*(3), 337–345.

Brown, S. R. (1971). The forced-free distinction in Q technique. *Journal of Educational Measurement, 8*(4), 283–287.

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

Brown, S. R. (1986). Q technique and method: Principles and procedures. In W. D. Berry & M. S. Lewis-Beck (Eds.), *New tools for social scientists: Advances and applications in research methods* (pp. 57–76). Beverly Hills, CA: Sage.

Brown, S. R. (1996). Q methodology and qualitative research. *Qualitative Health Research, 6*(4), 561–567.

Brown, S. R. (2002). Structural and functional information. *Policy Sciences, 35*(3), 285–304.

Carlson, J. M., & Trichtinger, R. (2001). Perspectives on entertainment television's portrayal of a racial incident: An intensive analysis. *Communication Review, 4*(2), 253–278.

Chadwick, B. A., Bahr, H. M., & Albrecht, S. L. (1984). *Social science research methods*. Englewood Cliffs, NJ: Prentice Hall.

Christensen, L. T., & Askegaard, S. (2001). Corporate identity and corporate image revisited: A semiotic perspective. *European Journal of Marketing, 35*(3/4), 292–315.

Cross, R. M. (2005). Exploring attitudes: The case for Q methodology. *Health Education Research, 20*(2), 206–213.

DeFleur, M. L., & Westie, F. R. (1963). Attitude as a scientific concept. *Social Forces, 42*(1), 17–31.

De Jonge, J., Frewer, L., Van Trijp, H., Renes, R. J., De Wit, W., & Joke, T. (2004). Monitoring consumer confidence in food safety: An exploratory study. *British Food Journal, 106*(10/11), 837–849.

Dennis, K. E. (1988). Q-methodology: New perspectives on estimating reliability and validity. In O. L. Strickland & C. F. Waltz (Eds.), *Measurement of nursing outcomes* (pp. 409–419). New York: Springer.

Durning, D. (1999). The transition from traditional to postpositivist policy analysis: A role for Q-methodology. *Journal of Policy Analysis and Management, 18*(3), 389–410.

Edwards, A. L., & Kenney, K. C. (1967). A comparison of the Thurstone and Likert techniques of attitude scale construction. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 249–256). New York: John Wiley.

Fairweather, J. R. (1981). Reliability and validity of Q-method results: Some empirical evidence. *Operant Subjectivity, 5*(1), 2–16.

Fairweather, J. R., & Swaffield, S. R. (2002). Visitors' and locals' experiences of Rotorua, New Zealand: An interpretative study using photographs of landscapes and Q method. *International Journal of Tourism Research, 4*(4), 283–297.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, Mass: Addison-Wesley.

Fyock, J., Koepke, C. P., Meitl, J., Sutton, S., Thompson, E., & Engelberg, M. (2001). Beneficiary decisionmaking: The impact of labeling health plan choices. *Health Care Financing Review, 23*(1), 63–75.

Gelineau, E. P. (1981). A psychometric approach to the measurement of color preference. *Perceptual and Motor Skills, 53*(1), 163–174.

Hamm, B. C., & Cundiff, E. W. (1969). Self-actualization and product perception. *Journal of Marketing Research, 6*, 470–472.

Hogg, M. A., & Vaughan, G. M. (1995). *Social psychology: An introduction*. London: Prentice Hall.

Kitzinger, C. (1986). Introducing and developing Q as a feminist methodology: A study of accounts of lesbianism. In S. Wilkinson (Ed.), *Feminist social psychology* (pp. 151–172). Milton Keynes: Open University Press.

Lecouteur, A., & Delfabbro, P. H. (2001). Repertoires of teaching and learning: A comparison of university teachers and students using Q methodology. *Higher Education, 42*(2), 205–235.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 5–53.

Likert, R. (1967). The method of constructing an attitude scale. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 90–96). New York: John Wiley.

Martin, W. S., & Reynolds, F. D. (1976). On the usefulness of Q-methodology for consumer segmentation. *Journal of the Academy of Marketing Science, 4*(1), 440–445.

McDonald, C., & Vangelder, P. (1998). *Esomar handbook of market and opinion research*. Amsterdam: Esomar.

McKeown, B. F. (2001). Loss of meaning in Likert scaling: A note on the Q methodological alternative. *Operant Subjectivity, 24*, 201–206.

McKeown, B., & Thomas, D. (1988). *Q methodology*. Newbury Park, CA: Sage.

Mrtek, R. G., Tafesse, E., & Wigger, U. (1996). Q-methodology and subjective research. *Journal of Social and Administrative Pharmacy, 13*(2), 54–64.
Pease, D., Boger, R., Melby, J., Pfaff, J., & Wolins, L. (1989). Reliability of the Q-sort inventory of parenting behaviours. *Educational and Psychological Measurement, 49*(1), 11–17.
Poiesz, T. B. C. (1989). The image concept: Its place in consumer psychology. *Journal of Economic Psychology, 10*(4), 457–472.
Prasad, R. S. (2001). Clinical research and methods: Development of the HIV/AIDS Q-sort instrument to measure physician attitudes. *Family Medicine, 33*(10), 772–778.
Previte, J., Pini, B., & Haslam-McKenzie, F. (2007). Q methodology and rural research. *Sociologia Ruralis, 47*(2), 135–147.
Risdon, A., Eccleston, C., Crombez, G., & McCracken, L. (2003). How can we learn to live with pain? A Q-methodological analysis of the diverse understandings of acceptance of chronic pain. *Social Science & Medicine, 56*(2), 375–386.
Roberts, J. S., Laughlin, J. E., & Wedell, D. H. (1999). Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement, 59*(2), 211–233.
Schlinger, M. J. (1969). Cues on Q-technique. *Journal of Advertising Research, 9*(3), 53–60.
Schmolck, P., & Atkinson, J., (2002). PQMethod (Version 2.11). Computer program, available at <http://www.qmethod.org>.
Seiler, L. H., & Hough, R. L. (1970). Empirical comparison of the Thurstone and Likert techniques. In G. F. Summers (Ed.), *Attitude measurement* (pp. 159–173). Chicago, Ill: Rand McNally.
Smith, S. M., & Albaum, G. S. (2004). Measurement and scaling in marketing research. In S. M. Smith & G. S. Albaum (Eds.), *Fundamentals of marketing research* (pp. 371–412). Thousand Oaks, CA: Sage.
Smith, A. P., Young, J. A., & Gibson, J. (1999). How now, mad-cow? Consumer confidence and source credibility during the 1996 BSE scare. *European Journal of Marketing, 33*(11/12), 1107–1122.
Sommers, M. S. (1963). Product symbolism and the perception of social strata. In S. A. Greyser (Ed.), *Toward scientific marketing* (pp. 200–216). Chicago, Ill: American Marketing Association.
Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.
Stainton Rogers, R. (1995). Q methodology. In J. A. Smith, R. Harré, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 178–192). London: Sage.
Staley-Gane, L., Flynn, L., Neitzel, K., Cronister, A., & Hagerman, R. J. (1996). Expanding the role of the genetic counsellor. *American Journal of Medical Genetics, 64*(2), 382–387.
Steller, M., & Meurer, K. (1974). Zur Reliabilität eines Q-Sort zur Veränderungsmessung. [On the reliability of the Q-sort method for measuring changes.]. *Psychologische Beitrage, 16*(4), 618–624.
Stephenson, W. (1953). *The study of behaviour: Q-technique and its methodology*. Chicago, Ill: University of Chicago Press.
Stephenson, W. (1963). Public images of public utilities. *Journal of Advertising Research, 3*(4), 34–39.
Stephenson, W. (1965). Definition of opinion, attitude and belief. *Psychological Record, 15*(2), 281–288.
Stephenson, W. (1968). The contribution of Q to attitude research. In L. Adler & I. Crespi (Eds.), *Attitude research on the rocks* (pp. 160–171). Chicago: American Marketing Association.
Stephenson, W. (1969). Evaluation of public relations programs. *Revista Internationale di Scienze Economiche e Commerciali, 17*, 166–184.
Stephenson, W. (1979). Homo ludens: The play theory of advertising. *Revista Internationale di Scienze Economiche e Commerciali, 26*, 634–653.
Stricklin, M., & Almeida, R., (2004). PCQ for Windows (Academic Edition 1.4). Computer program, available at <http://www.pcqsoft.com/>.
Swaffield, S. R., & Fairweather, J. R. (1996). Investigation of attitudes towards the effects of land use change using image editing and Q sort method. *Landscape and Urban Planning, 35*(4), 213–230.
Treadwell, D. F., & Harrison, T. M. (1994). Conceptualizing and assessing organizational image: Model images, commitment, and communication. *Communication Monographs, 61*(1), 63–85.
Valenta, A. L., & Wigger, U. (1997). Q-methodology: Definition and application in health care informatics. *Journal of the American Medical Informatics Association, 4*(6), 501–510.
Van Riel, C. B. M., Stroeker, N. E., & Maathuis, O. J. M. (1998). Measuring corporate images. *Corporate Reputation Review, 1*(4), 313–326.
Verbeke, W., & Viaene, J. (1999). Beliefs, attitude and behaviour towards fresh meat consumption in Belgium: Empirical evidence from a consumer survey. *Food Quality and Preference, 10*(6), 437–445.
Webler, T., Tuler, S., & Krueger, R. (2001). What is a good public participation process? Five perspectives from the public. *Environmental Management, 27*(3), 435–450.
Wood, R. T. A., Griffiths, M. D., Derevensky, J. L., & Gupta, R. (2002). Adolescent accounts of the UK national lottery and scratchcards: An analysis using Q-sorts. *Journal of Gambling Studies, 18*(2), 161–183.
Zikmund, W. G. (1997). *Exploring marketing research*. Chicago, IL: Dryden Press.