Mystery Shopping:
In-depth measurement of customer satisfaction

Martijn Hesselink, Ton van der Wiele

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	This paper will discuss the phenomenon Mystery Shopping in the field of customer satisfaction measurement techniques. By using the literature about Mystery Shopping definitions and restrictions of this instrument will be presented. Also, possible ways to present and use the gathered data will be shown. After the literature part of the paper some practical research will be presented. A Dutch Flexcompany introduced the instrument Mystery Shopping in addition to the already used measurement methods like customer satisfaction measurement with use of questionnaires. Some of the first results of the Mystery Shopping visits will be presented.	
Library of Congress Classification (LCC)	5001-6182	Business
	5546-5548.6	Office Organization and Management
	HF5415.5	Customer services, Marketing
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 10	Business Administration: general
	L 2	Firm Objectives, Organization and Behaviour
	M 31	Marketing
European Business Schools Library Group (EBSLG)	85 A	Business General
	100B	Organization Theory (general)
	240 B	Information Systems Management
	280 J	Marketing audit
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.05	Management organisatie: algemeen
	85.08	Organisatiesociologie, organisatiepsychologie
	85.40	Marketing
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Organisatieleer, informatietechnologie, prestatiebeoordeling	
	Klantgerichtheid, dienstverlening, prestatiebeoordeling, anonimiteit	
Free keywords	Customer Satisfaction, Mystery Guest, Service Quality	

Mystery Shopping: In-depth measurement of customer satisfaction

drs Martijn Hesselink
Erasmus University Rotterdam
Rotterdam School of Economics (Room H15-10)
PO Box 1738, NL 3000 DR Rotterdam, The Netherlands
Tel. int. +31 10 4081354
Fax int. +31 10 4089169
Email: hesselink@few.eur.nl

dr ir Ton van der Wiele
Erasmus University Rotterdam
Rotterdam School of Economics (Room H15-28)
PO Box 1738, NL 3000 DR Rotterdam, The Netherlands
Tel. int. +31 10 4081354
Fax int. +31 10 4089169
Email: vanderwiele@few.eur.nl
Personal page:
http://www.few.eur.nl/few/people/vanderwiele

Martijn Hesselink is a PhD-student at the Erasmus University In Rotterdam, The Netherlands. His research topic is about the influence of leadership on the service behaviour of employees, using networktheory. He also works since five years as a quality consultant at a flexcompany.

Ton van der Wiele is Associate Professor in the Department of Business and Management at the Rotterdam School of Economics.

Abstract
This paper will discuss the phenomenon Mystery Shopping in the field of customer satisfaction measurement techniques. By using the literature about Mystery Shopping definitions and restrictions of this instrument will be presented. Also, possible ways to present and use the gathered data will be shown. After the literature part of the paper some practical research will be presented. A Dutch Flexcompany introduced the instrument Mystery Shopping in addition to the already used measurement methods like customer satisfaction measurement with use of questionnaires. Some of the first results of the Mystery Shopping visits will be presented.

1. Introduction
A customer focus has become more and more important in today's competitive environment. Measuring customer satisfaction is a tool frequently used and also stimulated by the most recent version of the quality systems requirements according to the ISO 9000 series quality systems standard (ISO, 2000). Another way of measuring the quality of the service delivery is by making use of mystery guests, well-trained persons who behave as normal customers but are precisely observing what is going well and what can be improved in the service process as perceived by the customer.

This paper will discuss the conditions for mystery shopping and the advantages. A case study of mystery shopping within the context of a flex company is presented and results are shown.

2. Mystery shopping
Mystery shopping is very useful in service situations, for example in banks, restaurants, shops, travel agencies, airlines, car dealers etc. Examples of mystery shopping in the literature can most of the time be related to these type of organizations (Wilson,1998a)

The way mystery shopping is implemented is rather general:

* Anonymous persons are used as mystery guests

* The mystery guest is well trained and uses a predefined checklist

* The mystery guest behaves like a normal customer

* After the visit to the service point the checklist is filled in and an evaluation report is documented

* The evaluation report and the scores on the checklist are used for feedback

* The feedback has to lead to a number of improvement issues and actions

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. In this situation the mystery guest can be focused on the compliance to specific standards, guidelines or demands, or the mystery guest can be instructed to position the quality of the service on a scale. If a mystery guest is also used to visit not only the own service locations but also locations of competitors, benchmarking becomes a way to judge your own activities against those. Other goals of mystery shopping are:

* Measuring effectiveness of (training) programmes (Morrison et al, 1997)

* Testing if employees treat all people in the same way, or testing against discrimination (Morral, 1994; Tepper, 1994)

The results of mystery shopping can be used in various ways:

* Communication: the checklists make very clear what is expected

* Reports make very clear what is going wrong

* Training programmes can be improved

* Improvement activities can be focused on issues that are important for the customer

* Good behavior of employees can be rewarded

An important advantage of mystery shopping is the value of the measurement. Mystery guests are well trained and know the processes and the failure points and are thus able to measure the details. Mystery guests role is clearly to report these details and are supported by the detailed checklist they most of the time carry with them. This all will make the results of mystery shopping worthwhile. We all know that customers do not always notice defects in the processes, or even if they notice, they do not always complaint or make their dissatisfaction visible. Research of TARP (Technical Assistance Research Program) shows:

- 26 out of 27 dissatisfied customers of goods with a relative low price, don't tell their supplier about their dissatisfaction, however, 63% will never buy again

- 45% of customers who experience a problem with a relative low cost service, don't complaint; 45% of those will never come back

- 27% of the buyers of expensive products will not complaint, but 41% will not come back

- 37% of dissatisfied customers of expensive services do not complaint, however, 50% of the dissatisfied customers will not come back.

Customer satisfaction is important:

* 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 have already a relation with the company and some processes might already have been standardized.

3. Basics of Mystery Shopping
Mystery shopping has to be designed very well with specific attention to the following aspects: design of the process (goal, frequency and feedback), data gathering, mystery shoppers, and reporting.

Design of the process

The first step in the design is to define the goals for mystery shopping. The goals have to be made transparent and be used as the input for the checklist that will be used to measure against those goals. The checklist has to be developed for example by going through the process of the service delivery and paying attention to failure points in those processes and complaints that have been made in earlier customer contacts and surveys. The specific issues in relation to the quality of the service have to be formulated by the service organizations and should be related to the strategic positioning of the organization in the marketplace.

In most cases it is better to have more visits done by the mystery shopper at a location over a period of time, and also visits have to take place at different working hours (Leeds, 1992). This will reduce the effect of special situations, and will also give a way of measuring the effects of improvement that have been made after the results of the first visit of the mystery shopper.

Data gathering

The data gathering has to be linked to the specific situation and to the strategy of the organization. The checklist should cover not only the 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.

Mystery shoppers

Mystery shoppers should be independent, critical, objective and anonymous. The mystery shopping has to lead to a reliable picture of the normal way of working in the specific shop-location. Although mystery shopping has to be presented to employees as a way to improve services, there should not be a warning or other announcement to the location about the day and time the mystery shopping will take place.

Mystery shoppers need to be trained in relation to the instrument, but also in relation to the specific situation and context of the organization where the visits will take place.

Reporting

Cramp (1994) concludes that there is a shift in the type of information that is reported I relation to a mystery shopping. In first instance the focus was on subjective information based on open questions. The last ten years it has become more accepted to use checklists with objective statements. The best approach seems to be building on a combination of subjective and objective information.

In general, there is a need to feedback the results of the mystery shopping to the location management as soon as possible after the visit.

4. Literature on Mystery Shopping
A focused review of the literature shows that in selected journals in the areas of Quality Management (Quality Progress, TQM Magazine, and Sigma), of Marketing (Journal of Marketing, Tijdschrift voor Marketing, and European Journal of Marketing), and of General Economics (Harvard Business Review, California Management Review, and Holland Management Review) over a period of five years (between 1993 and 1997) only a very few research papers have been published on mystery shopping. A search on electronic databases gives some more results. These publications can be clustered in the following way:

- Banking: various publications are available on the implementation of mystery shopping in Banks (Holliday, 1994; Morral, 1994; Dorman, 1994; Hotchkiss, 1995; Hoffman, 1993; Stoval, 1993; Leeds, 1992 and 1995; Hanke, 1993, Tepper, 1994)

- Tourism: Mystery shopping seems to be a useful instrument also in service environments like hotels and in the travel and tourism industry (Anderson etc al, 2001; Erstad, 1998; Wilson and Gutmann, 1998)

- Relationship between mystery shopping results and rewarding: Examples of the link of the results of mystery shopping with bonus system, and or promotion system (McNerney, 1996; Eisman, 1993; Boyd, 1995)

- 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 (Cramp, 1994; Miles, 1993; Dwek, 1996; Burnside, 1994; Cobb, 1995)

The focus of the academic research papers is strongly related to aspects of reliability. Exploratory research has been undertaken by Wilson et al (1998a, 1998b, and 1998c) 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. Also Morrison et al (1997) 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.

5. Empirical Research

Organisational background

The studied organization is a service organisation at the far extreme end of the manufacturing/service continuum. It is an intermediary organisation with the aim to create the link between the person looking for a job (the flexworker) and the company that is looking for new employees (the customer). In the past the ultimate aim of flexcompanies has been to help people in a job and thereafter the role of the agency was finished. More recently, this organization has redefined its role because of developments in the employability market and in changes in the demands of the flexworkers. The new philosophy is to become a partner for flexworkers in their long-term career, and thereby being able to satisfy customers better. To strengthen this new philosophy the organization started in 2000 a 'Service Excellence Program', which is a way to create future business performances through a service quality leadership culture, focused on getting highly motivated, satisfied and quality orientated employees, who create satisfied and loyal flexworkers and customers. To create such a culture the organization was supported by the Disney Institute, which is a training division of Walt Disney Corporation. Together they developed the Chain of Excellence, which is the management model that describes the relationships between leaders, employees, flexworkers, customers and business results.

Erasmus University Rotterdam is developing a Service Monitor for the flexcompany, covering the steps in the Chain of Excellence. The objective of this monitor is to show the organisation periodically at what level the service provision is. The Erasmus research team is analysing the available data in a thorough way (Hesselink et al., 2002) and is also examining possible relationships in the Chain of Excellence (Van der Wiele et al., 2002). The main goal is to get an organisational dashboard on which the management can easily see the status of their division, region or branch.

One of the measurement techniques that are part of the Service Monitor is Mystery Shopping. This technique is a good addition to the already existing techniques like customer satisfaction surveys, assessments and complaint registrations.

Methodology for mystery shopping

The methodology used by the organisation is comparable with the way of measurement described in the first paragraph of this paper. The mystery guests behave like real customers who are looking for a suitable job. During the visit to one of the branches they have to observe everything that concerns the process. This means that they observe the attitude of the employees, their knowledge and skills, but also the cleanliness of the branch and the coffee facilities. The mystery guests gather all kind of information about flexwork and ask for a real job. In this situation it is extremely important to have a good shopping scenario so you can answer all the possible questions of the employees. Only with a good scenario it is possible to have clear answers to unexpected questions. A scenario is 'good' if it is very close to the real life situation of the mystery guest.

After the visit the mystery guests fill in the checklist, which will be described in the next paragraph, and write a short story about their findings and perceptions during the mystery shop visit. It is important that the mystery guests give their own opinion about the visit, based on their feelings and experiences. So the facts can be found in the checklist and the personal feelings of the mystery guest are written down in the 'prose-part' of the report.

Checklist

One of the instruments the company uses to make their mystery shopping visits as objective and measurable as possible is a multi-item checklist, which has to be filled in by the mystery guest after the visit. This checklist contains about one hundred single items, divided into seven different categories. Every item can be rated on either a yes/no scale or a 1-5 Likert scale. In table 1 some examples of questions are presented.

Table 1 Example of questions

Question no.	Question	Scale
36	The opening hours are clearly visible	1 2 3 4 5
64	The employees wear nametags.	no yes
83	The employee show they have a lot of knowledge	1 2 3 4 5

The mean results of the different categories are achieved 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.

Table 2 Categories and weights of the criteria

Category	N items	Importance factor
Search phase	16	1
Making appointment	11	2
Branch outside	11	1
Branch inside	26	1
First contact	15	5
Intake	13	5
General opinion	8	5

One of the major advantages of working with categories is the flexibility of changing, adding or removing underlying items from the list without actually changing the whole category. Therefore, longitudinal comparison between visits can be made, even when some questions are changed.

6. Results of the mystery visits

During a period of about six months in 2002, 23 mystery shop visits are executed; 5 different mystery guests did these visits. There is a large variation in the results of the visits. In table 3 the scores of the visits are presented, where a distinction is made between the seven categories of criteria.

Table 3 Results per category

Branch	Search phase	Making appointment	Branch outside	Branch inside	First contact	Intake	General appointment	Total (weighted)
A	2,94	4,83	4	3,16	4,82	4,83	-	4,43
B	4	4	3,7	4,13	4,64	4,17	-	4,28
C	2,67	5	4	3,83	4,64	3,83	-	4,14
D	3,56	-	3	4,14	4,36	4,17	-	4,13
E	3,91	-	3,4	4,04	4,18	4,17	-	4,07
F	2,55	4,67	4,56	4,15	3,79	3,77	4,38	4,01
G	2,81	5	4,1	3,8	3	4,83	-	3,95
H	2,69	4	3,22	2,81	3,82	4,67	-	3,89
I	3,75	-	3,8	3,77	3,93	3,69	3,88	3,82
J	2,25	5	4	3,84	3,45	4,17	-	3,81
K	2,69	4,67	3,2	3,16	3,82	4	-	3,74
L	4	-	3,75	3,09	3,62	3,18	4,25	3,66
M	2,5	4,5	2,33	3,04	4,27	2,67	-	3,36
N	4,27	-	4,1	4,11	3,82	2,33	-	3,33
O	3,91	-	3,43	2,21	2,54	3,15	4,38	3,29
P	3,31	4,17	3,9	3,16	2,73	2,83	-	3,04
Q	2,5	5	3,2	3,39	3,45	2,31	2,88	3,01
R	2,4	4	2,63	2,42	2,71	2,67	-	2,75
S	3,91	-	3,5	2,92	2,92	2,5	2,38	2,71
T	2,83	4,33	3,5	3,11	2,55	2,15	2,25	2,57
U	4,33	5	4,6	3	1	1	-	1,89
V	2,5	5	3	4,43	1	1	-	1,88

Not only between the various branches large differences can be found, also within the branches the categories show complete different scores (for example branch H, M and V).

Table 4: Variation between branches by category

1. Search phase	2.25 – 4.33
2. Making appointment	4.00 – 5.00
3. Branch outside	2.33 – 4.60
4. Branch inside	2.21 – 4.43
5. First contact	1.00 – 4.82
6. Intake	1.00 – 4.83

It seems that the tangibles (2/3/4) show a less wide variation and maybe are easier to repair; the people issues however are the most important items with the heaviest weight in the overall score and also show a wider variation. First of all it seems more difficult to have control over the employees; secondly, every customer might be different and thus perceptions of customers will always be difficult to define.

7. Follow up of mystery visits
The follow-up of Mystery Shopping is threefold:

- There is the regular communication line 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 will be published at the Intranet of the organization in order to show the whole organization, which branches have high scores and which branch low scores. After a few months the branches with the highest scores will be rewarded with a dinner or something. This way of reporting creates a sort of healthy competition between the various branches, which causes that employees of a branch wants to work as a team in order to reach the top-positions. Through this team spirit there is a strong focus on delivering and improving service excellence.

-   The scores of the visits during one year will also be part of the official reward and recognition structure at the company, which means that at the end of the year the actual results are compared with the objectives formulated at the beginning of the year.

8. Afterthoughts

Mystery Shopping is a good instrument to create an in-depth insight in perception of customers. It adds value to customer satisfaction data coming from surveys. Mystery shopping can well be used as an instrument to gather qualitative as well as quantitative information. It is also an instrument to gather objective as well as subjective data.

Mystery Shopping should be used in an open transparent way. By communicating through the whole organization the use of Mystery guests, it already gives a signal and stimulus to pay more attention to the perception of real customers. Good communication of the results of mystery shopping also can create positive stimuli for improvements.

The empirical research on mystery shopping is still scarce. Some researchers did study aspects of reliability. However, the rhetoric and the communication around the instrument might be a much more important factor in relation to the effects of Mystery Shopping in terms of stimulating employees to work on improvements and to become more customer focused.

References

Anderson D.N., Groves D.L., Lengfelder J. and Timothy D., (2001), A research approach to training: a case study of mystery guest methodology, International Journal of Contemporary Hospitality Management, 13 (2), 93-102.

Boyd M., (1995), A case for incentives, Incentive, 169 (2), 14.

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

Cobb R., (1995), Magical mystery lure, Marketing, October 26, 45.

Cramp B., (1994), Industrious espionage and Witnesses for the defence, Marketing, August 18, 17-18.

Dorman K.G., (1994), Mystery shopping resultscan shape your future, Bank Marketing 26 (8), 17-21.

Dwek R., (1996), Magic of mystery shopping, Marketing, October 17, 41-44.

Eisman R., (1993), Going undercover, Incentive, 167 (5), 70-71.

Erstad M., (1998), Mystery shopping programmes and human resource management, International Journal of Contemporary Hospitality Management, 10 (1), 34-38.

Finn A., (2001), Mystery shopper benchmarking of durable-goods chains and stores, Journal of Service Research, 3(4), 310-320.

Hanke E., (1993), Through the looking glass, Credit Union Management, August, 42, 45-46.

Hesselink M., Wiele A. van der and Boselie P., (2002), The importance of customer satisfaction in organisational transformation: a case of a Dutch temporary employment agency, ERIM Report Series Research in Management, ERS-2002-60-ORG.

Hoffman G., (1993), Customers can hone mystery shopping, Bank Marketing, 25 (8), 36-37.

Holliday K.K., (1994), Mutual funds, Bank Marketing, 26 (7), 23-31.

Hotchkiss D.A., (1995), What do your customers really think?, Bank Marketing, 27 (3), 13-20.

Leeds B., (1992), Mystery shopping offers clues to quality service, Bank Marketing, 24 (11), 24-26.

Leeds B., (1995), Mystery shopping: from novelty to necessity, Bank Marketing, 27 (6), 17-23.

McNerney D.J., (1996), The link to customer satisfaction, HR Focus, 73 (9), 1-4.

Miles L., (1993), Rise of the mystery shopper, Marketing, July 29, 19-20.

Morrall K., (1994), Mystery shopping tests service and compliance, Bank Marketing, 26 (2), 13-23.

Morrison L.J., Colman A.M. and Preston C.C., (1997), Mystery customer research: processes affecting accuracy, Journal of the Market Research Society, 39 (2), 349-361.

Stovall S.A., (1993), Keeping tabs on customer service, Bank Marketing, 25 (6), 29-33.

Tepper G.C., (1994), The merits of self-testing, Mortgage Banking, 54 (8), 76.

Wiele A. van der, Boselie P., Hesselink M., Empirical evidence for the relationship between customer satisfaction and business performance, MSQ, 12 (3), 2002, 184-193.

Wilson A.M.and Gutmann J., (1998), Public transport: the role of mystery shopping in investment decisions, Journal of the Market Research Society, 40 (4), 285-293.

Wilson A.M., (1998), The use of mystery shopping in the measurement of service delivery, Service Industries Journal, 18(3), 148-163.

Wilson A.M., (1998), The role of mystery shopping in the measurement of service performance, Managing Service Quality, 8 (6), 414-420.

Zeithaml V.A., Parasuraman A., Berry L.L., Delivering quality service; balancing customer perceptions and expectations, The Free Press, New York, 1990.

Publications in the ERIM Report Series Research∗ in Management

ERIM Research Program: "Organizing for Performance"

2003

On The Future of Co-operatives: Talking Stock, Looking Ahead
George W.J. Hendrikse and Cees P. Veerman
ERS-2003-007-ORG

Governance of Chains and Networks: A Research Agenda
George W.J. Hendrikse
ERS-2003-018-ORG

Mystery Shopping: In-depth measurement of customer satisfaction
Martijn Hesselink, Ton van der Wiele
ERS-2003-020-ORG