

Quality of services 'applicable to production'

Citation for published version (APA):

Govers, C. P. M. (1993). Quality of services 'applicable to production'. In R. Grubström, & H. Hinterhuber (Eds.), *Strategic and operational issues in production economics : proceedings of the seventh international working seminar on production economics, Igls, Austria, February 17-21, 1992* (pp. 385-397). (International Journal of Production Economics; Vol. 30-31). Elsevier. <https://doi.org/10.1016/0925-5273%2893%2990107-V>, [https://doi.org/10.1016/0925-5273\(93\)90107-V](https://doi.org/10.1016/0925-5273(93)90107-V)

DOI:

[10.1016/0925-5273%2893%2990107-V](https://doi.org/10.1016/0925-5273%2893%2990107-V)
[10.1016/0925-5273\(93\)90107-V](https://doi.org/10.1016/0925-5273(93)90107-V)

Document status and date:

Published: 01/01/1993

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Quality of services “applicable to production?”

C.P.M. Govers

*Graduate School of Industrial Engineering and Management Science, University of Technology,
Eindhoven, The Netherlands (In cooperation with the Frits Philips Institute for Quality Management)*

Abstract

Measuring quality judgements is usually aimed at evaluating the appraisal of a product or situation. Originating from the field of sociological research, the vignette method is aimed at uncovering the underlying principles on which judgements are based.

Two field surveys are reported in discussing the application and usefulness of this method in the field of quality research. The research design is outlined and the methods of the analysis of data are described.

In the first survey a complete factorial design is used and the data are analysed by the use of a loglinear effect model to determine interaction effects of independent variables (quality characteristics) on the dependent variable (the quality judgement).

The second survey is an incomplete design. As a measure of association, the θ -statistic is used to express the degree of domination of quality characteristics.

Introduction

In recent years, quality management has been given increasing attention as a fundamental prerequisite for efficient production. Quality concepts range from a more manufacturing-based approach to a transcendent approach [1]. Generally accepted is the user-based approach directed to meet external as well as internal customer's requirements.

From the point of view of strategic policy, the problem is how to define the critical (quality) characteristics (CQCs) of production. Product quality, flexibility, good value for money and reliability of delivery times, to name but a few factors, are increasingly being taken for granted. However, a general trend in the recognition of opportunities promotes a process of overall thinking. Hence, meeting customer's requirements deals with the concept of perceived quality being the customer's

judgement about a product's overall excellence or superiority [2].

The quality of production presents characteristics comparable to the quality of services. Quality dimensions which can be distinguished within service operations are: the tangible or intangible product quality in terms of the effects produced; the process quality, i.e. how the service is rendered, and the relational quality, i.e. the kindness of the service personnel as well as the client friendliness of the system [3]. The perceived quality results from a comparison of expectations with perceptions of performance. If we conceive the relations between consecutive production stages as an “internal supplier customer relation”, we encounter the same dimensions when speaking about quality of production.

To obtain more insight into customer expectations or judgements about the quality of services, we often use an interview or survey technique like the multiple-item scale SERQUAL developed by Parasharam et al. [4, 5]. However, in actual judgement behaviour the customer judges a situation as a whole. By implementing a survey composed

Correspondence to: C.P.M. Govers, Graduate School of Industrial Engineering and Management Science, University of Technology, Eindhoven, The Netherlands.

of isolated aspects, it can become difficult to discover which aspects dominate, and to what extent the interrelations between these aspects influence the judgement of what is a good and what is a bad service. Just as in social judgements, we have to deal with quite a number of aspects and complex interrelations.

Originating from studies in the field of social satisfaction, an important development was the realization that the technique of factorial survey approach based on the vignette method was generally applicable to a wide range of substantive areas involving the judgement of complex objects [6–8].

The purpose of this paper is to present this approach in the field of quality research. We discuss the practical use and outcomes of two surveys aimed at determining the relative weights of some CQCs of service quality. Furthermore, we will establish a connection between these research results and the possibilities to define and evaluate quality of production.

Research design

By measuring quality judgements in complex situations, we want to know what the underlying reasons are that make a judgement positive or negative. Cognitive structures based on interviewees' expressions of agreement or disagreement with a number of presented statements do not provide a satisfactory answer.

In the "vignette method", respondents are asked to give an *overall* judgement of fictitious situations. On a vignette, a situation that could be a real existing one is described by some short verbalizations, each containing a well-defined stimulus component. The method was developed by Rossi [6] in a research project on the subjectively experienced fairness of income distribution. Rossi created fictitious households wherein he described the age, educational level of husband and wife, their profession(s), income(s) and housing, number, gender and age of children. He put forth the

question of how "fair" the family income was: too much, just right or too little. By means of conjoint analysis, the sets of overall responses are decomposed to factorially designed stimuli, so that the effect of each stimulus component can be inferred from the respondent's overall evaluations of the stimulus.

The underlying assumption used to uncover the structure on which judgements are based is the belief that there is a relatively small number of characteristics of objects to which individuals pay attention. Furthermore, in many domains judgements are socially structured, i.e. there is more or less agreement among people on how much weight should be given to relevant characteristics and how these should be combined in order to make a judgement. For example, in choosing a motorcar, not all characteristics are salient. While there may seem to be an infinite number of ways in which one car differs from another, by and large car buyers pay attention only to a relatively small set of characteristics of the car they choose. In making quality judgements, it is presumed to be some kind of a rather consistent evaluating process, originating in a relatively regular way from the socially determined consensus on how such judgements should be made. So nowadays car buyers may generally weigh fuel efficiency as being more important than driver comfort. However, one individual may weigh fuel efficiency more or less heavily than is typical for all car buyers.

In factorial experiments, usually only a few dimensions with only a few levels within each dimension can be used. This disability lessens considerably the resemblance between the experiments and the real-life conditions. Orthogonality in experiments makes it possible to observe their effects uncontaminated by the usual overlapping in the real world, but is often bought at the price of oversimplification. Factorial surveys capture more faithfully the complexity of the real world and the conditions of real human judgements. At the same time, the surveys provide the ability to identify clearly the separate influences.

The use of conjoint analysis decomposes a set of overall responses to factorially

designed stimuli so that the "utility" of each stimulus can be interfered from the respondent's overall evaluations of the stimulus. The solution technique like multi-dimensional scaling (MDS) involves a type of analysis of variance in which the respondent's preferences serve as a criterion variable and the predictor variables are represented by the various factorial levels making up each stimulus [9, 10].

In the vignette design the following actions should be taken: identification of relevant characteristics, creation of vignettes and collection and analysis of data. In order to find out which factors or elements determine a quality judgement, a distinction should be made between factors that are of "absolute" importance, factors that can be influenced and factors that are difficult to control, such as natural circumstances or economic and political developments. If important factors are found that can be influenced, the cause/effect knowledge can be used to bring about a change in the quality judgement of the customer. Sometimes, a few interviews with experts may be sufficient. Other situations may require a research project on its own, for example, to investigate cognitive structures of potential elements.

Once the relevant factors have been identified, it is possible to concentrate on the terminology and formulations of characteristics to be used in the vignettes. By randomizing the sequence and the values of the vignette characteristics, the stimuli are combined to sets of vignettes which will be presented to the respondents. The obtained data set of the judgements can be analysed by one of the methods of multivariate analysis. In these cases, the quality judgement is the dependent variable and the vignette characteristics are the explanatory independent variables.

On the basis of two surveys, we will discuss the practical use and outcomes of this approach. The design and analysis of data for these surveys were suggested and developed by Dijkstra [7].

The first survey was aimed at determining the relative weights of some CQCs of campsites. Using a "multiplicative effect model", we

got an insight into the influence of the independent variables (the vignette characteristics) and the dependent variable (the judgement). The intention of this survey was to answer the following question: Which factors dominate the judgement of the attractiveness of a campsite?

The second survey was carried out in an industrial setting. The purpose of this research project was to identify quality features concerning products and (technical) services/assistance of a bakery additives company and to determine their importance as considered by (potential) clients. The choice and verbalization of the quality features was a special component of this research.

COCs of a campsite

Already for some years a Dutch regional association of about 16 campsite owners has been active in concerted promotion campaigns. A recent marketing survey among their visitors indicated that the most effective publicity was through hearsay. So, to improve the quality image, the association intended to set up a more formal quality assurance system. However, to formulate an adequate quality policy, they firstly had to find out the customers' expectations of (some) critical quality characteristics (CQCs) of a campsite.

To identify CQCs for this case, we made use of observations about quality aspects which became apparent from the marketing survey and of the outcomes of our discussions with experts. Since it was also our intention to find out whether or not the vignette method was suitable for this type of research, we decided to limit the number of vignette elements to five and to differentiate between only two values classes for each element. A large number of elements and values only induces more complicated descriptions per vignette, which affects the clarity of arrangement. The use of dichotomous values gives a better insight into the differences in judgements and diminishes scaling problems.

Table 1 outlines the quality dimensions of services, the translation of these dimensions into vignette elements as well as the two values linked with these elements. The formulation on the vignettes was condensed into keywords. However, the use of just keywords is not clear enough for the respondents. Therefore, the instructions given to the interviewees included a concise formulation of the terminology used. In the instruction the different values of the characteristics as used on the vignettes were described as follows:

A1 – "*Friendly treatment at the reception*". When you arrive, you have the feeling that you are a desired guest. People are courteous. They pay attention to your wishes and also offer ample information without you directly asking for it.

A2 – "*Detached treatment*". When you arrive, you get the feeling that you are just another number. People work very efficiently and pass on all the information but only if requested to do so.

B1 – "*Comprehensive level of facilities*". In addition to the basic facilities the estate is also provided with a restaurant, a swimming pool, a snackbar and, occasionally, a disco, a riding school and/or a tennis court.

B2 – "*Modest facilities*". The estate provides for a reception, ample sanitary facilities and simple sports and play grounds.

C1 – "*Quite a lot of privacy*". The campsite is subdivided into small areas and the standing plots are separated by green belts.

C2 – "*Limited privacy*". The campsite is quite open and the standing plots are situated close together.

D1 – "*Friendly social interaction*". The social contacts with the personnel are courteous, they take notice of your requests or observations and provide full support.

D2 – "*Stiff social interaction*". It is difficult to get into contact with the personnel. You do not know to whom questions can be addressed because personnel seems to have little time for you.

E1 – "*Intensive surveillance*". The staff is alert that everybody adheres to the camping rules. They take action when offences occur or are reported. You have the feeling that your camping place is well protected when you leave for an excursion.

E2 – "*Negligent surveillance*". The camping rules are not enforced. There is common nuisance, uncontrolled parking, unleashed dogs and camp fires although that is forbidden.

Based on the keywords, we can construct a number of k^n vignettes, with n being the number of characteristics and k the number of values for every characteristic. So, in our case we had $2^5 = 32$ vignettes. To explore a complete factorial design, we have to check that all

Table 1

The keywords used to formulate the CQCs and their values on the vignettes in relation to the quality dimensions of services

Dimension	Characteristic	Values
Process/relational quality	A Treatment at the reception	1 Friendly 2 Detached
Tangible product	B Level of facilities	1 Comprehensive 2 Modest
Intangible product	C Camping plot privacy	1 Quite a lot 2 Limited
Relational quality	D The social interaction with the staff	1 Friendly 2 Stiff
Process quality	E Surveillance of the campground rules	1 Intensive 2 Negligent

combinations of the independent variables lead to real situations, which is the case in this situation.

After an introduction, the respondents were asked to judge the different situations described on the vignettes presented (see Fig. 1). To avoid undesired interactions, the sequence of the characteristics on the vignettes was randomly alternated. During the high season (July–August 1988) a total number of 576 visitors equally spread over the 16 campsites of the association was interviewed. We split up the season into three periods of 16 days and visited the campsites once in a period. Upon each visit, 12 randomly selected campers were invited to give a judgement of a set of eight vignettes. The sets were composed in advance by drawing randomly from the 32 vignettes. So, we could easily check that all vignettes were almost equally represented and were judged 144 times on an average.

On the registration form we recorded the campsite code, the period, the weather conditions and the judgements on the eight vignettes. After the judgements of the vignettes the respondents were asked to give also their opinion concerning each of the five characteristics with respect to the campsite involved and to provide information regarding age, gender, nationality experience of camping, whether they explicitly had selected that specific campsite and whether the camping fee influenced the quality judgement.

Because one respondent may judge a situation as being "very good", while another

judges it as "satisfactory", we made a dichotomy "positive" and "negative" for the judgements. Amalgamation of "very good", "good" and "satisfactory" into "positive" and the other categories into "negative" is a more realistic representation of the respondents' judgements and simplified the analysis of the data.

If the judgement is dichotomous, an appropriate data representation model is a log linear model for a frequency distribution [11]. Since we were able to conduct a complete factorial design, the model is saturated. In this model a coefficient (τ) indicates the first- and higher-order influence of the independent variables (the vignette characteristics) on the dependent variable (the judgement of the vignette).

To illustrate the method of analysis, we present in Table 2 a frequency distribution with only two independent variables *A* and *B* and one dependent variable *C*, all three being dichotomous.

For this distribution we can develop the following multiplicative frequency model:

$$F_{ijk}^{ABC} = \eta * \tau_{ijk}^{ABC} * \tau_{ij}^{AB} * \tau_{ik}^{AC} * \tau_{jk}^{BC} * \tau_i^A * \tau_j^B * \tau_k^C \tag{1}$$

Verbally, this means: the frequency in cell *i* of *A*, cell *j* of *B*, and cell *k* of *C* is the product of an overall mean (η) and a number of parameters (τ). The parameter τ is a skewness measure defined by the ratio of the frequencies

Vignette no. 10		
Quite a lot of privacy		
Intensive surveillance		
Comprehensive level of facilities		
Detached treatment at the reception		
Stiff social interaction		
How do you judge a camping that conforms to this description?		
(1) very good	(2) good	(3) satisfactory
(4) insufficient	(5) bad	(6) very bad

Fig. 1. Example of a vignette used.

Table 2
Example of a frequency distribution for two independent variables (*A*, *B*) and the dependent variable *C*

		<i>C</i>	
		1	2
<i>A</i>	<i>B</i>		
1	1	38	16
1	2	30	44
2	1	22	26
2	2	50	12

in the elementary cells. From Table 2 we get

$$\tau_1^C = \left(\prod_i \prod_j \frac{f_{ij1}}{f_{ij2}} \right)^{1/8}, \quad \tau_1^C = 1/\tau_2^C \quad (2)$$

$$\tau_{11}^{AC} = \left(\prod_j \frac{f_{1j1}}{f_{1j2}} * \frac{f_{2j2}}{f_{2j1}} \right)^{1/8},$$

$$\tau_{11}^{AC} = \tau_{22}^{AC} = 1/\tau_{12}^{AC} = 1/\tau_{21}^{AC}$$

and so on; see also Eq. (4).

The conditional frequency ratios of the classes of the dependent variable indicate single and interaction effects. It may be noted that, for every parameter containing more than one variable, the interaction effect between these variables becomes increasingly weak as the value approaches 1. A coefficient value of 1 means "no influence", while the influence becomes stronger as the value (or its reciprocal) deviates more from 1, in the sense that a coefficient of 4 indicates an equally strong (positive) effect as a coefficient of 0.25 ($= \frac{1}{4}$) (negative).

As opposed to the frequency model [Eq. (1)], in the effect model the ratio of frequencies in the classes of a dependent variable is calculated for all class combinations of the independent variable. In the illustration we can explore the effect model by the ratio C_1/C_2 within cells A_1B_1, A_1B_2, A_2B_1 and A_2B_2 . This model then explains the skewness of C as an effect of the skewness of A and B . By dividing the multiplicative model [Eq. (1)] for $k = 1$ by the model for $k = 2$ ($F_{ij1}^{ABC}/F_{ij2}^{ABC}$), we find that η and all parameters without the variable C cancel each other; so, only the τ ratios for C_1/C_2 will remain. We will indicate the model ratio here by Ω . Using the following notation

$$\Omega^{ABC} = \gamma^{ABC} * \gamma^{AC} * \gamma^{BC} * \gamma^{\bar{C}} \quad (3)$$

From Table 2 these ratios (for example, only two of them) are

$$\begin{aligned} \gamma^{\bar{C}} &= \left(\frac{\tau_1^C}{\tau_2^C} \right) = (\tau_1^C)^2 \\ &= \left(\frac{38 * 30 * 22 * 50}{16 * 44 * 26 * 12} \right)^{1/4} = 1.5458 \end{aligned} \quad (4)$$

$$\begin{aligned} \gamma^{A\bar{C}} &= \left(\frac{\tau_{11}^{AC}}{\tau_{12}^{AC}} \right) = (\tau_{11}^{AC})^2 \\ &= \left(\frac{38 * 30 * 26 * 12}{16 * 44 * 22 * 50} \right)^{1/4} = 0.8232 \end{aligned}$$

Equation (3) will be read as the skewness of C is a product of the marginal distribution of C ($\gamma^{\bar{C}}$) and a number of parameters indicating the influence of the independent variables.

Using the γ parameters, the original distribution of a saturated model can be got back by applying the following calculation rules:

- Multiplying by γ -terms (half the number) for
 - all the independent variables in class 1,
 - all combinations of classes 1 of the independent variables,
 - all combinations of independent variables when the sum of the classes is an even number more than the number of independent variables.

Dividing by γ for the other situations.

The calculation is illustrated in Table 3.

The analysis of the influence of the vignette characteristics on the quality judgement is carried out on the basis of the observed frequency distributions. The γ parameters of Ω are presented as a ratio between the frequencies in the

Table 3
Calculation of the distribution of table 2 by γ terms

A	B	$\gamma^{\bar{C}}$	$\gamma^{A\bar{C}}$	$\gamma^{B\bar{C}}$	$\gamma^{\bar{C}}$
1	1	$C_1/C_2 = 1.5458 * 2.0351 * 0.8232 * 0.9171 = 38/16$			
1	2	$C_1/C_2 = 1.5458/2.0351 * 0.8232/0.9171 = 30/44$			
2	1	$C_1/C_2 = 1.5458/2.0351/0.8232 * 0.9171 = 22/26$			
2	2	$C_1/C_2 = 1.5458 * 2.0351/0.8232/0.9171 = 50/12$			

Note: Differences are caused due to rounding off.

classes "positive" and "negative" of the judgements (F_+/F_-); so, that a value $\gamma > 1$ indicates a positive influence on the quality judgement. In order to prevent too much attention being paid to parameters which differ very little from 1, we arbitrarily set the limits at $\Omega = 1.05$ and $\Omega = 0.95$ (within these limits the skewnesses vary between 51% and 49%, which means that there is practically no influence). Asserting this criterion, we found only first- and second-order effects. Table 4 summarizes the effects of class 1 (means "positive" description) characteristics on the quality judgements. The skewnesses are also expressed as percentage positive (F_+) and negative (F_-).

From Table 4 we can gather the influence of each characteristic (or combinations) on the quality judgement. For example, the influence of the combination of the *C* and *E* characteristics on the quality judgement is

$$\Omega^{CE} = \gamma^C * \gamma^E * \gamma^{CE} = 3.024 * 2.167 * 1.224$$

$$= 8.02 (F_+/F_- = 88.9/11.1)$$

Table 4
Summary of first- and second-order effects of a positive characteristic on the quality judgement

Characteristics					First-order effects		
A	B	C	D	E	$\gamma = F_+/F_-$	% F_+	% F_-
		*			3.024	75.1	24.9
			*		2.191	68.7	31.3
				*	2.167	68.4	31.6
*					1.986	66.5	33.5
	*				1.698	62.9	37.1
					Second order effects		
		*		*	1.224	55.0	45.0
			*	*	1.170	53.9	46.1
*		*			1.113	52.7	47.3
	*			*	1.088	52.1	47.9
		*	*		1.076	51.8	48.2
*				*	1.066	51.6	48.4

- A = treatment at the reception.
- B = level of facilities.
- C = camping plot privacy.
- D = social interaction with staff.
- E = surveillance of camping ground rules.

It was found that camping plot privacy was the most important factor and the level of facilities, in comparison with the other characteristics, the least. Only the combinations of *CE* and *DE* exercise some traceable influence (F_+ 55.0% and 53.9%, respectively). We also checked the judgements of each vignette separately. For all vignettes with three or more characteristics in class 2 (negative), the ratio was less than 1. This means that the number of negative judgements is higher than the positive ones.

The findings underline the importance of camping plot privacy (*C*) and the social interaction with the staff (*D*). The characteristic *C* is not easily to be changed directly, but *D* and also "surveillance" (*E*) can quite well be used to the management of a campsite to upgrade the quality judgement of customers.

The analysis of respondents' characteristics did not give evidence of the existence of certain subgroups within this population. Also, we did not find significant differences between periods of interviewing, nor between weather conditions.

As mentioned before, at the end of the interview, we also asked the respondents to give their evaluation of each characteristic on the same 6-point scale (very bad . . . very good) for the campsite involved. We wanted to get an idea about quality scores of the specific campings.

These scores were treated by analysis of variance and presented as "quality profiles" indicating the mean and standard deviation of the scores per characteristic for each of the 16 campsites involved. Based on this analysis, we concluded that only seven campsites had a score of "good" or "very good" for characteristic *D* and only four campsites had this score for characteristic *E*. Furthermore, two campsites scored for characteristic *E* "insufficient" or less.

COCs for rendering products and (technical) assistance

The second survey was conducted to determine CQC for a producer of improving and

release agents which are raw aid-materials (additives) for bakeries. The aim of this survey was to get an insight into the general judgement of (potential) bakers about the performance of the firm.

Because of the sensitiveness of the findings from a marketing point of view, we will pay attention only to the experiences and differences regarding this project and the research project at the campsites.

Firstly, we had to establish a good understanding of the "product" in relation to market segmentation for an adequate verbalization of vignette characteristics. Secondly, a test survey made obvious that we could not perform a complete factorial design, which had consequences for the analysis of data.

The market for the products can be divided into three segments: craft bakery; middle-sized bakery; and industrial bakery. A craft bakery delivers directly to consumers. Generally, a craft bakery uses less than 250 bags of 50 kg flour a week. Industrial bakeries serve the consumer by intermediate trade and realize a weekly flour turnover of more than 700 bags. The dividing line is arbitrary and, therefore, we introduced the category of middle-sized bakeries (between 250 and 700 bags a week). Among the approximately 3350 bakeries in the Netherlands, there are about 110 industrial ones which produce 55% to 60% of the total national bread turnover. This percentage also counts for the share of the firms turnover regarding industrial bakeries. They are served by direct deliveries and require tailor-made specials. The other segments are generally served via buying organizations or cooperatives. To stay in the market, it is needed to accompany this kind of products with good services like (technical) product information and field services such as demonstration, instruction and trouble shooting.

After interviews with experts and a study of the product and service range, we made up a list of possible characteristics. From this list we deleted the characteristics which are taken for granted (evident or submitted to governmental regulations) or go beyond the firms

control. Nevertheless, the number of characteristics remained too large; so, a reference group of firm's representatives was asked to decide which characteristics should be selected. During a brain-storming session, it became evident that they could not reach a concensus about a limited set of CQCs. The main reason was that the issues were determined mainly by the personal relationships between the representatives and their customers. So, after this session, all the mentioned elements were simply listed and the representatives were invited to score each element on a 5-point scale ranging from "very important" to "very unimportant". By analysis, we could cluster the elements into six homogeneous groups. Each cluster was amalgamated by a characteristic that seemed to represent best the elements in the cluster. This resulted in six vignette characteristics (partnership, information, assortment, delivery, communication, reliability).

On the vignettes, reference was made to a fictitious firm Nesup (New Supplier) and the characteristics were stated as follows (all positive formulations):

- A The firm Nesup is a partner in your activities (the company shows involvement; suggests technical and commercial innovations).
- B You will get additional information from the firm Nesup (the company provides uncalled information on consumer purchasing trends, materials and new products).
- C The assortment of Nesup is complete and clear (sufficient stock and clear instructions for the application).
- D Nesup always delivers at the right time and the right place.
- E Nesup's representative is easy to communicate with (has a good knowledge of branches, materials and techniques and understands the jargon).
- F Nesup is reliable, trustworthy and keeps its word.

The dichotomous values of the characteristics were varied by formulating in a confirmatory or negatory way. So, in total we had $2^6 = 64$ vignettes.

Each vignette was judged on a 6-point scale ranging from "very good" to "very bad". Besides the judgements of the vignettes, respondents who were clients were asked to judge the firm in question about each separate characteristic at the same scale.

Maximizing the number of vignettes to be judged weighted against interview time and interviewees concentration, it was decided to present to each respondent sets of ten vignettes. At the test survey, it appeared that vignettes with three (or more) negative statements led to extreme negative judgements. Therefore, vignettes with more than three negative statements and also the vignette with six positive statements (should be good by definition) were deleted. In total 41 vignettes have been judged.

The design can still be complete within the restriction to vignettes with just one, two or three negative values. When second-order and higher interactions are of no interest, an incomplete design is also preferable to reduce the number of parameters.

It was intended to stratify the survey in accordance with the market segmentation (about 32% craft, 17% middle-sized and 51% industrial bakeries). Therefore, we had to care for an adequate sampling of the vignettes. The 41 vignettes were arranged into categories I, II and III with, respectively, one, two and three negative values for the characteristic(s). So, cat. I contains #6 vignettes, cat. II #15 and cat. III #20. To prepare sets of ten vignettes, we calculated what would be the best draws per category so that, especially for cat. I and cat. II vignettes, the number of times a vignette will be judged by interviewees from each segment is more or less equal. We decided to use #2 draws from cat. I, #5 from cat. II and #3 from cat. III per set.

At the rate of 100 respondents, this results in a number of expected judgements per vignette for each category, as presented by \bar{x} (as an integer) in Table 5. The survey was conducted countrywide among 103 bakers based on stratified random sampling per market segment. Since we used an incomplete factorial design and measured the vignette judgements on an

ordinal scale, the method of analysis was adopted.

To determine the relative importance of the characteristics, the frequency patterns of the vignettes are ranked. The most commonly used measures of association for ordinal variables are those based on the numbers of concordant and discordant pairs in the sample. A pair of observations is *concordant* if the member that ranks higher on variable X also ranks higher on variable Y . A pair of observations is *discordant* if the member that ranks higher on X ranks lower on Y . Consider some pairs of individuals on (X, Y) :

$A (2, 1)$; $B (3, 5)$; $C (7, 4)$; and $D (7, 2)$.

The pairs $(A-B)$, $(A-C)$ and $(A-D)$ are concordant since the second individual is ranked higher than the first both in amount of X and Y (1 is the high and 7 the low end of the scales). $(A-B)$ equals 2, 3 and 1, 5, $(A-C)$ 2, 7 and 1,

Table 5
Expected number of judgements (\bar{x}) per vignette per category ($N = 100$)

Category	I	II	III	Total
# Vignettes	6	15	20	41
# Per set	2	5	3	10
Craft # Vign.	64	160	96	320
(32%) \bar{x}	11	11	5	
Mid. # Vign.	34	85	51	170
(17%) \bar{x}	6	6	2	
Ind. # Vign.	102	255	153	510
(51%) \bar{x}	17	17	8	
total #	200	500	300	1000
\bar{x}	34	34	15	

Note: From 100 respondents, we will get 1000 ($100 * 10$) judgements, of which 320 are from craft bakers (32% market share); two out of ten are from cat. I (64 vignettes). Category I numbers 6 vignettes; so, each will be judged at about 11 times on an average. Table 5 shows that a 2/5/3 scheme per set leads to an equal number of judgements for category I and II vignettes within each segment.

4 while (A-D) equals 2, 7 and 1, 2. Pairs (B-C) and (B-D) are discordant (3, 7 and 5, 4, and 3, 7 and 5, 2, respectively) while pair (C-D) is indifferent because the rank on the X scale is the same.

The general formula for calculating the concordance (C) of an *i*j* frequency table is

$$C = \sum_{i < k} \sum_{j < l} n_{ij}n_{kl}$$

where the first summation is over all pairs of rows *i < k* and the second summation is over all pairs of columns *j < l*. The number of discordant pairs *D* of observations is

$$D = \sum_{i < k} \sum_{j > l} n_{ij}n_{kl}$$

Several ordinal-ordinal measures of association are based on the difference *C - D* between the numbers of concordant and discordant pairs [12]. For each of these measures, the association is said to be positive if *C - D > 0* and negative if *C - D < 0*. In each case a reversal in the category orderings of one variable simply causes a change in the sign of the measure.

Of the *C + D* pairs of observations that are untied on both variables, *C/C + D* is the proportion of concordant pairs and *D/C + D* is the proportion of discordant pairs. The difference between these two proportions is conceived as a measure of distance and can be expressed as

$$\theta = \frac{C - D}{C + D}, \quad -1 \leq \theta \leq +1$$

By definition, the value of θ lies between -1 and $+1$. A value $|\theta| = 1$ implies that the relationship is monotone. Independence implies that $\theta = 0$, but the converse of this does not hold.

The results are presented in Table 6. The vignette numbers (first column) are ranked

from most negative judgements to least negative based on the θ 's (last column).

Column 2 shows the pattern of the negative statements in the vignettes. In column 3 the negative statements for category I and II vignettes are labeled. Column 4 presents the frequencies of the judgement scores (6 very bad 1 very good). In the fifth column the mean score is calculated; the upper part of the table contains the vignettes which are found unsatisfactory or worse (mean ≤ 4), while the lower part contains the vignettes which are found satisfactory or better.

The scores for the vignettes are indicator variables that index whether the judgement is less negative. Based on the profiles of the vignettes (column 2), we determined the order of the quality features as formulated in the vignette characteristics. Emphasis is laid on vignettes belonging to categories I and II with one or two negative values only. From category I vignettes the characteristic "Communication" (E) had the least influence on the general judgement (mean score 2.24). The characteristic "Delivery" (D) is dominant above the others. The difference is quite high (mean score 3.91). If this is the only negative characteristic, the supplier will be judged nearly at unsatisfactory on an average.

Bringing in vignettes with two negative values, for example, "delivery" and "reliability" (pattern DF, results in a mean score of 4.22). Here the cumulative effect of negative elements becomes clear.

Figure 2 shows the dominance of the characteristics for category I and category II vignettes. The dominance in Fig. 2 is indicated by a distance measure. However, there is a scaling difference between category I and II; so, a direct comparison of distances is not allowed. What can be seen is that two negative characteristics do not change the order, but it appears that characteristic F becomes more dominant when it occurs in combination with another negative characteristic.

The findings of this survey show the importance of the characteristics D and F. Figure 2 presents the results of the entire population. The ranking remains the same for the three

Table 6
Ranking of the judgements of the vignettes (mean scores ≤ 4)

Vig. No.	Characteristics						Neg. I & II	Judgement						Mean	Thêta θ
	A	B	C	D	E	F		6	5	4	3	2	1		
45	-		-	-				6	3	3	1	0	0	5.08	-
36	-			-	-	-		8	7	4	1	0	0	5.10	+ 0.01
15			-	-	-			4	5	3	1	1	0	4.71	+ 0.24
22			-	-		-		2	4	4	0	0	0	4.80	+ 0.04
8				-	-	-		0	8	3	0	0	0	4.73	+ 0.03
23	-			-	-			3	8	7	1	0	0	4.68	+ 0.08
12			-	-	-	-		4	2	6	2	0	0	4.57	+ 0.11
14			-	-	-	-		2	6	3	2	0	1	4.36	+ 0.04
53	-	-		-				2	6	9	1	0	0	4.50	+ 0.04
39	-			-	-			2	5	11	0	1	0	4.37	+ 0.11
42	-		-			-		2	4	2	2	2	0	4.17	+ 0.04
43	-		-		-			1	8	5	4	1	0	4.21	+ 0.01
6				-		-	DF	3	6	18	5	0	0	4.22	+ 0.06
36	-				-	-		1	4	6	2	0	1	4.07	+ 0.00
26		-	-			-		0	6	5	4	0	0	4.13	+ 0.01
29		-	-	-				2	4	11	2	2	0	4.10	+ 0.04
50	-	-				-		5	4	6	2	1	0	4.00	+ 0.05
21		-		-			BD	3	6	23	6	1	0	4.05	+ 0.03
10			-			-	CF	3	3	15	7	2	0	4.06	+ 0.04
13			-	-			CD	1	4	17	7	1	0	4.03	+ 0.00
4					-	-	EF	3	4	17	10	0	0	4.00	+ 0.05
37	-			-			AD	1	7	17	8	1	1	3.89	+ 0.01
7				-	-		DE	3	2	22	6	2	0	3.94	+ 0.01
5				-			D	3	3	19	9	0	1	3.91	+ 0.04
34	-					-	AF	2	6	16	5	5	0	3.85	+ 0.00
20		-			-	-		0	1	11	3	0	0	3.87	+ 0.04
18		-				-	BF	2	9	12	8	6	0	3.81	+ 0.02
57	-		-					1	2	6	4	2	0	3.73	+ 0.07
27		-	-			-		0	4	6	6	2	0	3.67	+ 0.03
51	-	-				-		0	4	3	4	4	0	3.47	+ 0.14
41	-		-				AC	0	3	11	13	5	0	3.36	+ 0.05
11			-			-	CE	0	2	14	8	9	0	3.27	+ 0.05
49	-	-					AB	1	3	7	19	3	1	3.32	+ 0.01
19		-				-	BE	1	0	12	19	5	0	3.27	+ 0.02
2						-	F	2	1	10	14	6	3	3.17	+ 0.09
25		-	-				BC	1	1	10	13	8	1	3.15	+ 0.01
35	-					-	AE	0	1	11	7	14	2	2.86	+ 0.20
9			-				C	0	0	5	15	16	0	2.69	+ 0.12
17		-					B	0	0	4	9	15	3	2.45	+ 0.26
33	-						A	1	1	1	9	17	5	2.38	+ 0.14
3					-		E	0	0	2	10	16	6	2.24	+ 0.05

market segments. However, there are some differences in weighting by the subgroups.

Although the characteristics represent a cluster of correlated items, we could nevertheless indicate CQCs to be used for performance improvement.

Especially in the bakeries survey, the analysis of clients' judgements on each separate characteristic demonstrated quite strong halo and leniency effects. These phenomena are quite well known in social psychology [14].

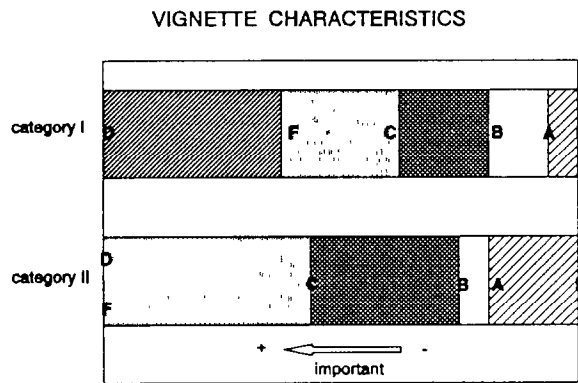


Fig. 2. Dominance of the vignette characteristics.

The halo effect is the propensity to reconcile judgements about specific characteristics with a general opinion about a person or situation.

The leniency effect or "positive bias" is the tendency to express less negative feelings or judgements, the more the subject matter becomes important. This means that the clients incline to judge the important factors concerning their (commercial) relations as positive. We found only "unsatisfactory" or worse judgement categories for the "less important" characteristics.

On the halo effect, we have an adequate check consisting of the model of the Likert scale [7, 13]. The basic idea of this model is that a latent, unmeasured factor is the exclusive cause of variation in a set of measured variables. For five out of the six vignette characteristics, this effect is demonstrated by the item rest correlations to prove internal consistency and by Cronbach's alpha to show some predictable factors concerning homogeneity. The alpha (0.85) and the item rest correlation for A (0.63), B (0.70), C (0.73); D (0.68) and F (0.59) are sufficiently large to identify the five elements as a subscale. The halo effect is generally present but not always visible. In the vignette method it can be proved that by adding up scores on specific judgements which fulfil the requirements of the Likert model, the sum scores are valid and reliable representatives of the latent factor (the general evaluation)

Table 7

Judgements about the company compared with the judgements of vignettes

Characteristics	Company		Vignette	
	Mean score	Rank	Mean score	Rank score
E Communication	2.67	1	2.24	6
B Information	2.42	2	2.45	4
A Partnership	2.46	3	2.38	5
C Assortment	2.30	4	2.69	3
F Reliability	2.28	5	3.17	2
D Delivery	1.94	6	3.91	1

much more than its characteristics or the general judgement itself [7].

The leniency effect is demonstrated by the judgements about the company corresponding to each vignette characteristic. In Table 7 the rankings of judgements about the company and the single negative vignette characteristics are compared. The rank numbers are in conformity with the theta statistic. The mean scores for the judgements range from very good (1) to very bad (6); a score < 4 means satisfactory or better. The leniency phenomenon holds the warning that the more negatively the vignette characteristics are evaluated, the more respondents will be reticent in expressing negative opinions relating to these aspects. Assessing customer perceptions of quality on single items may cause an unjustified rosy impression, leading to false conclusions that such characteristics are really experienced as "good".

Conclusions

We believe that it is of great importance to know how *external* as well as *internal* clients judge the performance of production in relation to quality, flexibility, price, delivery and so on. The advantage of the vignette method is that the attention is focused on uncovering the principles that underlie these judgements. In this approach the research question is shifted

from "What is the judgement of an existing situation?" to "Which factors determine this judgement?". Once the relative importance and the interrelations between these factors are determined, we can establish a strategic policy of production.

The vignette method owes its force to the circumstances that we can conduct a laboratory research setting (controlled stimuli) in a real-world situation. Because the respondent is asked to give a judgement of "fictitious" situations, it will be easier for him to keep a distance from a concrete subject matter. This will reduce leniency or extreme response effects and offer good possibilities to decompose general judgements of complex situations.

Our pilot projects were aimed at evaluating the application and usefulness of this method in quality judgement research.

For the vignette method to be applicable and meaningful, some conditions must be fulfilled: the features to be examined must be operationable; the features have to be manipulable and are not allowed to be of absolute importance. In that sense in the campsite project, for example, we did not take into account the characteristics concerning hygienic aspects. Furthermore, this method is limited. To create high reliability and still test realistic population size at random, only a few features can be taken into account. Another reason to limit the number of features is the mental grasp of the respondents. In spite of these restrictions, in our opinion both surveys have demonstrated that the technique is generally applicable.

To compare and to validate the two methods of analysis, we also analysed the data of the complete factorial design (campings) in the same way as the incomplete design (bakeries). We came to the same ranking of characteristics. However, the incomplete design does not give an insight into second-order and higher interactions. Mostly, these interactions become very weak (see Table 4). If there is no practical or theoretical objection to a complete design, or for some reasons interactions are of no interest, an incomplete design can be used equally well.

Acknowledgement

I thank L. Dijkstra for elaborating the research design and methods of analysis and for his assistance during the graduation projects of A. van Erp (campings) and G. van Geemen (bakeries project), who performed the field surveys.

References

- [1] Garvin, D.A., 1984. What does "product quality" really mean? *Sloan Management Review*, Fall 1984.
- [2] Zeithaml, V.A., 1987. Defining and relating price, perceived quality and perceived value. Report no. 87-101, Marketing Science Institute, Cambridge, MA.
- [3] Boomsma, S. and Borrendam, A., 1987. *Kwaliteit in Diensten (Quality in Services)*. Kluwer, Deventer.
- [4] Parasuraman, A., Zeithaml, V.A. and Berry, L.L., 1988. Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *J. Retailing*, 64 (Spring): 12-40.
- [5] Parasuraman, A., Berry, L.L. and Zeithaml, V.A., 1990. An empirical examination of relationships in an extended service quality model. Report no. 90-122, Marketing Science Institute, Cambridge, MA.
- [6] Rossi, P.H. and Nock, S.L., 1982. *Measuring Social Judgements: The Factorial Survey Approach*. Beverly Hills.
- [7] Dijkstra, L., 1992. A general design for quality judgement research. Paper presented at Workshop on Quality Management, EIASM, May 21, Maastricht.
- [8] Govers, C.P.M., 1992. The judgement of service quality. In P. Kunst and J. Lemmink (eds.), *Quality Management in Services*. van Gorcum, Assen/Maastricht.
- [9] Green, P.E., 1978. *Analyzing Multivariate Data*. Dryden Press, Hinsdale.
- [10] Green, P. and Tull, D., 1978. *Research for Marketing Decisions*. 4th ed. Prentice-Hall, Englewood Cliffs, NJ.
- [11] Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W., 1975. *Discrete Multivariate Analysis*. MIT Press, Cambridge.
- [12] Agresti, A., 1984. *Analysis of Ordinal Categorical Data*. Wiley, New York.
- [13] Dijkstra, L., 1991. The Likert attitude scale: theory and practice. Report TUE/BdK/ORS/91/03, University of Technology, Eindhoven.
- [14] Freedman, J.L., Sears, D.O. and Carlsmith, J.M., 1978. *Social Psychology*. Prentice Hall, Englewood Cliffs, NJ.