



## Journal of Services Marketing

Assessing the impact of mystery client traits on service evaluation

Pedro Quelhas Brito, Meena Rambocas,

### Article information:

To cite this document:

Pedro Quelhas Brito, Meena Rambocas, (2016) "Assessing the impact of mystery client traits on service evaluation", Journal of Services Marketing, Vol. 30 Issue: 4, pp.411-426, <https://doi.org/10.1108/JSM-01-2015-0004>

Permanent link to this document:

<https://doi.org/10.1108/JSM-01-2015-0004>

Downloaded on: 19 January 2018, At: 10:05 (PT)

References: this document contains references to 101 other documents.

To copy this document: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

The fulltext of this document has been downloaded 557 times since 2016\*

### Users who downloaded this article also downloaded:

(2016), "The effects of three customer-to-customer interaction quality types on customer experience quality and citizenship behavior in mass service settings", Journal of Services Marketing, Vol. 30 Iss 4 pp. 384-397 <<https://doi.org/10.1108/JSM-06-2014-0194>>

(2016), "When social media met commerce: a model of perceived customer value in group-buying", Journal of Services Marketing, Vol. 30 Iss 4 pp. 398-410 <<https://doi.org/10.1108/JSM-04-2014-0129>>



Access to this document was granted through an Emerald subscription provided by emerald-srm:363333 []

### For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit [www.emeraldinsight.com/authors](http://www.emeraldinsight.com/authors) for more information.

### About Emerald [www.emeraldinsight.com](http://www.emeraldinsight.com)

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

\*Related content and download information correct at time of download.

# Assessing the impact of mystery client traits on service evaluation

Pedro Quelhas Brito

Faculdade de Economia, Universidade do Porto, Porto, Portugal, and

Meena Rambocas

Department of Management Studies, The University of the West Indies (St. Augustine Campus), St. Augustine, Trinidad and Tobago

## Abstract

**Purpose** – This study aims to investigate the reliability of a mystery client (MC) as a service evaluation technique taking into consideration personal differences of the MC agents.

**Design/methodology/approach** – The ratings from 144 MCs from 355 evaluations of computer and electronic stores were cross analyzed with eight psychographic and demographic profile variables.

**Findings** – MCs who were highly involved in the product category were more critical of service responsiveness with respect to product demonstrations and listening to customer requirements. On the other hand, MCs with stronger faith in intuition were more inclined to rate services higher on empathy with respect to employees making a conscientious effort to understand customers' needs.

**Practical implications** – Depending on the service marketing goals, managers learn to define which aspects of MC profile they should consider or avoid during the recruitment as well as becoming more critical when they analyze the evaluation reports to avoid an interpretation bias.

**Originality/value** – The usefulness of the MC tool relies on its reliability and credibility as a marketing research technique. It was identified that the MC personality traits are more likely associated with marketing service evaluation variability.

**Keywords** Reliability, Service evaluation, Personality traits, Mystery client

**Paper type** Research paper

## 1. Introduction

### 1.1 Mystery client concept

Mystery client (MC) survey is a data collection technique relevant to business service quality assessment and scientific research. Although impactful in the insights that it can improve service quality, the North American MSPA organization (2014) cautions that the MC survey is not exclusive or sufficient, but a complementary method to other research tools. In the scientific literature, MC has been applied in several service research contexts: health (Gosselt *et al.*, 2007; Young *et al.*, 2009; Glasier *et al.*, 2010; Wong *et al.*, 2012), banking (Tarantola *et al.*, 2012), family planning (León *et al.*, 1994), hospitality tourism and transportation (Nathanail, 2008; Butcher *et al.*, 2009) and retailing (Finn and Kayandé, 1999; Gómez *et al.*, 2011; Wang *et al.*, 2012).

Table I combines several definitions of MC surveys. In marketing management where the service encounter is the main interface between a customer and service provider, the more general term “client” will be adopted here instead of “shopper” term. Although both terms overlap the same potential buyer function, the former emphasizes the service

experience process, while the latter is focused more on the transaction. From these definitions, it was possible to depict the following common features of the MC concept:

- designed to collect customer and competitive intelligence information;
- undercover or anonymous nature of the job;
- discloses the process of service provider–customer interaction rather than the outcome; and
- represents a learning opportunity for an organization, as the MC technique is rooted in third-party feedback, and so it is logically the element most closely related to an actual client.

Apart from the quality and quantity of the collected and reported data by the MC agent, there are two relevant criteria to take into consideration in the design and planning of the MC process: reliability and credibility. If subsequent measurement repetition of the same observation produces a similar outcome under similar conditions, we may classify that measurement as reliable (Trochim and Donnelly, 2007). That means that the obtained result is trustful, thus it is also credible. Such credibility is based on the evidence of objectivity of the measurement process in opposition to the subjectivity (prone to error) of the agent who performs it (Bocking, 2004).

The current issue and full text archive of this journal is available on Emerald Insight at: [www.emeraldinsight.com/0887-6045.htm](http://www.emeraldinsight.com/0887-6045.htm)



Journal of Services Marketing  
30/4 (2016) 411–426  
© Emerald Group Publishing Limited [ISSN 0887-6045]  
[DOI 10.1108/JSM-01-2015-0004]

Received 8 January 2015  
Revised 29 April 2015  
20 August 2015  
Accepted 7 October 2015

Table I Mystery client definitions

| Authors                        | Definition   |
|--------------------------------|--|
| Hurley (1998)                  | "Mystery shopping is an approach that aims at getting service providers to see their services as consumers see them . . . anonymous professionals show up as customers". p. 594  |
| Ford <i>et al.</i> (2011)      | "Mystery shoppers go to an organization to experience a service 'incognito'. They act as a typical customer to assess predetermined service standards of service quality are being met by employees delivering the service". p. 159  |
| Cook <i>et al.</i> (2002)      | "Mystery shopper is a surrogate guest/customer who, unbeknownst to the service provider. Evaluates the service against 'ideal' and then objectively reports out on their experience". p. 170   |
| Butcher <i>et al.</i> (2009)   | "Mystery shopping involves an outsider to visit the accommodation as a guest and report on how they perceived the level of service on key attributes". p. 392  |
| Tarantola <i>et al.</i> (2012) | "They act as normal or potential customers and make unannounced visits to the company . . . to evaluate and monitor customer satisfaction and quality of service in different sectors". p. 10104   |
| León <i>et al.</i> (1994)      | "Mystery-client technique method is characterized by a new 'client', who controls the client-provider interaction as she or he solicits services and is not obtrusive, since the provider is unaware that the client is conducting a test". p. 185   |
| Finn and Kayandé (1999)        | "Typically, the observer enters the outlet to be evaluated posing as an average customer and, immediately after engaging in what appears to be a normal customer interaction, completes a detailed report on various aspects of the service and shopping environment at the outlet". p. 1                |
| Wiele <i>et al.</i> (2005)     | "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". p. 533 |
| Wilson (1998)                  | "Mystery shopping, a form of participant observation, uses researchers to act as customers or potential customers to monitor the processes and procedures used in the delivery of a service". p. 148   |

Ultimately, under the service quality narrative and framework (Ranjan *et al.*, 2015; Urban, 2013), MC helps the organizations to control operation management efficiency, to detect problems and to avoid service failures. That specific topic will be examined in the next section.

### 1.2 Justification of mystery client method

The rationale for service evaluation activity is quite straightforward: quality and innovation. Service quality and the implementation of an innovation culture are the recognized strategies for improving competitive advantage, as well as the key area customers focus on during the service encounters (Bowen, 1990; Bitner, 1990; Jayawardhena, 2010; Kim *et al.*, 2012). Service evaluation is neither static nor universal, but appraised on a wide range of factors varying by provider-to-provider and environment-to-environment. However, although evaluations are subjective, the vast numbers of service evaluation frameworks provide comprehensive and thorough evaluation frameworks (Seth *et al.*, 2005). Attributes evaluated range from tangible physical elements such as service encounter appearance, waiting time and product display, to abstract intangible elements like perception of employee responsiveness, competence and satisfaction. Grönroos (1984) described services on two dimensions: functional (what is provided) and technical (how it is provided). Specifically, functional evaluation is the outcome related to whether customers receive utility from service interactions. Rust and Oliver (1994) extended this conceptualization to include environmental factors (conditions under which the service is performed). Farmer (1988) proposed a Service Attributes Quality Model, which bases service quality on three basic attributes: physical facilities and processes, people behavior and customer professional judgment. McDougall and Levesque (1994)

added a fourth dimension to service evaluation, which involves evaluating encounters based on the ease of accessing the service experience (enabling). Brady and Cronin (2001) introduced the concept of interpersonal and administrative interaction into service experience evaluation. Dagger *et al.* (2007) synthesized a hierarchical service evaluation model specific to the health-care industry and raises concerns of transferability to other service areas.

A modified version of Parasuraman *et al.*'s (1994) model recognized the multi-dimensionality of service evaluation, and future developments adopted an integrative perspective by analyzing service performance (Brady and Cronin, 2001; Dagger *et al.*, 2007) and incorporating objective factors such as service quality environment and quality of processes (Zhanga *et al.*, 2012; Urban, 2013).

### 1.3 Research goals

The quest to determine to what extent this method is objective, an accurate or reliable measure of the experience under observation has been attempted by several researchers. Service encounter is the strategic moment when because of an interaction process, the service is provided and the transaction is accomplished. The actor who experiences that service in terms of flow of time and performance judgment is the client (Cook *et al.*, 2002). Consequently, the role of (mystery) client as a collecting data entity is necessarily relevant. Morrison *et al.* (1997) discussed social and cognitive aspects, mostly relying on the different memory mechanisms of encoding, storage and retrieval of observed target information. Wilson (1998, 2001) analyzed the factors/techniques used to maximize the reliability of this diagnostic tool and the managers' justification for using it. Calvert (2005) criticized the technique, as it distorts reality because it considers a snapshot of service in a specific place, at a specific time. Other

authors pinpointed that individual demographical and behavioral characteristics can influence the quality of data collected (Wilson, 1998; Erstad, 1998; Finn and Kayandé 1999; Dagger *et al.*, 2007). León *et al.*, (1994) went further in proving an answer to the quest. Their investigation was the closest to our research goals. They trained 18 women to work in pairs as simulated clients to observe and rate 84 family planning specialists. By measuring the inter-rater agreement of six relevant variables, they found that the checklist-item scores were better than ratings. Although, methodologically insightful, this research was associated with health-care activities thus unrelated to service marketing context.

The aim of the study is to examine the impact of the effect of personal characteristics on the accuracy and consistency of MC technique. Let us consider the following scenario: two observers are simultaneously experiencing a service; if they belong to the same market segment and share similar demographic and psychographic characteristics, they should, in principle, evaluate that service in an exactly identical way. That is the main assumption for the sake of the scientific and managerial validity of the MC method. An MC is not a “true” client but someone who pretends to be and performs the role. Theoretically, to understand the rationale behind the ability to perform an MC task, we rely on two major explanatory branches: multiple personas and social role. Those two theoretical domains are discussed in the next section.

## 2. Theoretical foundations

### 2.1 Multiple personas

Some degree of variability in the way individuals express their personality seems to be a natural human trait. Excluding schizophrenia or psychiatric multiple personality disorder symptoms, multiple personality behavior coexists within a well-adjusted and coherent ego identity or core self (Raggatt, 2000). The self-theory discussion has helped to explain the integration and sane conviviality among those personas, but the self-construction process is also a function of significant others' representations (Andersen and Chen, 2002). The diversity of personas is dependent on interpersonal contextual cues to induce a repertoire of self-regulated relational selves, which is particularly relevant and manifest in consumer behavior (Bahl and Milne, 2010; Gould, 2010). The process of generating multiple personas is dynamic. If we rely on our past experience self-schemas to build those personas, we also have to nurture and scaffold possible selves to cope with future situations (Markus and Nurius, 1986).

Accepting the evidence that human beings exhibit such personality plasticity, the number of potential personas an individual can develop and reveal may vary. The self-concept differentiation was defined by Donahue *et al.* (1993, p. 834) as “the tendency to see oneself as having different personality characteristics in different roles and psychological adjustments”. Holding a high self-concept differentiation means having a highly fragmented identity to flexibly and adaptively respond to distinct role requirements. That division of self implies emotional and social relationship adjustments, and is more common in the contexts of frequent social role change and role performance dissatisfaction.

A possible psychological explanation of multiple personas is the simultaneous management of multiple impressions

(Leary and Allen, 2011). A process designated as self-presentational persona reflects a person's desire to project specific images deemed appropriate to a particular target or situation. The number of self-presentational profiles can range from a distinct mode, conveying a more idiosyncratic spin in tune with the individual's self-views to a normative mode. Within the latter, the size of repertoire of personas varies according to the degree of agreeableness, self-esteem and authenticity or Machiavellianism stance.

### 2.2 Role theory

Because the multifaceted nature of people explains the reason for multiple personas, the social instrumentalization of this characteristic has been studied by a social role theory. Rooted in the division of labor concept and based on the dramaturgical metaphor, the social role determines distinct occupational positions for the agents corresponding to a specific job status, thus the person and the role are independent concepts (Biddle, 1986).

The role represents a social construction which is structurally defined to fulfill not only biological needs but also work-related requirements. Apart from the stereotyped sex role classification, both genders can simultaneously occupy the roles of parents, care provider, mother/father, wife/husband and daughter/son (Barnett *et al.*, 1992; Martire *et al.*, 2000). In addition to the clear evidence of role accumulation in the non-normative condition, adult individuals execute work-related activities as employees in a hierarchical organization, being simultaneously subordinate *vis-à-vis* her/his supervisor and leaders of a team (Hong and Seltzer, 1995).

Occupying multiple roles demands resources of time and energy, and regardless of the flexibility span, those assets are limited, thus requiring some degree of work–family tradeoff (Greenhaus and Beutell, 1985). A common consequence of playing multiples roles within an organization is role conflict and role ambiguity occurring because of an expectation/knowledge deficiency about each manager-specific role (MaGee *et al.*, 1989; Shepherd and Fine, 1994). Drawing on the literature of consumer behavior/decision-making, buyer experience in a service encounter also represents a well-defined role where both service employees and consumers participate like actors on stage, performing according to a “service script” (Guiry, 1992) and facing similar psychological and sociological constraints (Otnes *et al.*, 1997). Buyers and employees behave interdependently and complementarily. Each should understand the other's positions and predicted role, as role expectations affect service quality and customer satisfaction (Solomon *et al.*, 1985).

If a human being inherently assumes different personas following specific situational cues or responding to her/his intrinsic choices and is also able to play multiple roles, then how can she/he compartmentalize such diversity among personas/roles that are not always congruent? To what extent is she/he competent to psychologically adjust or reorganize her/his personality characteristics according to the most appropriate role? This dispute around the person–role merger magnitude (Turner, 1978) raises a question about consistency (identity theory – Jackson and Smith, 1999) versus appearance (role theory – Biddle, 1986) in the process of “wearing in” and



“taking off” different roles. Surely, to attain the MC method objectives, the further and deeper the merger process is potentially more accurate will be her/his evaluations.

### 3. Mystery client potential individual differences

Behavioral dimensions of service operations should not be undermined, as regardless of how mechanistically employees are programmed to manage their activities, there are always in some extent personal contact and interaction with someone (Croson *et al.*, 2013). Both (mystery) clients and employees express and react emotionally which may influence their response/performance (Van Eerde and Peper, 2008; Wang *et al.*, 2012; Urda and Loch, 2013).

Previous research on consumer behavior literature extensively examines the influence of individual characteristics on service evaluation (Iacobucci and Ostrom, 1993; Hopkins *et al.*, 2009). Some relevant characteristics emerging are demographics and psychographic. Given the interactive and personal nature in production and consumption of services, these traits are considered especially relevant to service providers. We adopt a conceptualization of individual traits to emphasize distinctive qualities of an individual which shape personal character. Based on a review of the literature, personal characteristics are expected to have a significant influence on service evaluation (Dabholkar and Bagozzi, 2002). Six characteristics were examined (five psychological and one demographical) to ascertain the statistical invariance in the MC rating. Each of these factors is relevant in consumer-related research from previous empirical studies. The hypotheses rooted on the theoretical predictions are formulated in next sub-sections.

#### 3.1 Gender

There is consensus that men and women act differently in their environment and evaluate their encounters differently (Iacobucci and Ostrom, 1993). Men are more likely to pay attention to functional components of their service experiences, while women pay more attention to relational aspects of their encounters. Therefore, in rating services, men are likely to pay attention to attributes that pertain to the core service, while relationship attributes are more salient to women (Sharma *et al.*, 2012). Although, to the best of our knowledge, the impact of gender difference on MC has not been examined, the literature suggests that gender might have some influence on personal evaluation (Levy, 1988):

*H1.* The association between MC’s gender and each service evaluation variable is statistically significant.

#### 3.2 Opinion leadership

Reynolds and Darden (1971) suggested that individual difference relates to a person’s inherent leadership characteristic. Generally, opinion leaders have a dominating influence on the behavior of others and in particular, members of their social group. They are translators of marketing stimuli and act as intermediaries between the marketer and social peers. Opinion leaders are self-confident and share product information with members of her/his social peer. These characteristics manifest in the shopping practices. For instance, customers with more product-related knowledge are

more likely to have enhanced information-processing capabilities (Alba and Hutchinson, 1987; King and Balasubramanian, 1994; Bont and Schoormans, 1995). The influence of opinion leadership trait on MC evaluation remains, to the best of our knowledge, largely unanswered. MCs who are opinion leaders will have a clear idea of what to expect from their service experience encounters and will be more assertive in their evaluations taking into account more details. It is in this regard that MC will be more vigilant and less tolerant for perceived non-conformities:

*H2:* 1 The ability to influence others (component of opinion leadership) is positively associated with the ratings of each service evaluation variable.

*H2:* 2 The agreeableness with the information received from others (component of opinion leadership) is not statistically associated with the ratings of each service evaluation variable.

#### 3.3 Consumer involvement

Consumer involvement towards a product category reflects the degree of personal relevance and importance (Zaichkowsky, 1994). However, the conceptualization of product involvement has evolved in recent times and extended to include affective and emotional components (Zaichkowsky, 1994; Kim and Sung, 2009). Relative to the impact of involvement in service evaluation, Goodman *et al.* (1995) concluded that customers who are more involved with a product/service category are more likely to see each aspect of the service encounter as personally relevant. It is reasonable to expect therefore that MC, who consider service category as highly relevant to their life, would be more inclined to evaluate service attributes in detail:

*H3:* 1 The involvement in product category is positively associated with the ratings of each service evaluation variable.

*H3:* 2 The product ignorance (component of consumer involvement) is not statistically associated with the ratings of each service evaluation variable.

#### 3.4 Need for emotions

The need for emotion describes an individual propensity to actively seek out and enjoy emotional stimuli, and have an expressed preference to use their own emotions during personal interactions (Raman *et al.*, 1995). Generally, consumers with higher emotional propensity exhibit emotional responses to environmental stimuli (Harris and Moore 1990) and tend to evaluate service encounters more positively compared to individuals who experience negative emotions or are neutral. Gardner (1985) contends that a positive mood increases the likelihood that performance and behavior are evaluated on the positive end, whilst a negative mood is more likely to influence a negative evaluation of service outcomes:

*H4.* The association between need for emotion and each service evaluation variable is statistically significant.

### 3.5 Faith in intuition

Faith in intuition was developed by Epstein *et al.* (1996), originally named intuitive–experiential scale, and it correlates significantly with personality traits of conscientiousness and esoteric thinking. According to Epstein *et al.* (1996), there are two systems of human processing, namely, rational thought and experiential thinking, which influence each other. The rational system is a conscious, analytical, slower and relatively free of emotion. In contrast, the experiential system is a pre-conscious, fast, automatic, holistic and associated with affection. Following this concept, we refer to faith in intuition as an intuitive spirit that measures the confidence in one's feeling and immediate impressions as a basis for decision-making:

H5. The association between the components of faith in intuition and each service evaluation variable is statistically significant.

### 3.6 Need for cognition

Cacioppo and Petty (1982) described a person's need for cognition as the tendency to engage in and enjoy cognitive deliberations. The authors explained that individuals with higher cognition preferences tend to scrutinize marketing stimuli, engage in extensive elaboration and reasoning and enjoy relatively effortful cognitive tasks. These individuals pay attention to intricate details on messages and environmental stimuli when forming opinions and attitudes (Cacioppo *et al.*, 1982). These individual differences can influence how people understand, evaluate and interpret their environment and affect every facet of marketing. On the other hand, individuals with lower cognitive need are more passive in their approach and lack the ability and motivation to think extensively about the stimuli:

H6. The association between the components of need for cognition and each service evaluation variable is statistically significant.

### 3.7 Overview

We have two streams of theoretical elaborations to interpret the prospect of service evaluation differences between MC:

- 1 *Rationalist view*: As well-trained professional MCs tend to strictly follow the prescribed procedures by mechanically repeating previously learned role script and pursuing an optimal decision-making goal to guide their activities (Yanof, 2012; Croson *et al.*, 2013). Because the MC job design is highly standardized with predefined evaluation criteria and structured responses, we expect that the MC evaluation will always be equal regardless of individual demographical and psychological profile.
- 2 *Behavioral operations view*: MC professionals are also human beings; although acting in a deliberate manner, there are necessarily individual differences which hardly lead to uniform results or to a unique outcome.

Even when environment and contextual factors were under control according to pre-defined parameters, personal heterogeneity among (MC) agents produces some variability (Croson *et al.*, 2013; Urda and Loch, 2013).

To what extent do the service evaluations' ratings vary as a function of the differing MC's psychographical and demographical profile? Assuming that all MC participants in our study perform the same task, share the same training, role script and data-collecting instrument and motivation, are there statistical significant differences in their ratings between two mutually exclusive MCs' profile? Which theoretical view is more likely to pertain? According to the rationalist view, the hypothetical answer would be "no". If the behavioral operations view is retained, we expect that a variety of ratings will be in tune with MCs' heterogeneity.

## 4. Methodology

### 4.1 Research design and statistical techniques

The data collection mirrored the sequential approach of Erstad (1998). The MCs were recruited from a sales management course offered to students pursuing master's qualifications at a local university. The use of students is justified in this case, as they represent a homogeneous group in terms of education and age. The potential MCs received intensive training designed to provide skills and expertise on how to conduct effective MC surveys. At the end of the training session, each student was assigned to a group that consisted of three to four members. Further, 36 retail stores were selected for evaluation based on two selection criteria: type of product sold (retailers of consumer electronic products) and location of the retail store (located in Porto Metropolitan area, Portugal). Overall, 144 MCs conducted 355 evaluations of the stores and salespersons. Over 200 h of contact was spent between mystery shoppers and salespersons, and this facilitated cross comparisons, given that the MCs rated different stores. Given the independent multiple visits, the comparisons were not biased. Hair *et al.* (2010) advised that sample sizes should achieve a minimum of five observations for each dependent variable; in our case, given that we evaluated eight personal characteristics (seven psychological and one demographical), the minimum sample criterion was satisfied. To maintain consistency and control, each group interacted with the same salesperson in each store. The MC adopted the role of prospective buyers of electronic products (laptop, MP3/4 player or digital camera). Each member played a specific role: one acted as the prospective buyer, while the others played the role of accompanying friends and had the task of observing and recording data, to reduce the memory burden on the MC (Morrison *et al.*, 1997). Relative to the data-collection process, the type of interaction between the MC and service personnel was standardized and scripted. This allowed consistency with interaction but also gave the salesperson the opportunity to act freely as possible. The same MC team visited at least two other stores, allowing each member to perform different tasks active versus passive observation/interaction. Simultaneously, each of the 36 stores was visited by at least two teams. That is the same employee was observed several times. Relative to interaction, the MCs were required to wait for 5 min maximum near the location of the product they were interested in purchasing. If the MC received no assistance within 5 min, the MC then actively sought help. All participants memorized a script assuring that they verbalized exactly the same content in the same way – non-verbal

language included (Appendix). Immediately after the service encounter, each MC independently completed a standardized service evaluation form that evaluated eight specific criteria rooted in Parasuraman *et al.*'s (1994) service quality model and satisfaction with the overall service encounter: salesperson empathy, assurance, responsiveness, reliability, appearance of people, appearance of physical, environment satisfaction and service adhering to expectations. The service evaluation instrument consisted of 32 items measuring these 8 facets of the MC experience on a seven-point Likert scale. To maintain consistency, negatively worded items were reversed coded. The first six were measured on multi-item scales and the last two on a single-item scale. The use of Parasuraman *et al.*'s (1994) service quality model is not new to MC research and has been used in multiple service evaluation studies (Dawes *et al.* 2000; Wiele *et al.*, 2005).

To analyze the data, the study used a series of steps to condense, purify and categorize the data. This process started with analyzing the MC personality using exploratory factor analysis (EFA) and k-mean non-hierarchical cluster analysis. Second, the analysis of variance (ANOVA) was used to determine whether group differences existed on each of the 32 evaluation criteria. This facilitated cross-referencing of the average stores and facilitated the detection of group differences. For example, Store 2 (including four salespersons working there) was evaluated ten times by three separate MC groups. Each group had, on average, 3.4 members. The rating of each group was calculated and cross-referenced to other groups who visited the same store. This referencing was carried out on each service variable measured. The cross-referencing was set at 95 per cent confidence level and sought to determine whether variations existed in store evaluation. Multiple regression analysis was subsequently used to ascertain the significance of each characteristic in explaining MC evaluation, and multiple analysis of variance (MANOVA) was used to examine the statistical difference between varying MC profiles. As the rationale behind this research is to assess if the individual traits affect the way they evaluate service providers, a necessary condition to assure accuracy in that measurement is to control the sources of variance. We accomplished that goal in three ways:

- 1 the use of a homogeneous group of MCs in terms of education and age;
- 2 at least one-third of the observations/evaluations were made in the same stores and with the same salesperson by several teams of MCs ensuring that all were submitted to the same stimuli; and
- 3 the time span of that observation/evaluation was one week which rendered unlikely variations in store environment and changes in retail or product/category promotional activities.

There are other methodological alternatives. An experimental research design under laboratory settings would guarantee that all participants watched the same video exhibiting in-store environment and a sales process interaction simulation. Still, the lack of ecological validity would undermine that option. If all MC participants visited a single store and interacted exactly with the same salesperson, it would fail effectiveness because of the suspicion about the intention. The mystery circumstance quickly ceased to work. To overcome such

limitations, we considered 36 stores (real context) and multiple visits of different MC teams interacting with the same salesperson allowing to generate data and making it feasible to apply to the multivariate analysis. Therefore, underlying the service/store evaluation (as dependent variables) and MCs' profile (as independent variables) relationship, the unique source of variance was the MCs' individual traits.

## 4.2 The measurement instrument

### 4.2.1 Individual profile

The administration of the profiling instrument occurred at the briefing session. The design of the instrument facilitated the subsequent grouping of MCs into distinct psychographic profiles. The instrument consisted of 82 items, which were MCs' disposition on eight characteristics (seven psychological and one demographical), namely: opinion leadership (with two components), level of involvement in purchase decision (with two components), need for emotions, faith in intuition and need for cognition and gender. For measurement, the study used the constructs adopted from previous studies (Table II).

### 4.2.2 Service evaluation

The service evaluation instrument consisted 32 items measuring 8 facets of the MC experience on a seven-point Likert scale. The service attributes measured aspects of service quality dimensions proposed by Parasuraman *et al.* (1988) but modified it to fit the context of the study. Specifically, the instrument evaluated salesperson empathy, assurance, responsiveness, reliability, appearance of people, appearance of physical environment, satisfaction and service adhering to expectations (Appendix). To maintain consistency, negatively worded items were reverse-coded. The first six were measured on multi-item scales and the last two on a single-item scale.

## 5. Findings

In total, 52 per cent were female. On average, MC participants were 23.14 years old (SD = 2.99).

### 5.1 Determining the mystery client profile

First, EFA using the principal component approach with varimax rotation was used as the primary data-reduction technique. The use of EFA is justified given its ability to systematically reduce a number of variables to simplify the analysis and maintain parsimony. Items with low correlations ( $r < 0.30$ ) and communalities ( $< 0.50$ ) were removed from further analysis. To determine the reliability, Cronbach's alpha statistic was calculated for the five psychographic characteristics. The first characteristic was *Opinion Leadership* on eight items adopted from Reynolds and Darden (1971). One item was deleted from further analysis because of low correlations. The retained items generated a two-factor solution and explained 78 per cent of total variance. Factor 1 comprised four variables relating to participant's ability to influence others. The reliability for this retained scale was moderately high – 0.934. For each of the subsequent constructs, Table III summarizes the percentage of total variance explained (EFA output), their components, reliability value (Cronbach's alpha) and the corresponding items. Comparing with the original structure (Table III), some of the



Table II Individual characteristics

| Type   | No. of items | Measurement scale   | Reported reliability            | Description   |
|--|--------------|---|---------------------------------|---|
| <b>Opinion leadership (Reynolds and Darden, 1971)</b>                  | 8            | Seven-point Likert (1 – Strongly disagree to 7 – Strongly agree)    | 0.79 and 0.73                   | Opinion leadership is reflected in two basic dimensions: first, the extent to which a person exerts influence and gives information to others on his/her own opinions and second, the extent to which the person actively seeks information from others   |
| <b>Consumer involvement profile (CIP) (Laurent and Kapferer, 1985)</b> | 16           | Five-point Likert (1 – Strongly disagree to 5 – Strongly agree)     | 0.80; 0.90; 0.88; 0.82 and 0.72 | Conceptualize involvement as a multifaceted construct determined by five antecedents, namely: perceived importance of the relevance of the product; perception of risk in purchase which relates to the consequences of mis-purchase; the symbolic value assigned to the product class; hedonic or emotional value attributed to the product class; and interest in an enduring relationship with the product class |
| <b>Need for emotion (NEF) (Raman et al., 1995)</b>                     | 12           | Five-point Likert (1 – Strongly disagree to 5 – Strongly agree)     | 0.87                            | Evaluates individual propensity to look for emotional situations and stimuli and to use emotion in their interaction with world. The scale focuses on short-term emotions   |
| <b>Faith in Intuition (FI) (Epstein et al., 1996)</b>                  | 12           | Five-point scale (1 – Completely false to 5 – Completely true)      | 0.87                            | It measures confidence in decision making which is reflected in several aspects including personality, adaptation, self-fulfillment, interpersonal relations  |
| <b>Need for Cognition (NFC) (Cacioppo and Petty, 1982)</b>             | 34           | Nine-point Likert (4 – Completely disagree to 4 – Completely agree) | 0.80                            | Relates to the extent to which an individual engages in and enjoys thinking   |

items were dropped because of the low statistical significance ( $p > 0.1$ ) and inter-item correlations.

The second stage involved classifying MCs into homogeneous and distinctive groups based on the personality traits derived from the first stage of our analysis (factor analysis described above). To define the classifications, the study used a non-hierarchical k-mean clustering approach. The k-mean approach is a common method to partition data sets into predefined number of groups or k groups (Wagstaff et al., 2001). The process involved three steps: first, we defined the amount of groups (K) and using a random process to generate cluster centers for each group. Second, each observation was assigned to one of the clusters based on similarity of each center and iteratively refined, so that each observation resides closest to its respective center. Finally, each center is updated to reflect the mean of the observations classified into the specific group. The objective of this approach revolves on grouping MCs into taxonomies based on personal characteristics. Two groups ( $K = 2$ ) were defined for each personal characteristic (Group 1 – high and Group 2 – low). The numbers indicate how strong the MC ratings were on that particular characteristic. MCs with stronger ratings were classified into Group 1 – or high category, whereas lower ratings were classified into Group 2 – low category. A standard *F*-test showed that the difference between both groups was statistically different on each personal characteristic. Table IV

summarizes the mean score of the high and low groups for each personal characteristic.

## 5.2 Analysis of service evaluation criteria by mystery client group: cross analysis

The general aim of the cross-referencing analysis was to explore whether service evaluation criteria can vary across individual groups. Given that our MC evaluated different stores, it was imperative to test the consistency of MC service ratings. The study cross analyzed the ratings of each MC on each service criteria using one-way ANOVA analysis. The purpose of the ANOVA was to examine whether the MC ratings were related to individual characteristics. ANOVA compared the mean rating of each store on the 32 items on the questionnaire which measured 8 service dimensions outlined in our research model (assurance, empathy, tangible-people, tangible-place, reliability, responsiveness, conformance to expectations and satisfaction). The research design ensured that each group visited at least two stores previously visited by other groups and interviewed the same salesperson, hence conceptually justifying the cross analysis. In addition, the research was designed with a relatively balanced group size which allowed us to identify significant differences in MC ratings. In terms of process, the cross analysis involved the following stages:

- For each store visited, each group's average rating on each service item was compared.



Table III Measurement of individual characteristics

| Construct (%total expl.)              | Components (Reliability)     | Items  |
|---------------------------------------|------------------------------|--|
| Opinion leadership<br>(78%)           | Ability to influence (0.934) | My friends and neighbors often ask my advice about electronic/digital products<br>I sometimes influence the type of electronic/digital products my friends buy<br>I feel that I am generally regarded by my friends and neighbors as a good source of advice about electronic /digital products<br>My friends usually come to me, more than I go to them, for information about electronic /digital products |
|                                       | Agreeableness (0.770)        | I often seek out the advice from my friends regarding which electronic /digital products I should buy<br>I spend a long time talking to my friends regarding which electronic /digital products I should buy<br>My friends and neighbors usually give me good advice on what brands of electronic /digital products I should buy   |
| Consumer involvement<br>profile (71%) | Involvement (0.874)          | Buying computers/related products is like buying gifts for myself<br>It gives me pleasure to purchase electronic/digital products<br>I attach a great deal of importance to electronic/digital products<br>One can say electronic /digital products interest me a lot  |
|                                       | Ignorance (0.787)            | Choosing electronic /digital products is rather complicated<br>When I face a shelf of electronic /digital products, I always feel a bit at a loss to make my choice<br>When one purchases electronic /digital products, one is never certain of one's choice   |
| Need for emotion<br>(70%)             | Need for emotions (0.704)    | Experiencing strong emotions is not something I enjoy very much<br>I look forward to situations that are less emotionally involving<br>I don't look forward to being in situations that others have found to be emotional<br>I prefer to not get involved with the emotional aspects of any situation  |
| Faith in intuition (66%)              | Faith in others (0.8115)     | My initial impressions of people are almost right<br>I trust my initial feelings about people<br>I am a very intuitive person<br>I believe trusting to my hunches<br>When it comes to trusting people, I can usually rely on my "gut feelings"   |
|                                       | Self assured (0.790)         | I often have a clear visual image of things<br>I am good at visualizing things   |
| Need for cognition<br>(72%)           | Cognition preference (0.703) | I would rather do something that requires little thought than something that challenges my thinking abilities<br>I find little satisfaction in deep and lengthy deliberations<br>Thinking is not my idea of fun  |
|                                       | Cognition practice (0.707)   | I like tasks that require little thought once I've learnt them<br>I only think as hard as I have to  |
|                                       | Trust in cognition (0.646)   | More often than not, more thinking just leads to more errors<br>I usually become hesitant about making important decisions after thinking about them   |

Table IV Non-hierarchical cluster analysis

| Personal characteristics |                                | Group 1:<br>High | Group 2:<br>Low |
|--------------------------|--------------------------------|------------------|-----------------|
| Char. 1                  | Ability to influence (Op.Lead) | 0.519**          | -1.477**        |
| Char. 2                  | Agreeab (Op.Lead.)             | 0.570*           | -1.269*         |
| Char. 3                  | Involv (Cons.Inv.)             | 0.728**          | -0.956**        |
| Char. 4                  | Ignorance (Cons.Inv.)          | 0.791**          | -0.88**         |
| Char. 5                  | Need for emot (NFE)            | 1.13**           | -0.595**        |
| Char. 6                  | Faith in others (FI)           | 0.673**          | -0.743**        |
| Char. 7                  | Self assured (FI)              | 0.456**          | -1.26**         |
| Char. 8                  | Cogn Prefe (NFC)               | 0.749**          | -0.862**        |
| Char. 9                  | Cogn practice (NFC)            | 1.05**           | -0.645**        |
| Char. 10                 | Trust in Cogn (NFC)            | 0.934**          | -0.745**        |

Notes: Mean contrast are significant ( \*\* $p < 0.05$ ; \* $p < 0.1$ ) according to the  $F$ -test

- For each store visited, an ANOVA compared the average rating of each group on each service item at a significance level of 95 per cent.

ANOVA allowed us to identify statistical differences between our group means by using the  $F$ -ratio.

In summary, the result showed that there are significant differences in intergroup comparison on each service evaluation criteria. This means that individuals who rated the same store, varied in their service evaluation. By statistically comparing the different MC groups ( $p < 0.05$ ), Table V shows the 32 service evaluation attributes. The most profound influence occurred in the perception of physical store appearance (dusty) and the lowest perception of salesperson ethics. This provided the evidence that suggested significant intra-personal influence on service evaluation that may be unique to an individual, which suggests the need to engage in

Table V A comparison of inter-group rating on each service attribute

| Store/salesperson evaluation criteria                       | Store/salesperson attributes           | No. of stores with significant deviation based on comparison of intergroup ratings | % of statistical deviation based on 36 stores |
|---|--|--|---|
| Assurance   | 1.1 – Competency of vendor             | 13   | 36  |
|   | 1.2 – Trustfulness of vendor           | 11   | 31  |
|   | 1.3 – Clarity of vendor                | 8  | 22  |
|   | 1.4 – Commitment                       | 14   | 39  |
|   | 1.5 – Proactive                        | 13   | 36  |
| Empathy   | 2.1 – Courtesy of vendor               | 7  | 19  |
|   | 2.2 – Pleasantness of vendor           | 9  | 25  |
|   | 2.3 – Honesty of vendor                | 8  | 22  |
|   | 2.4 – Ethics of vendor                 | 4  | 11  |
|   | 2.5 – Shows interest                   | 10   | 28  |
|   | 2.6 – Try to understand customer needs | 9  | 25  |
|   | 2.7 – Interest in client recoded       | 5  | 14  |
|   | 2.8 – Help client                      | 10   | 28  |
|   | 2.9 – Listen attentively               | 8  | 22  |
|   | 2.10 – Disposition of vendor           | 7  | 19  |
| Tangible evidence – people                                  | 3.1 – Employee dress                   | 9  | 25  |
|   | 3.2 – Hygiene of vendor                | 10   | 28  |
| Tangible evidence – place                                   | 4.1 – Cleanness                        | 16   | 44  |
|   | 4.2 – Attractive window                | 11   | 31  |
|   | 4.3 – Easy to find products            | 11   | 31  |
|   | 4.4 – Store appearance                 | 11   | 31  |
|   | 4.5 – Access to products               | 12   | 33  |
|   | 4.6 – Lighting of store                | 8  | 22  |
| Reliability   | 5.1 – Efficiency of vendor             | 12   | 33  |
|   | 5.2 – Persuasion of vendor             | 14   | 39  |
|   | 5.3 – Accurate information             | 12   | 33  |
| Responsiveness  | 6.1 – Insistence of vendor (pushy)     | 7  | 19  |
|   | 6.2 – Ask all necessary question       | 8  | 22  |
|   | 6.3 – Suggest appropriate products     | 8  | 22  |
|   | 6.4 – Demonstration                    | 6  | 17  |
| Extent to which service encounter conformed to expectations | 7                                      | 19   |   |
| Satisfaction with service encounters                        | 11                                     | 31   |   |

further analysis using multivariate techniques such as multiple regressions and MANOVA.

### 5.3 Creating summated scales for service evaluation criteria

Although the use of multi-item scales enhance precision and improve reliability, multiple items lends complication to data analysis (Spector, 1992). In this regard, to maintain parsimony and simplify the analysis, summated scales were constructed. Summated scales involve collapsing the data into a single variable. In our study, the service evaluation instrument consisted of 32 questions representing the eight attributes of service quality we intended to measure. In this regard, we simplified the MANOVA and regression analysis by collapsing the 32 items into 8 summated scores. Summated scales were calculated by taking the average of the items in each scale. To account for internal consistency, unidimensionality of each scale was verified through EFA and the Cronbach's alpha statistic was calculated. Furthermore, to adequately compare the relative effects of personality

characteristics on each service attribute, we standardized the ratings. Standardization converts the variables into a common scale (mean = 0; SD = 1) making the variables comparable (Hair *et al.*, 2010).

Table VI presents the descriptive statistics on each service attribute, including the standardized mean and standard deviation and reliability index.

### 5.4 Multiple regression analysis

To analyze the relationship between each service evaluation criterion and MC profile, the study used multiple regression analysis. The technique involved assessing the character of the relationship between each of service attribute (dependent variable) and personal characteristics. To maintain parsimony, confirmatory perspective to multiple regressions was used where the individual characteristics were specified as independent variables. Eight regression variates were formed (one for each service attribute). The regression variates are represented below:

Table VI Service evaluation criteria

| Store evaluation criteria                                   | No. of items in the service attribute | Summated score |       | Chronbach's alpha | Standardize score |    |
|---|---------------------------------------|----------------|-------|-------------------|-------------------|----|
|   |                                       | Mean           | SD    |                   | Mean              | SD |
| Extent to which service encounter conformed to expectations | 1                                     | 3.18           | 0.946 | –                 | 0                 | 1  |
| Satisfaction with service encounters                        | 1                                     | 3.37           | 0.995 | –                 | 0                 | 1  |
| Empathy   | 10                                    | 42.67          | 7.83  | 0.845             | 0                 | 1  |
| Assurance   | 5                                     | 23.73          | 6.4   | 0.887             | 0                 | 1  |
| Responsiveness  | 4                                     | 31.15          | 5.67  | 0.690             | 0                 | 1  |
| Reliability   | 3                                     | 14.68          | 3.62  | 0.752             | 0                 | 1  |
| Tangible evidence – place                                   | 6                                     | 30.61          | 6.40  | 0.738             | 0                 | 1  |
| Tangible evidence – people                                  | 2                                     | 10.74          | 2.80  | 0.604             | 0                 | 1  |

Note: Number of stores visited = 36 ( $p < 0.05$ )

$$Y_i = ABI + AG + IN_v + IG + NFE + FIO + SA + CP + CPR + TIC + Gender_{0=Male, 1=female}$$

CPR = Cognitive Practice (NFC);  
 TIC = Trust in cognition (NFC);  
 Gender = (Male; Female).

Where:

- $Y_i$  = Ratings on each service attribute (salesperson empathy, assurance, responsiveness, reliability, appearance of people, tangibles, satisfaction and service adhering to expectations);
- ABI = Ability to Influence (Opin. Leader);
- AG = Agreeableness (Opin. Lead);
- $IN_v$  = Involvement (Cons. Invol.);
- IG = Ignorance (Cons. Invol.);
- NFE = Need for emotions;
- FIO = Faith in others (FI);
- SA = Self-Assured (FI);
- CP = Cognitive preference (NFC);

The results of the regression analysis are summarized in Table VII.

Table VIII shows that significant relationships ( $p < 0.05$ ) between MC involvement in product category and the ratings assigned to four service evaluation criteria: empathy, assurance, responsiveness and perception of service personnel. This suggests that MCs with different levels of involvement assigned different ratings on three service evaluation criteria. To ascertain the nature of this variation, we used MANOVA in the two profiles (high and low involvement) confirming that MCs' involvement in the product category was a significant discriminator in service evaluations (Roy's largest root  $p = 0.002$ ). MCs were more likely to assign higher ratings on the

Table VII Multiple regression analysis

| $X_i$                          | Extent to which service encounter conformed to expectations |                                      |         |           |                |             |        | Tangible evidence – place | Tangible evidence – people |
|--------------------------------|---|--------------------------------------|---------|-----------|----------------|-------------|--------|---------------------------|----------------------------|
|                                | Y   | Satisfaction with service encounters | Empathy | Assurance | Responsiveness | Reliability |        |                           |                            |
| Ability to influence (Op.Lead) | 0.017   | 0.078                                | -0.028  | 0.049     | -0.012         | 0.064       | 0.107  | -0.027                    |                            |
| Agreeab (Op.Lead.)             | -0.058  | -0.032                               | -0.006  | 0.050     | -0.041         | -0.009      | 0.069  | 0.027                     |                            |
| Involv (Cons.Inv.)             | -0.110  | -0.061                               | -0.084  | -0.125*   | -0.141**       | -0.036      | 0.095  | -0.118*                   |                            |
| Ignorance (Cons.Inv.)          | -0.017  | -0.029                               | -0.035  | 0.050     | -0.064         | 0.041       | 0.016  | 0.042                     |                            |
| Need for emot (NFE)            | 0.018   | 0.023                                | 0.000   | -0.045    | -0.037         | 0.021       | -0.001 | 0.003                     |                            |
| Faith in others (FI)           | 0.068   | 0.011                                | 0.129*  | 0.118*    | 0.074          | 0.080       | 0.071  | 0.091                     |                            |
| Self assured (FI)              | 0.003   | -0.048                               | -0.030  | 0.019     | -0.010         | 0.000       | -0.044 | 0.039                     |                            |
| Cogn Prefe (NFC)               | 0.004   | -0.026                               | 0.029   | 0.001     | 0.001          | 0.000       | 0.058  | 0.057                     |                            |
| Cogn Practice (NFC)            | 0.010   | 0.023                                | -0.032  | 0.029     | -0.033         | -0.074      | -0.003 | -0.116                    |                            |
| Trust in cogn (NFC)            | -0.008  | -0.026                               | 0.001   | -0.015    | 0.068          | -0.022      | 0.017  | -0.040                    |                            |
| D_Gender                       | 0.009   | 0.043                                | 0.103   | -0.064    | 0.109          | -0.025      | -0.025 | -0.009                    |                            |
| $R^2$                          | 0.021   | 0.015                                | 0.037   | 0.029     | 0.026          | 0.017       | 0.035  | 0.032                     |                            |

Notes: \*  $p < 0.05$ ; \*\*  $p < 0.01$

Table VIII Hypothesis test results

| Hypothesis   | Results      |
|--|--------------|
| H1. The association between MC's gender and each service evaluation variable is statistically significant  | Rejected     |
| H2.1. The ability to influence others (component of opinion leadership) is positively associated with the ratings of each service evaluation variable                                    | Rejected     |
| H2.2. The agreeableness with the information received from others (component of opinion leadership) is not statistically associated with the ratings of each service evaluation variable | Not rejected |
| H3.1. The involvement in product category is positively associated with the ratings of each service evaluation variable  | Not rejected |
| H3.2. The product ignorance (component of consumer involvement) is not statistically associated with the ratings of each service evaluation variable                                     | Not rejected |
| H4. The association between need for emotion and each service evaluation variable is statistically significant   | Rejected     |
| H5. The association between the components of faith in intuition and each service evaluation variable is statistically significant   | Not rejected |
| H6. The association between the components of need for cognition and each service evaluation variable is statistically significant   | Rejected     |

service responsiveness when they were less involved in the category ( $M_{\text{high}} = -0.130$ ;  $M_{\text{low}} = 0.194$ ;  $p = 0.002$ ). Further analysis into each item of the empathy and responsiveness service criterion showed two significant areas: listen attentively ( $M_{\text{high}} = 0.106$ ;  $M_{\text{low}} = 0.137$ ,  $p = 0.023$ ) and demonstration ( $M_{\text{high}} = -0.115$ ;  $M_{\text{low}} = 0.148$ ,  $p = 0.014$ ), respectively.

Table VIII also shows significant relationships ( $p < 0.05$ ) between the MCs' faith in intuition and the ratings assigned to two service evaluation criteria: empathy and assurance. This indicates that MCs who vary in faith in intuition characteristics rated their service evaluation differently. The MANOVA confirmed that faith in intuition impacted more the ratings on service empathy (0.045) than assurance (0.083). MCs with high faith in intuition assigned higher ratings on service empathy ( $M_{\text{high}} = 0.088$ ;  $M_{\text{low}} = -0.123$ ). No statistical difference existed in the assurance service criterion. Further analyses into each item in the empathy evaluation variable revealed that the item "salesperson taking time to understand needs of customers" was significantly different between the two profiles ( $M_{\text{high}} = 0.094$ ;  $M_{\text{low}} = -0.140$ ,  $p = 0.030$ ).

## 6. Conclusions

### 6.1 Overview

Given the increased competitiveness of today's service environment, managers rely on service evaluation tools such as the MC to evaluate the quality of customer experiences in service environments. However, contributions made by Morrison *et al.* (1997) and Finn and Kayandé (1999) suggest

that data can be biased, subjected to personal influences based on the MC individual characteristic. Our study tested this proposition by investigating the influence of MC personal characteristics on eight key service evaluation criteria. Our results found very little evidence to support the impact of MC's characteristics on evaluation (Table VIII). Although four out of the eight hypotheses were partially supported (H2.2, H3.1, H3.2 and H5), only two (H3.1 and H5) meet the behavioral perspective: MCs' involvement in product category and MC faith in intuition. However, those two predictors affected only half of the criterion variables. The evidence showed that MCs who scored high on involvement were more critical of the service provider's willingness to help customers and provide prompt services (responsiveness). Perhaps the explanation for this variance rests on the MCs' knowledge in the product category and their general inclination to engage in communication with others. Due to this information-seeking tendency, they have considerable amount of knowledge in the product category, the skills and competencies to understand and evaluate the technical details. This assertion suggests that because the MC who are highly involved in a product category are more likely to seek information about the product category may be more critical of the type and quality of interactions. Therefore, they are less likely to be convinced. The findings also showed that MC with higher faith in intuition-rated service provider caring and attention toward customers higher.

### 6.2 Theoretical discussions

Theoretically, the capability to perform an MC function is rooted in multiple personas, trait and social roles. Although some of the variability in the way individuals express their personality seems naturally intrinsic, as well as the social role, both non-normative and normative, represents a subsequent need for social integration, the MC job responds to its own logic. Our personality plasticity allows us to flexibly adjust but our social condition forces us to fit by matching others' expectations in tune with a well-defined role script. This study analyzes the outcome of how individuals manage those two conditions under a specific relevant marketing research activity: evaluating a service. Theoretically, there are two opposing perspectives. Ultimately, the rationalist view is the only one that matters to assure an acceptable return on MC investment.

MC is a useful tool for evaluating real-time marketing service routines. However, concerns on reliability and validity have cast doubt on the legitimacy of the technique for service evaluation. Surprisingly, only a handful of academic content examined this problem (León *et al.*, 1994; Morrison *et al.*, 1997; Wilson, 1998, 2001; Dawes *et al.*, 2000), and even within this contribution, research paucity, loopholes and deficits exist. Our study addresses this gap by investigating the extent to which different MCs would report similar results despite their differing personality. The results suggest that MCs can provide highly reliable data despite the variation in MCs' individual profiles. However, consideration should be given to MCs' involvement in product category and faith in intuition. Although the impacts were marginal, the study reinforced the need for a standardized research design to minimize the possible contamination of individual idiosyncratic influences. The findings also suggest that the



agency responsible for the MC must exercise caution in selecting MCs to carry out the evaluation, as well as invest in training strategies that alert the MC of the possibility of personal biases in evaluation. These results also can be useful to scrutinize the researcher's personal traits on other methods such as participant observer (Bowen, 2008).

### 6.3 Managerial and social implications

Because of its impact on quality and profitability in high-contact marketing service, these findings have several implications for practitioners. First, it provides empirical evidence that supports the potential subjectivity in the MC evaluation. Our findings show that the ratings on tangible and intangible service factors vary by personal disposition mainly to the extent to which the MC considers him/herself as involved in the category. It shows that the nature of the interaction is perceptual and varies from individual to individual. However, we argue that with instrument structure and employee training, the effect of opinion leadership on service evaluation can be curtailed. For instance, questions and response on the survey evaluation form should be clearly worded and highly structured leaving very little room for ambiguities and interpretation. Second, by alerting the potential MC of the personal tendency to engage in subjective evaluation beforehand, they will be more aware and prepared to recognize and avoid possibility of contaminating the survey. The findings also have implications for the selection of the MC. Recruiters should avoid MCs who are egotistical, or have the desire to individuate themselves from peers. But, recruiters should select MCs with an "open" personality, impartial and neutral. However, the MC should match a specific personality profile (Finn and Kayandé, 1999). Finally, underlying the rational view of marketing services operations, a deviation to the standard behavioral normality is not a "good thing", as it compromises operational efficiency. However, competitive advantage depends on the ability of service providers to personalize (Surprenant and Solomon, 1987). Similarly, while quality calls for standardization, productivity engagement impacts directly employee's compensations (Luria *et al.*, 2014). In franchising retailing format, the dilemma arises between standardization and flexibility or adaptation to local specificities (Pardo-del-Val *et al.*, 2014). Those three conflicting goals – efficiency versus personalization, quality versus productivity and standardization versus flexibility – are difficult to reconcile. However it is important to measure exploration effort – standardization, efficiency and quality – and exploitation practices, personalization, productivity and flexibility (Zhanga *et al.*, 2012). A possible solution could be developing two different teams of MCs. One team would match a psychographical profile focusing on the traditional orthodox role of service evaluation. An alternative MC team should be specialized to capture the innovative employee approaches and their strategies to offer more imaginative and personalized services. Using specific psychographic tests, it would be feasible to previously assess MC applicants during the recruiting process and then allocate them to specific tasks.

The scope of the application of MC covers virtually every service activity whether profit- or non-profit-oriented such as hospitality, retailing, health care, banking, etc. The ultimate major goal behind the implementation of the MC method is

service improvement and failure correction. Therefore, every step accomplished to contribute to enhance the confidence and accuracy of the data collected is indirectly instrumental to provide quality and social welfare.

### 6.4 Limitations and future research

Despite the contributions to service marketing literature, this study is not without limitations. First, the study investigated the impact of key personality traits on one service industry – retail computer and electronic sectors. It would be useful to extend this investigation to other service domains like banking, hospitality or transportation by incorporating in a future research the type of services' variable into the MC service evaluation. Second, the study is limited to examining the differences on homogeneous sample – comprising a restricted age cohort. Additionally, the study only examined the effect of personality traits in isolation of other individual variables that may account for differences in consumer behavior. Our results are contingent to a European cultural background; changing the socio-cultural context may vary the kind, the diversity and intensity of MC personality effects on collected data accuracy (Hopkins *et al.*, 2009).

Future research should consider the effects of demographical factors such as age and gender on the general reliability and validity of MC surveys. As it was mentioned earlier in Section 4.1, an experimental factorial design is an alternative data-collection method particularly useful in fine-tuning detailed analysis of some contextual interference on MCs' reliability. The MC method is commonly applied in exploitation quality practices, as it evaluates the current well-established "normal" activities. Although more complicated to develop and implement, it would be challenging to measure the exploration/innovative quality practices (Zhanga *et al.*, 2012). Apart from the conventional service quality and job performance assessments, a measurement of employee satisfaction in the perspective of client, as well as the evaluation of the external customer mind-set, would be also interesting to include in the service evaluation instrument (Iyer and Johlke, 2015). We tend to assume that salespeople's mission is to please customers, but under some situations, they should manage problematic customers and unethical behavior (Suquet, 2010; Madupalli and Poddar, 2014). Such circumstances require specific skills, which also should be analyzed in the MC research process.

### References

- Alba, J.W. and Hutchinson, J.W. (1987), "Dimensions of consumer expertise", *Journal of Consumer Research*, Vol. 13 No. 4, pp. 411-454.
- Allison, P., Denver, S. and Dickson, D. (2010), "conceptual model for mystery shopping motivations", *Journal Hospitality Marketing and Management*, Vol. 19 No. 6, pp. 629-657.
- Andersen, S.M. and Chen, S. (2002), "The relational self: an interpersonal social-cognitive theory" *Psychological Review*, Vol. 109 No. 4, pp. 619-645.
- Bahl, S. and Milne, G.R. (2010), "Talking to ourselves: a dialogical exploration of consumption experiences", *Journal of Consumer Research*, Vol. 37, pp. 176-195.

- Barnett, R.C. and Marshall, N.L., Pleck, J.H. (1992), "Men's multiple roles and their relationship to men's psychological distress" *Journal of Marriage and the Family*, Vol. 54, pp. 358-367.
- Berry, L. (2000), "Cultivating service brand equity", *Journal of Academy Marketing Science*, Vol. 28 No. 1, pp. 128-137.
- Biddle, B.J. (1986), "Recent development in role theory", *Annual Review of Sociology*, Vol. 12, pp. 1267-1292.
- Bitner, M. (1990), "Evaluating service encounters the effects of physical surroundings and employee responses", *Journal Marketing*, Vol. 54 No. 2, pp. 69-82.
- Bocking, S. (2004), *Nature's Experts: Science, Politics, and the Environment*, Rutgers University Press, New Brunswick, NJ.
- Bont, C. and Schoormans, J. (1995), "The effects of product expertise on consumer evaluations of new-product concepts", *Journal Economic Psychology*, Vol. 16 No. 4, pp. 599-615.
- Bowen, D. (1990), "Interdisciplinary study of service: some progress, some prospects", *Journal of Business Research*, Vol. 20 No. 1, pp. 71-79.
- Bowen, D. (2008), "Consumer thoughts, actions, and feelings from within the service experience", *The Service Industries Journal*, Vol. 28 No. 10, pp. 1515-1530.
- Brady, M.K. and Cronin, J.J. (2001), "Customer orientation effects on customer service perceptions and outcome behaviors", *Journal of Service Research*, Vol. 3 No. 3, pp. 241-251.
- Butcher, K., Sparks, B. and Kennedy, J.M. (2009), "Predictors of customer service training in hospitality firms", *International Journal of Hospitality Management*, Vol. 28 No. 3, pp. 389-396.
- Cacioppo, J. and Petty, R. (1982), "The need for cognition", *Journal of Personality and Social Psychology*, Vol. 42 No. 1, pp. 116-131.
- Calvert, P. (2005), "It's a mystery: mystery shopping in New Zealand's public libraries", *Library Review*, Vol. 54 No. 1, pp. 24-35.
- Chan, K. and Mishra, S. (1990), "Characteristics of the opinion leader: a new dimension", *Journal of Advertising*, Vol. 19 No. 3, pp. 53-60.
- Cook, L.S., Bowen, D.E., Chase, R.B., Dasu, S., Stewart, D.M. and Tansik, D.A. (2002), "Human issues in service design", *Journal of Operations Management*, Vol. 20, pp. 159-174.
- Crosan, R., Schultz, K., Siemsen, E., Yeo, M.L. (2013), "Behavioral operations: the state of the field", *Journal of Operations Management*, Vol. 31, pp. 1-5.
- Dabholkar, P. and Bagozzi, R. (2002), "An attitudinal model of technology based self service: moderating effects of consumer traits and situational factors", *Journal of Academy Marketing Science*, Vol. 30 No. 3, pp. 184-201.
- Dagger, T., Sweeney, J. and Johnson, L. (2007), "A hierarchical model of health service quality: scale development and investigation of an integrated model", *Journal of Service Research*, Vol. 10 No. 2, pp. 123-142.
- Dawes, J., Sharp, B. and Adelaide, N.T. (2000), "The reliability and validity of objective measures of customer service: mystery shopping" *Australian Journal of Market Research*, Vol. 8 No. 1, pp. 29-46.
- Dawson, J. and Hillier, J. (1995), "Competitor mystery shopping: methodological considerations and implications for the MRS code of conduct", *Journal Marketing Research Society*, Vol. 37 No. 4, pp. 417-428.
- Donahue, E.M., Robins, R.W., Roberts, B.W. and John, O.P. (1993), "The divided self: concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation", *Journal of Personality and Social Psychology*, Vol. 64, pp. 834-846.
- Epstein, S., Pacini, R., Raj, V.D. and Heier, H. (1996), "Individual differences in intuitive – experiential and analytical – rational thinking styles", *Journal of Personality and Social Psychology*, Vol. 71 No. 2, pp. 390-405.
- Erstad, M. (1998), "Mystery shopping programmes and human resource management", *International Journal of Contemporaneity Hospitality Management*, Vol. 10 No. 1, pp. 34-38.
- Farmer, J.H. (1988), "A conceptual model of service quality", *International Journal of Operations and Production Management*, Vol. 8 No. 6, pp. 19-29.
- Finn, A. and Kayandé, U. (1999), "Unmasking a phantom: a psychometric assessment of mystery shopping", *Journal of Retailing*, Vol. 75 No. 2, pp. 195-217.
- Ford, R.C., Latham, G.P. and Lennox, G. (2011), "A new tool for coaching employee performance improvement", *Organizational Dynamics*, Vol. 40, pp. 157-164.
- Gardner, M. (1985), "Mood states and consumer behavior: a critical review", *Journal of Consumer Research*, Vol. 12, pp. 281-300.
- Gilbert, F. and Warren, W. (1995), "Psychographic constructs and demographic segments", *Psychology and Marketing*, Vol. 12 No. 3, pp. 223-237.
- Glazier, A., Manners, R.J., Loudon, J.C. and Muir, A. (2010), "Community pharmacists providing emergency contraception give little advice about future contraceptive use: a mystery shopper study", *Contraception*, Vol. 82, pp. 538-542.
- Gómez, R.S., González, I.S. and Suárez, L.V. (2011), "Service quality control mechanisms in franchise networks", *The Service Industries Journal*, Vol. 31 No. 5, pp. 713-723.
- Goodman, P., Fichman, M., Lerch, F. and Snyder, R.P. (1995), "Customer-firm relationships, involvement, and customer satisfaction", *Academic Management Journal*, Vol. 38 No. 5, pp. 1310-1324.
- Gosselt, J.F., Hoof, J.J., Jong, M.D. and Prinsen, S. (2007), "Mystery shopping and alcohol sales: do supermarkets and liquor stores sell alcohol to underage customers", *Journal of Adolescent Health*, Vol. 41, pp. 302-308.
- Gould, S.J. (2010), "To thine own self(ves) be true": reflexive insights for etic self theory from consumers' emic constructions of the self, *Consumption Markets and Culture*, Vol. 13 No. 2, pp. 181-219.
- Greenhaus, J.H. and Beutell, N.J. (1985), "Sources of conflict between work and family roles", *Academy of Management Review*, Vol. 10, pp. 76-88.
- Grönroos, C. (1984), "A service quality model and its marketing implications", *European Journal of Marketing*, Vol. 18 No. 4, pp. 36-44.

- Guiry, M. (1992), "Consumer and employee roles in service encounters", in Sherry, J. and Sternthal, B. (Eds), *Advances in Consumer Research*, Association for Consumer Research, Provo, UT, pp. 666-672.
- Hair, J., Black, W.C., Babin, B.J. and Tatham, R. (2010), *Multivariate Data Analysis*, 7th ed., Pearson, NJ.
- Harris, W. and Moore, D. (1990), "Affect intensity as an individual difference variable in consumer response to advertising appeals", *Advances on Consumer Research*, Vol. 17, pp. 792-797.
- Hodgkinson, G., Fox, J.L. and Smith, E.S. (2008), "Intuition: a fundamental bridging construct in the behavioral science", *British Journal Psychology*, Vol. 99 No. 1, pp. 1-27.
- Hong, J. and Seltzer, M.M. (1995), "The psychological consequences of multiple roles: the nonnormative case", *Journal of Health and Social Behavior*, Vol. 36 No. 4, pp. 386-398.
- Hopkins, S., Nie, W. and Hopkins, W. (2009), "Cultural effects on customer satisfaction with service encounters", *Journal of Service Science*, Vol. 2 No. 1, pp. 45-56.
- Hurley, R. (1998), "Customer service behavior in retail settings: a study of the effects of service provider personality", *Journal of Academy Marketing Science*, Vol. 26 No. 2, pp. 115-127.
- Iacobucci, D. and Ostrom, A. (1993), "Gender differences in the impact of core and relational aspects of services on the evaluation of service encounters", *Journal of Consumer Psychology*, Vol. 2 No. 3, pp. 257-286.
- Iyer, R. and Johlke, M.C. (2015), "The role of external customer mind-set among service employees", *Journal of Services Marketing*, Vol. 29 No. 1, available at: <http://dx.doi.org/10.1108/JSM-09-2013-0237>.
- Jackson, J.W. and Smith, E.R. (1999), "Conceptualizing social identity: a new framework and evidence for the impact of different dimensions", *Personality and Social Psychology Bulletin*, Vol. 25, pp. 120-135.
- Jayawardhena, C. (2010), "The impact of service encounter quality in service evaluation: evidence from a business-to-business context", *Journal Business Industrial Marketing*, Vol. 25 No. 5, pp. 338-348.
- Kim, D.Y., Kumarb, V. and Kumarb, U. (2012), "Relationship between quality management practices and innovation", *Journal of Operations Management*, Vol. 30, pp. 295-315.
- Kim, J. and Sung, Y. (2009), "Dimensions of purchase-decision involvement: affective and cognitive involvement in product and brand", *Journal of Brand Management*, Vol. 16 No. 8, pp. 504-519.
- King, M. and Balasubramanian, S. (1994), "The effects of expertise, end goal, and product type on adoption of preference formation strategy", *Journal of Academy Marketing Science*, Vol. 22 No. 2, pp. 146-159.
- Laurent, G. and Kapferer, J.N. (1985), "Measuring consumer involvement profiles", *Journal of Marketing Research*, Vol. 22, pp. 41-53.
- Leary, M.R. and Allen, A. (2011), "Self-presentational persona: simultaneous management of multiple impressions", *Journal of Personality and Social Psychology*, Vol. 101, pp. 1033-1049.
- León, F.R., Quiroz, G. and Brazzoduro, A. (1994), "The reliability of simulated clients' quality-of-care ratings", *Studies in Family Planning*, Vol. 25 No. 2, pp. 184-190.
- Levy, J.M. (1988), "The influence of sex roles on judgment", *Journal of Consumer Research*, Vol. 14, pp. 522-530.
- Luria, G., Yagil, D. and Gal, I. (2014), "Quality and productivity: role conflict in the service context", *The Service Industries Journal*, Vol. 34 No. 12, pp. 955-973.
- McDougall, H.G. and Levesque, T.J. (1994), "Benefit segmentation using service quality dimensions: an investigation in retail banking", *International Journal of Bank Marketing*, Vol. 12 No. 2, pp. 15-23.
- Madupalli, R.K. and Poddar, A. (2014), "Problematic customers and customer service employee retaliation", *Journal of Services Marketing*, Vol. 28 No. 3, pp. 7-17.
- MaGee, G.W., Ferguson, C.E. and Seers, A. (1989), "Role conflict and role ambiguity: do the scales measure these two constructs?", *Journal of Applied Psychology*, Vol. 74, pp. 815-818.
- Markus, H. and Nurius, P. (1986), "Possible selves", *American Psychologist*, Vol. 41, pp. 954-969.
- Martire, L.M., Stephens, M.A.P. and Townsend, A.L. (2000), "Centrality of women's multiple roles: beneficial and detrimental consequences for psychological well-being", *Psychology and Aging*, Vol. 15 No. 1, pp. 148-156.
- Morrison, L.J., Andrew, C.M. and Preston, C.C. (1997), "Mystery customer research: cognitive processes affecting accuracy", *Journal of the Market Research Society*, Vol. 39 No. 2, pp. 349-361.
- Nathanail, E. (2008), "Measuring the quality of service for messengers on the Hellenic railways", *Transportation Research*, Vol. 42, pp. 48-66.
- Oliver, R. (1999), "What is consumer loyalty", *Journal of Marketing*, Vol. 63, pp. 33-44.
- Otnes, C., Lowrey, T.M. and Shrum, L.J. (1997), "Toward an understanding of consumer ambivalence", *Journal of Consumer Research*, Vol. 24, pp. 80-93.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), "Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria", *Journal of Retailing*, Vol. 70 No. 3, pp. 201-230.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "Servqual", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.
- Pardo-del-Val, M., Martínez-Fuentes, C., López-Sánchez, J.I. and Rata, B.M. (2014), "Franchising: the dilemma between standardization and flexibility", *The Service Industries Journal*, Vol. 34 Nos 9/10, pp. 828-842.
- Raggatt, P.T.F. (2000), "Mapping the dialogical self: towards a rationale and method of assessment", *European Journal of Personality*, Vol. 14, pp. 65-90.
- Raman, N.V., Chattopadhyay, P. and Hoyer, W.D. (1995), "Do consumers seek emotional situations: the need for emotion scale", *Advances in Consumer Research*, Vol. 22, pp. 537-537.
- Ranjan, K.R., Sugathan, P. and Rossmann, A. (2015), "A narrative review and meta-analysis of service interaction quality: new research directions and implications", *Journal of Services Marketing*, Vol. 29 No. 1, pp. 3-14.



- Reynolds, F.D. and Darden, W.R. (1971), "Mutually adaptive effects of interpersonal communication", *Journal of Marketing Research*, Vol. 8 No. 4, pp. 449-454.
- Rust, R.T. and Oliver, R.L. (1994), "Service quality: insights and managerial implications from the frontier", *Service Quality: New Directions in Theory and Practice*, Sage Publications, New York, NY, pp. 1-19.
- Sadowski, C. and Cogburn, H.E. (1997), "Need for cognition in the big five factor structure", *The Journal of Psychology*, Vol. 131 No. 3, pp. 307-312.
- Schaninger, C. and Sciglimpaglia, M.D. (1981), "The influence of cognitive personality traits and demographics on consumer information acquisition", *Journal of Consumer Research*, Vol. 8 No. 2, pp. 208-216.
- Seth, N., Deshmukh, S.G. and Vrat, P. (2005), "Service quality models: a review", *International Journal of Quality and Reliability Management*, Vol. 22 No. 9, pp. 913-949.
- Sharma, P., Chen, I.S. and Luk, S.T. (2012), "Gender and age as moderators in the service evaluation process", *Journal of Services Marketing*, Vol. 26 No. 2, pp. 102-114.
- Shepherd, C.D. and Fine, L.M. (1994), "Scaling and measurement: role conflict and role ambiguity reconsidered", *Journal of Personal Selling & Sales Management*, Vol. 14 No. 2, pp. 57-65.
- Solomon, M.R., Surprenant, C., Czepiel, J.A. and Gutman, E.G. (1985), "A role theory perspective on dyadic interactions: the service encounter", *The Journal of Marketing*, pp. 99-111.
- Spector, P.E. (1992), *Summated Rating Scale Construction An Introduction*, SAGE Publication, Thousand Oaks, CA.
- Suquet, J.B. (2010), "Drawing the line: how inspectors enact deviant behaviors", *Journal of Services Marketing*, Vol. 24 No. 6, pp. 468-475.
- Tarantola, C., Vicard, P. and Ntzoufras, I. (2012), "Monitoring and improving Greek banking services using Bayesian networks: an analysis of mystery shopping data" *Expert Systems with Applications*, Vol. 39, pp. 10103-10111.
- Trochim, W.M. and Donnelly, J.P. (2007), *Research Methods Knowledge Base*, 3rd ed., Atomic Dog Publishing, NJ.
- Turner, R.H. (1978), "The role and the person", *American Journal of Sociology*, Vol. 84 No. 1, pp. 1-23.
- Urban, W. (2013), "Perceived quality versus quality of processes: a meta concept of service quality measurement", *The Service Industries Journal*, Vol. 33 No. 2, pp. 200-217.
- Urda, J. and Loch, C.H. (2013), "Social preferences and emotions as regulators of behavior in processes", *Journal of Operations Management*, Vol. 31, pp. 6-23.
- Van Eerde, W. and Peper, P. (2008), "Deviant service behavior: coming soon to a theatre near you?", *The Service Industries Journal*, Vol. 28 No. 1, pp. 85-93.
- Wagstaff, K., Cardie, C., Rogers, S. and Schrödl, S. (2001), "Constrained k-means clustering with background knowledge", *ICML*, Vol. 1, pp. 577-584.
- Wang, E.S.T., Tsai, B.K., Chen, T.L. and Chang, S.C. (2012), "The influence of emotions displayed and personal selling on customer behavior intention", *The Service Industries Journal*, Vol. 32 No. 3, pp. 353-366.
- Wiele, T., Hesselink, M. and Iwaarden, J. (2005), "Mystery shopping: a tool to develop insight into customer service provision", *Total Quality Management*, Vol. 16 No. 4, pp. 529-541.
- Wilson, A.M. (1998), "The use of mystery shopping in the measurement of service delivery", *The Service Industries Journal*, Vol. 18 No. 3, pp. 148-163.
- Wilson, A.M. (2001), "Mystery shopping: using deception to measure service performance", *Psychology & Marketing*, Vol. 18 No. 7, pp. 721-734.
- Wong, M.L., Chan, R. and Tan., H.H. (2012), "Sex work and risky sexual behaviors among foreign entertainment workers in urban Singapore: findings from mystery client survey", *Journal of Urban Health*, Vol. 89, pp. 1031-1044.
- Yanof, T.G. (2012), "Paradoxes of rational choice theory", in Roeser, S., Hillerbrand, R., Sandin, P. and Peterson, M. (Eds), *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, Springer, pp. 499-516.
- Young, G.J., Meterko, M.M., Mohr, D., Shwartz, M. and Lin, H. (2009), "Congruence in the assessment of service quality between employees and customers: a study of a public health care delivery system", *Journal of Business Research*, Vol. 62, pp. 1127-1135.
- Zaichkowsky, J.L. (1994), "The personal involvement inventory: reduction, revision, and application to advertising", *Journal of Advertising*, Vol. 23 No. 4, pp. 59-70.
- Zhanga, D., Linderman, K. and Schroeder, R.G. (2012), "The moderating role of contextual factors on quality management practices", *Journal of Operations Management*, Vol. 30, pp. 12-23.

## Further reading

- Hair, J., Anderson, R., Tatham, R. and Black, W. (1995), *Multivariate Data Analysis*, 4th ed., Prentice Hall International, Upper Saddle River, NJ.
- Iacobucci, D., Ostrom, A. and Grayson, K. (1995), "Distinguishing service quality and customer satisfaction: the voice of the consumer", *Journal of Consumer Psychology*, Vol. 4 No. 3, pp. 277-303.
- Surprenant, C.F. and Solomon, M.R. (1987), "Predictability and personalization in the service encounter", *Journal of Marketing*, Vol. 51, pp. 86-96.

## Appendix

### Mystery client script

- "Good morning" (greetings).
- "I am interested in a Mp3/4 or digital camera" (Then wait until the employee asks you something).
- "I just want it to download music and videos", (MP3/4) or "I just want to take pictures, no hassles", (digital camera). Never express preference for any specific brand or format. Always ask to see more products.
- If you are asked about price range, tell them you "want the less expensive" – if they do not ask, do not tell. Just ask to see more types.
- Never express or provide any clue about your preferences or technical specification whatsoever (keep mentioning what is pinpointed in Point 3.).



- After having seen several models/brands – and when, apparently, there is no more left to see – say this: “Which brand or model do you advise me to buy or do you thing fits my goals?”
- Listen his/her answer and then reply: “My cousin (or friend, brother) told me that this brand [. . .] at [. . .]. (competitor retail chain) is cheaper!!”
- Speaking to your colleague (make sure the employee listen the dialogue too) say this “ Do you know what happened to my friend Jo, his Mp3/4 or digital Camera was under warranty time and when it broke down, he brought the device here at the service repair and they complicated everything [. . .]”. Then listen the reaction and do not add more.
- “Is it possible to inform me when a more appropriate model is launched?”
- “Is it possible to let me know when there is a promotion of that brand/model?”
- Ask for a brochure – if they do have it ask a photocopy of technical specifications to analyze it at home.
- Ask the seller for his/her business card.
- “Thanks” (good-by salutation).

#### Store/salesperson evaluation (attributes)

Note: the digits indicate the corresponding order of the measured attribute shown on Table VI:

- 1 Assurance (semantic differential seven-point scale):
  - Technically incompetent [. . .] competent
  - Trustful [. . .] untruthful
  - Confuse or muddy in his/her presentation [. . .] clear/enlighten
  - Commitment [. . .] irresponsible/detachment
  - Proactive [. . .] reactive
- 2 Empathy (seven-point scale: 1 = *It does not apply at all* [. . .] to 7 = *Applies completely*):
  - He/she politely interacted with the customer

- He/she reveled pleasure/happiness in pursuing his/her job
  - He/she seemed to develop their arguments in an honest/sincere way
  - His/her behavior (verbally and non-verbally) seemed to be ethically correct
  - He/she took seriously and showed interest/concern to help the customer
  - He/she made their best to understand customer’s needs
  - He/she seemed to be selfishly interested in selling instead of helping the customer (r)
  - He/she is truly concerned in helping the client
  - He/she listened attentively what we wanted
  - He/she was serene/calm in finding the solution
- 3 Tangible evidence – people (semantic differential seven-point scale):
    - Employee well dressed [. . .] dowdy
    - Well shaved/proper make-up [. . .] careless
  - 4 Tangible evidence – place:
    - Dirty [. . .] clean
    - Attractive window-shopping [. . .] uninteresting
    - Easy to find the products [. . .] difficult
    - Messy [. . .] tidy
    - Easy access to products [. . .] physically inaccessible
    - Poor lighting [. . .] good lighting
  - 5 Reliability (semantic differential seven-point scale):
    - Efficient (quick response) [. . .] inefficient (slow)
    - Unconvincing [. . .] very persuasive
    - Offer accurate information [. . .] inaccurate

#### Corresponding author

Pedro Quelhas Brito can be contacted at: [pbrito@fep.up.pt](mailto:pbrito@fep.up.pt)

**This article has been cited by:**

1. Davis Kathryn Simons, Kathryn Simons Davis, Mohan Mayoor, Mayoor Mohan, Rayburn Steven W., Steven W. Rayburn. 2017. Service quality and acculturation: advancing immigrant healthcare utilization. *Journal of Services Marketing* 31:4/5, 362-372. [\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)