Customer Expectations, Priorities, Needs, and “Voice”

Although customers seldom spark true innovation (for example, they are usually unaware of state-of-the-art developments), their input is extremely valuable. Obtaining valid customer input is a science itself. Market research firms use scientific methods such as critical incident analysis, focus groups, content analysis and surveys to identify the “voice of the customer.” Noritaki Kano developed the following model of the relationship between customer satisfaction and quality (Fig. 2.4).

The Kano model shows that there is a basic level of quality that customers assume the product will have. For example, all automobiles have windows and tires. If asked, customers don’t even mention the basic quality items, they take them for granted. However, if this quality level isn’t met the customer will be dissatisfied; note that the entire “Basic quality” curve lies in the lower half of the chart, representing dissatisfaction. However, providing basic quality isn’t enough to create a satisfied customer.

The “Expected quality” line represents those expectations which customers explicitly consider. For example, the length of time spent waiting in line at a checkout counter. The model shows that customers will be dissatisfied if their quality expectations are not met; satisfaction increases as more expectations are met.

The “Exciting quality” curve lies entirely in the satisfaction region. This is the effect of innovation. Exciting quality represents unexpected quality items. The customer receives more than they expected. For example, Cadillac pioneered a system where the headlights stay on long enough for the owner to walk safely to the door. When first introduced, the feature excited people.

Competitive pressure will constantly raise customer expectations. Today’s exciting quality is tomorrow’s basic quality. Firms that seek to lead the market must innovate constantly. Conversely, firms that seek to offer standard quality must constantly research customer expectations to determine the currently accepted quality levels. It is not enough to track competitors since expectations are influenced by outside factors as
well. For example, the quality revolution in manufacturing has raised expectations for service quality as well.

**Garden Variety Six Sigma Only Addresses Half of the Kano Customer Satisfaction Model**

Some people, including your author, believe that even Six Sigma doesn’t go far enough. In fact, even “zero defects” falls short. Defining quality as only the lack of nonconforming product reflects a limited view of quality. Motorola, of course, never intended to define quality as merely the absence of defects. However, some have misinterpreted the Six Sigma program in this way.

One problem with “garden variety” Six Sigma is that it addresses only half of the Kano model. By focusing on customer expectations and prevention of nonconformances and defects, Six Sigma addresses the portion of the Kano model on and below the line labeled “Expected Quality.” While there is nothing wrong with improving these aspects of business performance, it will not ensure that the organization remains viable in the long term. Long-term success requires that the organization innovate. Innovation is the result of creative activity, not analysis. Creativity is not something that can be done “by the numbers.” In fact, excessive attention to a rigorous process such as Six Sigma can detract from creative activities if not handled carefully. As discussed previously, the creative organization is one which exhibits variability, redundancy, quirky design, and slack. It is vital that the organization keep this paradox in mind.

**Quality Function Deployment**

Once information about customer expectations has been obtained, techniques such as quality function deployment (QFD) can be used to link the voice of the customer directly to internal processes.

Tactical quality planning involves developing an approach to implementing the strategic quality plan. One of the most promising developments in this area has been policy deployment. Sheridan (1993) describes policy deployment as the development of a measurement-based system as a means of planning for continuous quality improvement throughout all levels of an organization. Originally developed by the Japanese, American companies also use policy deployment because it clearly defines the long-range direction of company development, as opposed to short-term.

QFD is a customer-driven process for planning products and services. It starts with the voice of the customer, which becomes the basis for setting requirements. QFD matrices, sometimes called “the house of quality,” are graphical displays of the result of the planning process. QFD matrices vary a great deal and may show such things as competitive targets and process priorities. The matrices are created by interdepartmental teams, thus overcoming some of the barriers which exist in functionally organized systems.

QFD is also a system for design of a product or service based on customer demands, a system that moves methodically from customer requirements to specifications for the product or service. QFD involves the entire company in the design and control activity. Finally, QFD provides documentation for the decision-making process. The QFD approach involves four distinct phases (King 1987):

1. **Organization phase.** Management selects the product or service to be improved, the appropriate interdepartmental team, and defines the focus of the QFD study.
Please check the appropriate box in response to the following statement: “The customer service representative was knowledgeable.”

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree

Intensity scales are very easy to construct. They are best used when respondents can agree or disagree with a statement. A problem is that statements must be worded to present a single side of an argument. We know that the respondent agrees, but we must infer what he believes. To compensate for the natural tendency of people to agree, statements are usually presented using the converse as well, for example, “The customer service representative was not knowledgeable.”

When using intensity scales, use an odd-numbered scale, preferably with five or seven categories. If there is a possibility of bias, order the scale in a way that favors the hypothesis you want to disprove and handicaps the hypothesis you want to confirm. In this way you will confirm the hypothesis with the bias against you—a stronger result. If there is no bias, put the most undesirable choices first.

- Semantic differential format—In this format, the values that span the range of possible choices are not completely identified; only the end points are labeled. For example,

  Indicate the number of times you initiated communication with your customer in the past month.

  few  □  □  □  □  □  □  □  many

  The respondent must infer that the range is divided into equal intervals. The range seems to work well with seven categories.

  Semantic differentials are very useful when we do not have enough information to anchor the intervals between the poles. However, they are very difficult to write well and if not written well the results are ambiguous.

Survey Development Case Study*

This actual case study involves the development of a mail survey at a community hospital. The same process has been successfully used by the author to develop customer surveys for clientele in a variety of industries.

The study of service quality and patient satisfaction was performed at a 213 bed community hospital in the southwestern United States. The hospital is a nonprofit, publicly funded institution providing services to the adult community; pediatric services are not provided. The purpose of the study was to:

1. Identify the determinants of patient quality judgments.
2. Identify internal service delivery processes that impacted patient quality judgments.
3. Determine the linkage between patient quality judgments and intent-to-patronize the hospital in the future or to recommend the hospital to others.

To conduct the study, the author worked closely with a core team of hospital employees, and with several ad hoc teams of hospital employees. The core team included the Nursing Administrator, the head of the Quality Management Department, and the head of Nutrition Services.

The team decided to develop their criteria independently. It was agreed that the best method of getting information was directly from the target group, in-patients. Due to the nature of hospital care services, focus groups were not deemed feasible for this study. Frequently, patients must spend a considerable period of time convalescing after being released from a hospital, making it impossible for them to participate in a focus group soon after discharge. While the patients are in the hospital, they are usually too sick to participate. Some patients have communicable diseases, which makes their participation in focus groups inadvisable.

Since memories of events tend to fade quickly (Flanagan, 1954), the team decided that patients should be interviewed within 72 hours of discharge. The target patient population was, therefore, all adults treated as in-patients and discharged to their homes. The following groups were not part of the study: families of patients who died while in the hospital, patients discharged to nursing homes, patients admitted for psychiatric care.

The team used the Critical Incident Technique (CIT) to obtain patient comments. The CIT was first used to study procedures for selection and classification of pilot candidates in World War II (Flanagan, 1954). A bibliography assembled in 1980 listed over seven hundred studies about or using the CIT (Fivars, 1980). Given its popularity, it is not surprising that the CIT has also been used to evaluate service quality.

CIT consists of a set of specifically defined procedures for collecting observations of human behavior in such a way as to make them useful in addressing practical problems. Its strength lies in carefully structured data collection and data classification procedures that produce detailed information not available through other research methods. The technique, using either direct observation or recalled information collected via interviews, enables researchers to gather firsthand patient-perspective information. This kind of self-report preserves the richness of detail and the authenticity of personal experience of those closest to the activity being studied. Researchers have concluded that the CIT produces information that is both reliable and valid.

This study attempted to follow closely the five steps described by Flanagan as crucial to the CIT: (1) establishment of the general aim of the activity studied; (2) development of a plan for observers or interviewers; (3) collection of data; (4) analysis (classification) of data; and (5) interpretation of data.

**Establishment of the general aim of the activity studied**

The general aim is the purpose of the activity. In this case the activity involves the whole range of services provided to in-patients in the hospital. This includes every service activity between admission and discharge. From the service provider’s perspective the general aim is to create and manage service delivery processes in such a way as to produce a willingness by the patient to utilize the provider’s services in the future. To do this, the service provider must know which particular aspects of the service are remembered by the patient.

Our general aim was to provide the service provider with information on what patients remembered about their hospital stay, both pleasant and unpleasant. This information was to be used to construct a new patient survey instrument that would be sent to recently discharged patients on a periodic basis. The information obtained would be used by the managers of the various service processes as feedback on their performance, from the patient’s perspective.
Interview plan
Interviewers were provided with a list of patients discharged within the past 3 days. The discharge list included all patients. Nonpsychiatric patients who were discharged to “home” were candidates for the interview. Home was defined as any location other than the morgue or a nursing home. Interviewers were instructed to read a set of predetermined statements. Patients to be called were selected at random from the discharge list. If a patient could not be reached, the interviewer would try again later in the day. One interview form was prepared per patient. To avoid bias, 50% of the interview forms asked the patient to recall unpleasant incidents first and 50% asked for pleasant incidents first. Interviewers were instructed to record the patient responses using the patient’s own words.

Collection of data
Four interviewers participated in the data collection activity; all were management level employees of the hospital. Three of the interviewers were female, one was male. The interviews were conducted when time permitted during the interviewer’s normal busy work day. The interviews took place during the September 1993 time period. Interviewers were given the instructions recommended by Hayes (1992) for generating critical incidents.

A total of 36 telephone attempts were made and 23 patients were reached. Of those reached, three spoke only Spanish. In the case of one of the Spanish-speaking patients a family member was interviewed. Thus, 21 interviews were conducted, which is slightly greater than the 10 to 20 interviews recommended by Hayes (1992). The 21 interviews produced 93 critical incidents.

Classification of data
The Incident Classification System required by CIT is a rigorous, carefully designed procedure with the end goal being to make the data useful to the problem at hand while sacrificing as little detail as possible (Flanagan, 1954). There are three issues in doing so: (1) identification of a general framework of reference that will account for all incidents; (2) inductive development of major area and sub-area categories that will be useful in sorting the incidents; and (3) selection of the most appropriate level of specificity for reporting the data.

The critical incidents were classified as follows:

1. Each critical incident was written on a 3 x 5 card, using the patient’s own words.
2. The cards were thoroughly shuffled.
3. Ten percent of the cards (10 cards) were selected at random, removed from the deck and set aside.
4. Two of the four team members left the room while the other two grouped the remaining 83 cards and named the categories.
5. The ten cards originally set aside were placed into the categories found in step 4.
6. Finally, the two members not involved in the initial classification were told the names of the categories. They then took the reshuffled 93 cards and placed them into the previously determined categories.

The above process produced the following dimensions of critical incidents:

- Accommodations (5 critical incidents)
- Quality of physician (14 critical incidents)
- Care provided by staff (20 critical incidents)
- Food (26 critical incidents)
- Discharge process (1 critical incident)
• Attitude of staff (16 critical incidents)
• General (11 critical incidents)

**Interpretation of data**
Interjudge agreement, the percentage of critical incidents placed in the same category by both groups of judges, was 93.5%. This is well above the 80% cutoff value recommended by experts. The setting aside of a random sample and trying to place them in established categories is designed to test the comprehensiveness of the categories. If any of the withheld items were not classifiable it would be an indication that the categories do not adequately span the patient satisfaction space. However, the team experienced no problem in placing the withheld critical incidents into the categories.

Ideally, a critical incident has two characteristics: (1) it is specific and (2) it describes the service provider in behavioral terms or the service product with specific adjectives (Hayes, 1992). Upon reviewing the critical incidents in the General category, the team determined that these items failed to have one or both of these characteristics. Thus, the 11 critical incidents in the General category were dropped. The team also decided to merge the two categories “Care provided by staff” and “Attitude of staff” into the single category “Quality of staff care.” Thus, the final result was a five dimension model of patient satisfaction judgments: Food, Quality of physician, Quality of staff care, Accommodations, and Discharge process.

A rather obvious omission in the above list is billing. This occurred because the patients had not yet received their bill within the 72 hour time frame. However, the patient’s bill was explained to the patient prior to discharge. This item is included in the Discharge process dimension. The team discussed the billing issue and it was determined that billing complaints do arise after the bills are sent, suggesting that billing probably is a satisfaction dimension. However, the team decided not to include billing as a survey dimension because (1) the time lag was so long that waiting until bills had been received would significantly reduce the ability of the patient to recall the details of their stay; (2) fear that the patient’s judgments would be overwhelmed by the recent receipt of the bill; and (3) a system already existed for identifying patient billing issues and adjusting the billing process accordingly.

**Survey item development**
As stated earlier, the general aim was to provide the service provider with information on what patients remembered about their hospital stay, both pleasant and unpleasant. This information was then to be used to construct a new patient survey instrument that would be sent to recently discharged patients on a periodic basis. The information obtained would be used by the managers of the various service processes as feedback on their performance, from the patient’s perspective.

The core team believed that accomplishing these goals required that the managers of key service processes be actively involved in the creation of the survey instrument. Thus, ad hoc teams were formed to develop survey items for each of the dimensions determined by the critical incident study. The teams were given brief instruction by the author in the characteristics of good survey items. Teams were required to develop items that, in the opinion of the core team, met five criteria: (1) relevance to the dimension being measured; (2) concise; (3) unambiguous; (4) one thought per item; and (5) no double negatives. Teams were also shown the specific patient comments that were used as the basis for the categories and informed that these comments could be used as the basis for developing survey items.

Writing items for the questionnaire can be difficult. The process of developing the survey items involved an average of three meetings per dimension, with each meeting lasting approximately two hours. Ad hoc teams ranged in size from four to eleven members. The process was often quite tedious, with considerable debate over the precise wording of each item.