Mystery shopping: A tool to develop insight into customer service provision

Ton Van der Wiele, Martin Hesselink & Jos Van Iwaarden

To cite this article: Ton Van der Wiele, Martin Hesselink & Jos Van Iwaarden (2005) Mystery shopping: A tool to develop insight into customer service provision, Total Quality Management & Business Excellence, 16:4, 529-541, DOI: 10.1080/14783360500078433

To link to this article: http://dx.doi.org/10.1080/14783360500078433

Published online: 24 Jan 2007.

Article views: 755

View related articles

Citing articles: 22
Mystery Shopping: A Tool to Develop Insight into Customer Service Provision

TON VAN DER WIELE, MARTIN HESSELINK & JOS VAN IWAARDEN

Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands

ABSTRACT Reaching service excellence through a focus on the customer, demands more than just measuring customer satisfaction by means of questionnaire surveys. Nowadays, it is not sufficient anymore to have service excellence in services, processes and relationships. Now is the time to create excellence in customer experiences, as the only way to create competitive advantage in the market. Organizational change should then be driven by a broader focus on customers’ expectations and multiple ways of measuring customers’ satisfaction. The case of a service company in The Netherlands (a temporary employment agency) illustrates the way to service excellence as an organizational change process. The case supports the need for a broad focus on measurements in order to be able to monitor and to direct organizational changes. Customer satisfaction data based on surveys are needed; however, they will not be sufficient for continuing the change process over time. To achieve that, other measurements, such as mystery shopping, may give more stimuli to change. Therefore, this research argues that mystery shopping can be a useful instrument in addition to the more often-used survey methods.

KEY WORDS: Service excellence, customer satisfaction, mystery shopping, mystery guest

Introduction

Quality thinking has evolved over time from product quality, via process quality and systems quality, towards organisational quality (Dale, 2003) and the quality of relationships with all stakeholder parties. The current trend from organizational quality towards quality of relationships is a move from inward looking (development of internal quality management systems) to focusing on stakeholders outside the organization. At the same time there is a trend to pay more attention to customers for two reasons. First, services become more important in developed societies and the customers’ role is more prominent in services. Secondly, strong customer relationships are becoming the only way to compete in many business sectors.
The latest ISO 9000:2000 standard includes a drive towards more customer orientation (ISO 9000, 2000). However, in practice, it means more often a kind of basic survey to gather some customer feedback data through satisfaction scores that are seldom used for organizational development or improvement. From the perspective of the quality award models or business excellence models there might be more attention to the overall process of why customer satisfaction measures are in place, how they are executed in relation to other activities and how they are used to develop organizational improvement (EFQM, 2004; NIST, 2004). A far more challenging focus is to go beyond customer satisfaction and develop excellent customer experiences according to the ideas described by Pine & Gilmore (1999).

Customer Experience

‘Building a customer experience which genuinely creates loyalty is the newest and hardest battleground business has faced. This is the most important debate in business today’ (Shepherd, 2004). Pine & Gilmore (1999) have stimulated the discussion on customer satisfaction, customer focus and customer experiences by means of their book The Experience Economy. Great customer experiences are built on seven philosophies, according to Shaw & Ivens (2002). These seven philosophies are easy to understand and indicate what you need to look for. They are:

- Recognize great customer experiences as a source of long-term competitive advantage.
- Great customer experiences are created by consistently exceeding customers’ physical and emotional expectations.
- Great customer experiences are differentiated by focusing on stimulating planned emotions.
- Great customer experiences are enabled through inspirational leadership, an empowering culture and empathetic people who are happy and fulfilled.
- Great customer experiences are designed outside in rather than inside out.
- Great customer experiences are revenue generating and can significantly reduce costs.
- Great customer experiences are an embodiment of the brand.

These seven philosophies for building great customer experiences become a virtuous circle, reinforcing them over and over again.

Other authors also stress the important role of customer satisfaction in order to be able to turn them into advocates for the company (Smith & Wheeler, 2002). They also indicate that customers are looking for service experiences that complement their lifestyle, and that they are willing to pay for that. The cost of the coffee in a €3.50 cappuccino at Starbucks may only be 20 cents. Apparently, customers are willing to pay for the experience. The economic value is moving from commodities, via goods and services, towards experiences. This implies that companies have to look for more differentiation in their markets and have to satisfy the demands of each individual customer. The only way to create positive customer experiences is to have a perfect balance and integration between processes (in order to produce the service and the product), settings (the environment in which the experience is performed) and employees (the people who
have to create the relationships with the customers). The integration and balance of these three inputs have to be aligned with the service standards or service excellence goals (Disney Institute, 2001).

One of the aspects Shaw & Ivens (2002) mention about creating great customer experiences is the important role of leadership. This role is also emphasized by Church (1995), who states that leadership behaviour is significantly related to service performance and organizational performance. In particular, empowerment has a large influence on the satisfaction rating of customers (Church, 1995). This finding is supported by House (1992) and Yukl (2002). These authors argue that subordinates will perform better towards the customer, when they are inspired by charismatic leaders who are able to convey their own strong motivation, enthusiasm and commitment.

The Case of a Dutch Flex Company

In 1977, the Dutch government, the existing labour unions and employers-associations set up a new temporary employment agency. The mission of this organization was to help those people who were experiencing difficulties in finding a suitable job. Today, the administrative and cultural heritage of this agency, based on feelings of social responsibility in terms of legitimacy and fairness, is deeply rooted in the organization. This is the case, even though there was a turnaround five years ago towards a more market-driven orientation and a focus on efficiency and effectiveness.

In the early days, the core business of the flex company was acting as an intermediary for flex workers. Nowadays, the organization offers a full-service package of products, such as recruitment and selection, long-term detachments, outplacement, and reintegration. Recruitment and development are the company’s central issues in its role as an intermediary in temporary work contracts and its role of employee development centre. The flex company has a long-term vision to add value to its two customer groups: (1) to the paying customers (the clients), who are the people who hire the flex workers, by supporting them to find capable flex workers and (2) to the flex workers (the candidates), who are the people who are looking for temporary work, by giving them the opportunity to get a suitable job.

Because competition on price or new product innovations is nearly impossible, the flex company made the decision to excel in service towards candidates and clients, and thus to implement a service culture and behaviour. The flex company started by looking for a benchmark organization in order to determine the nature and essence of excellent service. In this way, the rather abstract concept ‘service’ could be made more tangible for the organization. Worldwide, the Walt Disney Company is considered by many writers (e.g. Capodagli & Jackson, 1998) to be the world leader in the field of service management. The flex company decided to involve the Disney Institute in its transformation process and to ask for help with the training of its top managers. The first step was to formulate a new vision and mission-statement for the organization, which had to be clear and unambiguous for all the employees of the flex company. According to Capodagli & Jackson (1998), an important condition for achieving service excellence is to have a clear organizational mission statement, which must be accepted by every member of ‘the cast’ (a typical Disney label for its employees). Together with the development of the mission-statement, other company standards and statements were developed, for example the four Service Standards (Accessibility, Reliability, Service Provision and
Efficiency). These Service Standards are the building blocks of the framework of the flex company’s service excellence programme.

The first step in the service excellence programme was leadership training for top managers at the Disney Institute in Florida. During this course, the managers were confronted with service excellence in practice. This ‘see-with-your-own-eyes’ experience is extremely important in creating trust in the programme. At Disney World, the managers experienced service excellence and its benefits, and how to gain a service excellence status. The leadership training for all top managers is consistent with the nature of the change programme, which is based on a change in leadership style (Levy & Merry, 1986, p. 58). Consequently, the managers have the most important role in the organization during the implementation. They are the persons who have to transfer and translate the ideas and implications of the programme and of the concept of service excellence to their employees. Commitment of management is a necessary condition for a quality-driven programme, such as the service excellence programme (see also Capodagiali & Jackson, 1998; Prabhu & Robson, 2000; Maister, 2001). It is argued that when top management does not give priority to the implementation of the programme, there is no reason for lower level employees to be committed to the ideals.

The chain of excellence, which is shown in Figure 1, can be summarized as follows:

- Excellent leadership will be achieved by means of clear vision and mission statements, and the distinct role model behaviour of managers who are proud of their organization and know what is expected of them.
- Excellent employees will be developed by means of excellent leadership, recruitment and selection procedures, excellent education and training facilities, and clear guidelines about what is expected from them.
- Satisfied flex workers will be created by means of excellent employees who see every flex worker as a VIP (very individual person), for whom every little detail is covered and for whom the employees carefully assess what the best service is in each specific situation.
- Satisfied customers will be brought about in a service experience that is excellent, by means of the actions of satisfied flex workers and excellent employees who treat their customers as individuals for whom every little detail is covered and for whom the employees carefully assess what the best service would be in each specific situation.
- Satisfied customers who have become loyal will spread the word to others and will build long-term relationships and thereby ensure excellent financial results (Heskett et al., 1992).

![Figure 1. Chain of excellence](image-url)
Although the supposed relationships between the separate blocks in the chain of excellence seem straightforward, a lot of research has been undertaken during the last decades to prove those relationships. The relationship between, on the one hand, the first two blocks, which can be put in the field of human resource management, and, on the other hand, the last block (performance measurements) has been the subject of considerable study (e.g. Arthur, 1994; Huselid, 1995; Guest, 1997; Youndt et al., 1996). Although a positive relationship between these two concepts or particular parts of these concepts is presumed to exist, a direct linear relationship between human resource activities and operational performance is hard to find. Nevertheless, the managers at the flex company believe in their role as a leader and make sure that everything will be done to confirm a positive HRM-performance relationship. Customer satisfaction and performance has also been the subject of considerable examination (Zeithaml, 2000). This relationship has the same characteristics as the HRM-performance relationship. Initially, the connection between the two concepts appears to be clear, but empirical research comes up with mixed results.

Customer Satisfaction Measurements

Customer focus has always been one of the basic elements of quality management. Increased attention for quality management during the last two decades has brought about higher priority for the importance of customer satisfaction. Many authors recognize the importance of customer satisfaction (e.g. Dale, 2003; Zeithaml et al., 1990; Heskett et al., 1990; Horovitz & Jurgens-Panak, 1992). Most authors agree on the various reasons for the importance of customer satisfaction. The reasons are:

- Satisfied customers buy more and become loyal.
- Satisfied customers buy additional products and services.
- Satisfied customers support positive word of mouth.
- Satisfied customers pay less attention to advertisements of competitors.
- Satisfied customers deliver more ideas and suggestions.
- Satisfied customers cost less because these customers already have a relation with the company and some processes might already have been standardized.

Having a strong customer focus has become very important in today’s competitive environment. Measuring customer satisfaction by means of surveys is frequently used and is also stimulated by the latest version of the ISO 9000 quality systems standard (ISO 9000, 2000). However, there are also completely different ways of measuring the quality of service delivery. Organizations can, for example, measure the quality of service delivery by making use of mystery guests, which are well-trained people who behave as normal customers but who are accurately observing what is going well and what can be improved in the service process as perceived by them.

The questionnaire survey type of measurement focuses mainly on the mean values in relation to customer satisfaction. Changes over time in the mean values are quite often either minimal or cannot be explained by what the organization is able to manage. Therefore, it seems very useful to link other measurements to these surveys, for example measurements of the extremes (good and bad service quality) through critical incident analyses or through mystery shopping (Wilson, 1998a).
A lot of different measurements exist in the flex company. To determine if the relevant items are being measured according to the chain of excellence, the flex company is developing a service monitor in cooperation with Erasmus University Rotterdam. The objective of this monitor is to provide the flex company periodically with an assessment of the current level of service provision. The research team is also examining the presupposed relationships in the chain (Van der Wiele et al., 2002; Hesselink et al., 2002) and is analysing available data from service excellence research in a more thorough way (Boselie et al., 2001). One goal is to get an organizational dashboard from which management can easily derive the performance of its divisions, regions and branches. Another goal is to create a balanced and integrated overview of performance indicators that make clear what the most important issues in the change process are.

Satisfaction measurements that are in place at the flex company cover:

- Satisfaction of employees with the human resource management practices, with leadership, with the culture etc (by means of a general employee satisfaction survey, which is done every two years).
- Satisfaction of the candidates (flex workers) with the way they are supported by the organization and the extent to which they are satisfied with the jobs offered. The items in the questionnaire are grouped into two factors: (1) aspects of matching quality and (2) aspects of service quality. This survey is done on a continuous basis. Matching quality deals with the actual match of a person to a job (the ‘what’ or outcome quality). Service quality provides insight into the way the match (or service) is delivered by the flex company (the ‘how’ or relational quality).
- Satisfaction of clients (i.e. the organizations that hire flex workers). This measurement also shows a grouping of the items of the questionnaire around matching and service quality. This survey is also done on a continuous basis.

Mystery Shopping

Mystery shopping can be used for various purposes. Most of the time the goal is to measure the quality of the service delivery to the customer. Mystery shopping can also be used to benchmark by sending mystery guests to branch offices of competitors. Other goals of mystery shopping are:

- Measuring effectiveness of (training) programmes (Morrison et al., 1997).
- Testing if customers are treated equally (e.g. testing against discrimination) (Morrall, 1994; Tepper, 1994).

An important advantage of mystery shopping is the quality of the measurement. Mystery guests are well trained and know the processes and are therefore able to measure the critical failure points. The first step in the design of mystery shopping is to define goals. These goals have to be made transparent and used as the input for the checklist. This checklist has to be developed by going through the process of the service delivery and by paying attention to failure points. In most cases it is better to have multiple visits done at the same location over a period of time; these visits should take place at different
times of the day (Leeds, 1992). This will reduce the effect of special situations and will also provide means of measuring the effects of improvements made after earlier visits. The second step in the design of mystery shopping is data gathering. The gathered data should not only cover general applicable service quality dimensions (Zeithaml et al., 1990) but should also reflect the key performance indicators defined by the organization in relation to its vision and mission. The mystery shoppers that gather the data should be independent, critical, objective and anonymous. Mystery shopping has to lead to a reliable picture of the normal way of working in a branch office. Although the mystery shopping instrument has to be presented to employees as a way to improve services, individual visits should not be announced beforehand to avoid abnormal behaviour by the employees. The final step in the design of mystery shopping is reporting the results. Cramp (1994) concludes that there is a shift in the type of information that is reported in relation to a mystery shopping. Traditionally, reporting was based on subjective information and open questions. More recently, reporting also covers objective measurements based on checklists. The reports should be presented to responsible managers as soon as possible after the visit.

Various papers have been published on mystery shopping. The research in this area that has been published can be categorized in four clusters:

- Tourism: mystery shopping is a useful instrument in service environments like hotels and in the travel and tourism industry (Erstad, 1998; Wilson & Gutmann, 1998; Anderson et al., 2001).
- The relationship between mystery shopping results and rewarding: examples of the link of the results of mystery shopping with a bonus system, and or promotion system (Eisman, 1993; Boyd, 1995; McNerney, 1996).
- Marketing: a number of publications can be found on the prerequisites and conditions for mystery shopping, and the rules to be followed when implementing mystery shopping (Miles, 1993; Burnside, 1994; Cramp, 1994; Cobb, 1995; Dwek, 1996).

The focus of the academic research papers is strongly related to aspects of reliability. Exploratory research has been undertaken by Wilson (1998a, 1998b) and Wilson & Gutmann (1998) in the UK on the role of mystery shopping in the measurement and management of the service delivery process, and leading to conclusions in relation to the reliability of the technique. Morrison et al. (1997) also studied the reliability of mystery shopping results based on cognitive psychology. Finn (2001) describes research in retail chains focusing on the quality of mystery shopping data and the number of visits needed.

In the Dutch flex company, mystery shopping has been introduced in parallel to survey measurements. Mystery shopping is used in the flex company to evaluate the service provider on dimensions identified as important to potential candidates (job seekers), rather than dimensions identified as important to clients (job providers). The customer surveys show the perceptions of customers, while mystery shopping makes it possible
to look at the execution of processes from the point of the intentions of the organization (e.g. strategy, service standards as defined by the flex company). During 2003, 237 mystery shopping visits were executed. This large number of mystery shopping visits exceeds the minimum number of 150 that is necessary to rank different branches that belong to the same chain organization (Finn, 2001). The visits have been executed in multiple measurement rounds. Some branches have been visited once during the year, while others received up to six visits. The number of visits to a branch is the outcome of a random process instead of revisiting only poorly performing branches. One of the instruments the organization uses to make the results of these visits measurable and as objective as possible, is a multi-item checklist, which has to be filled out by the mystery shopper after the visit. In Table 1 some examples of the used items are presented.

The overall score of a branch-office is calculated by multiplying the weighted averages of the categories with an importance factor. The categories, the number of underlying items and the accompanying importance factors are shown in Table 2. The importance factors are used to emphasize the categories reflecting interpersonal contact between the potential candidate and the employee of the flex company. These categories have a higher importance factor than the categories about the physical properties of the branch, the search phase, and the way of making an appointment. However, this does not mean that these last mentioned categories are not at all important.

The follow-up of the mystery shopping visits is threefold:

- There is regular communication towards the specific branch and its responsible managers. They have to pick up all the critical points in the report and make improvements where possible.
- The scores of each visit are published on the Intranet of the flex company in order to show the whole organization which branches have high scores and which have low scores. After a few months the branches with the highest scores are rewarded with a dinner. This way of reporting creates competition between the various branches, which stimulates the employees of a branch to work as a team in order to reach the top positions. This team spirit creates a strong focus on delivering service excellence.
- At the end of the year, the actual results are compared with the objectives formulated at the beginning of the year. The scores of the visits during that year are part of the official reward and recognition system of the company.

<table>
<thead>
<tr>
<th>Question no.</th>
<th>Question</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>The opening hours are clearly visible</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>64</td>
<td>The employees are friendly</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>83</td>
<td>The employees show they have a lot of knowledge</td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>
Mystery shopping is a useful instrument to create an in-depth insight into perceptions of potential customers. It adds value to customer satisfaction survey data. Mystery shopping can well be used as an instrument to gather qualitative as well as quantitative information. Mystery shopping should be used in an open and transparent way. By communicating the use of mystery guests through the whole organization, it already gives a signal and stimulus to pay more attention to the perceptions of real customers. Good communication of the results of mystery shopping can also create positive stimuli for improvements. In relation to that, the role of the managers in the organization is very important. They are responsible for the communication of the results of the visits. They also have to coordinate the actions for improvement and have to show role modelling behaviour in taking actions.

Quantitative Results of the Mystery Shopping

In Table 3 and Figure 2 a summary of the results of the 237 mystery shoppings is presented. Table 3 shows all six measurements; however, most branches have been visited only once \( N = 167 \), twice \( N = 41 \), or three times \( N = 15 \). Figure 2 compares the means of visit one, two and three (visits four, five and six have been omitted because of their low numbers). The average scores for each of the seven categories vary between 3.33 and 4.60 on a five-point scale. The appointment phase is the category that scores highest, while the intake category scores lowest.

To assess the usefulness of mystery shopping as a method to improve the performance of a service organization, the results of follow-up measurements should be compared to the results of earlier measurements. From the database of 237 visits, 41 pairs of two sequential visits of the same branch could be selected. Based on these 41 pairs, a Wilcoxon non-parametric signed ranks test has been performed. The results of this test are shown in Table 4. From this table, it can be concluded that three out of the seven categories have improved significantly \( P < 0.05 \): branch outside, branch inside, and first contact. The improvements of these three categories are also meaningful, based on their Z-values, which indicate that the scores of the follow-up measurements are considerably higher than the scores of the earlier measurements.

The meaning of these results is twofold. On the one hand, the mystery shopping instrument can be used to stimulate improvements and to measure these improvements within a short time frame. On the other hand, the results indicate that mystery shopping stimulates improvements that are easy to achieve. Examples of this kind of improvement are related
### Table 3. Mystery shopping results

<table>
<thead>
<tr>
<th>N</th>
<th>1st measurement</th>
<th>2nd measurement</th>
<th>3rd measurement</th>
<th>4th measurement</th>
<th>5th measurement</th>
<th>6th measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>167</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>41</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>15</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>7</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>5</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>2</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
</tbody>
</table>
to the looks of the branch office (e.g. the building is clean; recent newspapers and magazines in the waiting room; coffee and tea facilities) and the way customers are addressed during their first contact with the branch (e.g. waiting time not too long; friendly employee; provision of good information about the procedures of the company).

At the moment it is not yet clear whether mystery shopping is able to stimulate more profound improvements over the longer term. The more difficult improvements that are related to service delivery and image (e.g. the ability of employees to show empathy to customers; high knowledge level of employees; trust in the company; extent to which customers feel that they are important) might have been stimulated by mystery shopping, however they probably take much longer to show up in the results of follow-up measurements. Within the one-year time frame of this research these improvements have therefore not been found.

Table 4. Results of the Wilcoxon non-parametric signed ranks test

<table>
<thead>
<tr>
<th>Delta</th>
<th>Z-value</th>
<th>Sign. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search phase (first – /– last)</td>
<td>−0.82</td>
<td>0.20</td>
</tr>
<tr>
<td>Making appointment (first – /– last)</td>
<td>−0.46</td>
<td>0.32</td>
</tr>
<tr>
<td>Branch outside (first – /– last)</td>
<td>−1.69</td>
<td>0.05</td>
</tr>
<tr>
<td>Branch inside (first – /– last)</td>
<td>−1.74</td>
<td>0.04</td>
</tr>
<tr>
<td>First contact (first – /– last)</td>
<td>−1.72</td>
<td>0.04</td>
</tr>
<tr>
<td>Intake (first – /– last)</td>
<td>−0.82</td>
<td>0.21</td>
</tr>
<tr>
<td>General opinion (first – /– last)</td>
<td>−0.67</td>
<td>0.25</td>
</tr>
<tr>
<td>Weighted mean (first – /– last)</td>
<td>1.31</td>
<td>0.10</td>
</tr>
</tbody>
</table>

* The change is significant (p<0.05)
Conclusions

Satisfaction measurements are, in general, not very well linked to goals and actions within organizations. Too often, the customer satisfaction measurements are conducted because of external pressure and therefore not much is done with the results of the measurements. In the case of external pressure, quite often a standard survey questionnaire is used, which is not specific for the organization. Even if the right questions are asked, the scores do not indicate what has to be done. Over time scores might change; however, the changes are most often correlated with general exogenous developments instead of a direct consequence of improvement actions within the organization.

Customer satisfaction measurements by means of questionnaires are frequently used; however, a broader perspective should be developed to give real meaning to the concept of customer focus. Mystery shopping is seen as an efficient and effective instrument to gain more in-depth knowledge of the customers’ perception of service delivery (Finn, 2001). The research in this paper has shown that the mystery shopping instrument is useful for improving service delivery.

Besides mystery shopping, there are also other ways of developing content-rich customer relationships, for example:

- Small group interviews and discussions.
- Segmenting customers into target groups: potential customers, current customers, and exit interviews with customers who did not come back after a certain period of time.
- Bringing customers into your organization
- Stimulating contacts between employees and customers.

The richness is in the combination of different measurements and in the actions that will follow after the measurement. It is an important task for the managers of an organization to ensure that multiple measurements are combined and to ensure that the follow-up actions will really take place. It is not the measurements that make the difference but the improvement actions.

References


Burnside, A. (1994) In-store spies sniff out poor service, Marketing, 28 April, pp. 32–33.


