

Experimental Mystery Shopping

Are Businesses Spending Their Research Dollars Wisely?

By: Perceptive Sciences Corporation

As defined by Alan Wilson (2001), mystery shopping is “a form of participant observation [that] uses researchers to deceive customer-service personnel into believing that they are serving real customers or potential customers.” Like characters in a detective story, mystery shoppers can be viewed as covert agents that solve crimes and benefit the greater good. In this case they uncover poor customer service and poor customer relations, which when remedied improve the shopping experience for consumers at large.

Business executives have clearly adopted this romantic picture of mystery shopping. Approximately \$1 billion dollars is being spent on mystery shopping annually worldwide (Maret, 2005). However, the critical question addressed in this white paper is: are businesses spending their dollars wisely?

Additionally, other questions of interest in addition to the above stated critical question addressed by this paper are:

- Is mystery shopping a reliable technique for collecting data?
- What are the drawbacks of using mystery shoppers?
- How can mystery shopping improve to provide valid and reliable data?

Businesses desire that their employees treat all customers according to company policy. They have sought mystery shopping in droves as a “tool to motivate personnel” and as a “diagnostic tool” that would allow them to identify the strengths and weaknesses of their service model (Wilson, 2001). Monitoring personnel performance is critical because “ninety percent of unhappy customers leave a place of business because of inattentive, impolite employees” (Semenak, 2005).

Mystery shopping may indeed be able to provide some business solutions. However, mystery shopping must use experimental and statistical techniques to produce reliable data and must account for normal cognitive limitations and biases. This requires that researchers design and manage mystery-shopping programs carefully.

“Mystery shopping must use experimental and statistical techniques to produce reliable data”

First, mystery shopping will fail to produce reliable data if researchers engage in poor population sampling from the population of stores or customers. In behavioral market research, it is critical to both sample randomly and to sample from the population of interest. This ensures basic statistical assumptions are met, which allows for generalization from the representative sample to the entire population.

Imagine a situation in which these techniques are not applied. An individual with free time applies on-line to be a paid mystery shopper and chooses an assignment, thereby self-selecting instead of being randomly selected. This individual is sent to a garden center to mystery shop despite no gardening experience, which suggests this individual falls outside of the customer population. Therefore, this mystery shopper will not provide any data one could use to determine the perceptions and desires of an average garden center customer.

Second, mystery shoppers are a unique population. These individuals have the time and inclination to be mystery shoppers. By definition, they are willing to deceive store owners and employees by pretending to be customers. In addition, mystery shoppers know what they are doing and why. They understand the study objectives and therefore can knowingly or unknowingly bias their responses to be consistent with the study goals. Therefore, there is great potential for collecting biased data.

Although this might not be as critical for determining how many minutes passed between the mystery shopper entering the store and being greeted by an employee, it is extremely critical for answering questions about customer preferences. Researchers have demonstrated that individuals are bad at predicting their future preferences when they are forced to provide justifications for their choices and furthermore that people do not seem to know why they feel the way that they do (Wilson, Lisle, Schooler, Hodges, Klaaren, & LaFleur, 1993). This suggests that more implicit or indirect methods are needed to assess preferences.

Third, the mystery-shopping scenario lacks external validity. Real customers are not neutral observers. For example, they have goals, time constraints, and expectations. Imagine that a bank uses mystery shopping to determine how long the teller lines are at noon and how pleased or displeased the customers are with the wait times. A mystery shopper being paid by the hour is likely to report less frustration with a long wait than an employee trying to deposit a paycheck during lunch.

Fourth, the average mystery shopping procedure requires extraordinary memory -much beyond the capabilities of normal individuals. Shoppers spend an average of 10-15 minutes in an assigned store and then go home to write a report detailing the mystery shopping experience that takes approximately 10-15 minutes to complete (Semenak, 2005).

“The likelihood of a mystery shopper producing a delayed accurate detailed report is very low”

Although this procedure may seem reasonable, cognitive psychologists have demonstrated that memory is reconstructive (Koriat, Goldsmith, & Pansky, 2000). Memory can be influenced when the information is encoded, stored, or retrieved. Information can be added or subtracted based on previously known or intervening information. Essentially, individuals typically just remember the basic gist of a situation. As such, the likelihood of a mystery shopper producing a delayed accurate detailed report is very low.

Lastly, and worst of all, mystery shopping lacks experimental control. Because the research is observational and correlational, it is impossible to determine what variable may actually be causing changes. For example, say a business gets a mystery shopping report stating that employees do not smile enough and the business implements a new company policy that requires employees to smile more. Then, a future mystery shopping report reveals that employees seemed happy and were smiling. One may want to assume that the new policy caused the change in employee behavior. However, because of the lack of experimental control, there is no way to know what caused the change.

Perceptive Sciences Corporation is a science based market research, user interface design, and user testing firm, employing experts in the fields of cognitive psychology, information sciences and human factors studies. Perceptive Sciences serves best-in-class technology based companies and market leaders in a wide range of industries in the U.S. and Europe.