

Applying Fuzzy Multiple Criteria Decision-Making Method and Service Failure Index to Evaluate Restaurants' Performance

Hsi-Tien Chen¹

¹ Department of Leisure Industry Management, National Chin-Yi University of Technology, Taiwan

Correspondence: Hsi-Tien Chen, Department of Leisure Industry Management, National Chin-Yi University of Technology, Taiwan. E-mail: tien7845@gmail.com

Received: July 27, 2014

Accepted: August 24, 2014

Online Published: October 16, 2014

doi:10.5430/jms.v5n4p14

URL: <http://dx.doi.org/10.5430/jms.v5n4p14>

Abstract

When a severe service failure occurs, loss to the firm is generally inevitable, but little research has been conducted on how to control the frequency of severe service failure. The purpose of this study is to propose a service failure control process and model by assessing whether the failure frequency in various service failure categories meets minimum tolerance levels. An initial restaurant-related service failure classification model was constructed based on literature review and expert suggestions. The present study applies the fuzzy Delphi Method to screen these initial restaurant-related service failure categories to produce a formal restaurant-related service failure classification model. Next, a questionnaire survey is used to collect data of the frequency and severity of service failure categories. The tolerable upper limits for the frequency of service failures are determined according to the severity of each category. Service failure indices are defined, and a statistical method is used to test whether failure indices values of various categories meet the required level for the business. An illustrative example of a restaurant chain is used to demonstrate the process of implementing the service failure control model. The results show that the restaurant had seven unqualified failure categories and corresponding priorities for improvement, which can help to establish a service quality improvement strategy and resource allocation plan. This study offers a new and progressive way of thinking about service failure control for industry and academic.

Keywords: service failure, service recovery, fuzzy Delphi method, multiple criteria decision making, failure severity, service failure control model

1. Introduction

Customers are regarded as assets of an enterprise, and their satisfaction directly affects profits (Fornell, 1992). Therefore, satisfying customers' needs is crucial for any enterprise. Because service possesses intangible, inseparable, variable, and perishable characteristics (Fisk, Brown, & Bitner, 1993), even if enterprises have comprehensive service plans and rigid quality control, service failures remain inevitable (Johnston & Hewa, 1997). When a service failure occurs, not only does it affect customer satisfaction regarding the particular service, but it can also influence the customer's overall evaluation on the enterprise, resulting in negative effects such as negative word-of-mouth and loss of sales, which can further lead to irreparable damage to the image of the enterprise (Spreng, Harrell, & Mackoy, 1995; Priluck, 2003).

When service failures occur, the enterprise should restore customer service as soon as possible. Even so, not all service recovery actions can completely satisfy the customers. If service recovery is executed improperly, it may exacerbate customer discontent, leading to even greater losses (Smith, Bolton, & Wagner, 1999). Hart, Heskett and Sasser (1990) found that over 50% of enterprises have intensified negative customer impressions because of improper handling of service failures. Service failure and inappropriate service recovery are critical reasons customers switch providers (Johnston & Hewa, 1997). McCollough, Berry and Yadav (2000) found even if practitioners take appropriate service recovery actions to restore customer satisfaction and loyalty, customer satisfaction is lower than before the service failure. Maxham and Netemeyer (2002) suggested that enterprises should succeed in providing service the first time and properly compensate customers for the service failure. Therefore, the importance of assessing and managing service failures may be similar to service recovery.

Numerous studies related to service failures and service recovery have been conducted, however, most of the studies focused on classifying service failure and recovery (e.g., Bitner, Booms, & Tetreault, 1990; Hoffman, Kelley, &

Rotalsky, 1995; Lin, Huang, & Huang, 2003), exploring customer satisfaction following service failure and recovery (e.g., Maxham & Netemeyer, 2002; McCollough et al., 2000), identifying the role of equity in service failure and recovery (e.g., Smith, et al. 1999), clarifying the relationships between service failure, service recovery, and other variables (e.g., Ha & Jang, 2009), providing cross-cultural comparisons of service failure and service recovery strategies (e.g., Mueller, Palmer, Mack, & McMullan, 2003) and so forth. In particular, the critical incident technique (CIT) is a qualitative research method that has been widely used in service failure and service recovery classification research (e.g., Hoffman et al., 1995; Mueller et al., 2003). However, qualitative research does not provide quantitative information about the customers, but explores their emotions and motivations (Dickens, 1987). Both quantitative and qualitative data may be complementary. Grounded theory methodology may include these two types of data. Most importantly, different types of data on the same subject when compared are conducive to generating theory (Glaser & Strauss, 1967). It evokes the research motivation of this paper.

Based on the above discussion, we believe the severity of service failure differs according to the failure categories, and each category can lead to a different degree of loss in customer satisfaction and business performance. Based Pareto's 80/20 rule, practitioners should pay special attention to assessing and managing severe service failures to ensure quality service and customer satisfaction and enhance business performance. However, relevant research on how to assess and manage the frequency of service failure in different categories according to the severity of service failure is rare. In addition, past studies on service failure paid little attentions to uncertain and vague issue about customer opinions. To address the issues, this study integrated Fuzzy theory approaches and developed a testing process of service failure to remedy the shortage in previous service failure studies. It can objectively and effectively evaluate and manage service performance and provide a monitoring method of service failure frequency according severe degrees in various failure categories.

This study uses literature review to construct an initial restaurant-related service failure classification model. Based on fuzzy multiple criteria decision-making approaches, we apply fuzzy Delphi method to screen and develop a formal restaurant-related service failure classification model. Furthermore, fuzzy theory is used to calculate the severity of each failure category, and a tolerable upper limit of failure frequency for each failure category can then be identified according to the severity of service failure. Ultimately, this study defines a service failure index and uses a statistical testing method (or control chart of service failure) to identify whether the frequency of the various failure categories in restaurants are qualified. The results can serve as a reference for the restaurant industry and academia. Also, this approach can serve as an alternative method of service quality control and evaluation of customer satisfaction.

2. Literature Review

2.1 Classification, Frequency and Severity of Service Failures

Bitner et al. (1990) stated that service failures occur when core services are delivered at a level below customers' minimum expectations. To develop a "service failure classification," they examined customers of the airline, hotel, and restaurant industries and found that service failures can typically be divided into three groups ("service delivery system failures," "customer needs and requests," and "unprompted/unsolicited employee action") and 12 categories. Hoffman et al. (1995) referenced the framework of Bitner et al. (1990) and divided service failures for restaurants in the United States into three groups and 14 categories. Following the same framework, Lin et al. (2003) divided service failure for restaurants in Taiwan into three groups and 22 categories. These studies classified service failures and recovery strategies from the customers' perspectives. Customer satisfaction may vary with different service failure categories. Subsequently, some studies such as Mueller et al. (2003), Ha and Jang (2009), and Chen (2013) used the concept of service failure classification in different research.

Smith et al. (1999) stated that, in addition to being influenced by various service failure categories, customer satisfaction is affected by the severity of service failure, based on resource exchange principles. The greater the severity of service failure implies that the greater the likelihood that the customer will respond negatively. Kelley and Davis (1994) thought that service failures may vary considerably across the dimensions of timing, severity and frequency. Hence, customer satisfaction may also be affected by frequency of service failure. That is to say, the frequency and severity of service failure are both quite important factors affecting customer satisfaction except for the timing of service failure occurrence. When a service failure occurs, the severity of the failure will determine the level of recovery required to restore customers' satisfaction. In a word, different categories and severities of service failures result in different levels of losses for enterprises, as well as the higher the frequency of service failure, the larger the enterprises' profit loss is. The enterprise managers should pay more attentions to frequent and severe failure categories. We believe that the frequency of service failures should be evaluated and managed according to

the corresponding severity of failure categories in order to identify whether reach the levels of quality service.

2.2 Fuzzy Theory

Human thinking and perception of things around us is rather vague. Fuzzy set theory, introduced by Zadeh (1965), provides a means for representing uncertainties and dealing with problems in a vague environment. This theory transforms linguistic variables into fuzzy sets to replace the crisp set. The values of linguistic variables are not numbers but words or sentences in a natural or artificial language. The concept is very useful in situations that are too complicated or ill-defined to be appropriately described by traditional quantitative expressions (Zadeh, 1975). Fuzzy theory has been successfully applied in various fields (i.e., Hsia, Chen, & Chen, 2008; Tang, Tzeng, & Wang, 1999). For example, “severity,” “possibility,” “importance,” and “satisfaction” are used popularly, but their meaning varies with the situations.

In this study, customer opinions were investigated to rank and calculate the severity of service failure in all categories. Because arithmetic operation of a triangular fuzzy number is more easily calculated and understood, especially in fuzzy multiplication (Chen, 2000), the values of linguistic variables were defined by a triangular fuzzy number, as shown in Table 1. Since fuzzy arithmetic means are not crisp values, they cannot be ranked. Thus, fuzzy ranking methods were employed, referred to as ‘defuzzification,’ to obtain non-fuzzy values. In general, the center of area (COA) method is an easy and efficient method (Tang et al., 1999), thus, we used it to calculate non-fuzzy severities for all failure categories.

2.3 Fuzzy Delphi Method

Linstone (1985) demonstrated that the Delphi method, designed by Dalkey and Helmer (1963), is a subjective prediction method based on expert opinions and selecting useful data from questionnaires and meetings. However, related studies (e.g., Hsu, 1998; Tsai & Chen, 2011) indicated that the Delphi method has some shortcomings. To overcome these shortcomings, Murray, Pipino and Gigch (1985) were the first to integrate fuzzy theory into the Delphi method. Subsequently, Ishikawa et al. (1993) employed the concept of cumulative frequency distribution and fuzzy integrals to integrate expert opinions into fuzzy numbers and introduced the fuzzy Delphi method. Hsu (1998) claimed that the fuzzy Delphi method has the following advantages compared to Delphi method: (1) it reduces investigation time and costs; (2) individual experts’ opinions can be clearly expressed without distortion; (3) the semantic structure of the predicted items can be clearly expressed; (4) the fuzziness during the investigation process is considered; and (5) the calculation process is simple and can address multi-level, multi-attribute, and multi-scheme decision-making problems. Since then, the fuzzy Delphi method has been widely applied (e.g., Horng, Hsu, Liu, Lin, & Tsai, 2011; Tsai & Chen, 2011). Therefore, we adopted the fuzzy Delphi method with triangular fuzzy numbers to facilitate the screening and development of a restaurant-related service failure classification framework in this study. The method possesses higher stability and representativeness even while small sample was collected (Hsu, 1998).

Table 1. The scaling values for triangular fuzzy numbers of linguistic variables

The values of linguistic variables	Likert 7-point scale	Triangular fuzzy numbers
Extremely non-severe (Extremely impossible)	1	(0, 0, 0.1)
Non-severe (Impossible)	2	(0, 0.1, 0.3)
Slightly non-severe (Slightly impossible)	3	(0.1, 0.3, 0.5)
Ordinary severe (Ordinary possible)	4	(0.3, 0.5, 0.7)
Slightly severe (Slightly possible)	5	(0.5, 0.7, 0.9)
Severe (Possible)	6	(0.7, 0.9, 1.0)
Extremely severe (Extremely possible)	7	(0.9, 1.0, 1.0)

3. Research Methods

3.1 Development of a Restaurant-related Service Failure Classification Framework

Service failure is defined in this paper as “items identified by consumers as unsatisfactory when consumers believe that the services or products rendered by the restaurant are substandard.” To ensure the comprehensiveness of the service failure groups and categories for the restaurant industry, this study used the service failure classification framework proposed by Bitner et al. (1990), Hoffman et al. (1995) and Lin et al. (2003) as a classification basis. Also, the restaurant scenarios of this study were considered, and two scholars with extensive practical experience in

restaurant-related fields were invited to engage the reviewing jobs of failure categories. After discussing and receiving revised suggestions from the scholars, the failure categories were reorganized, summarized, merged, and amended. Finally, an initial restaurant-related service failure classification model was developed and presented in Table 2. It comprises three major groups and 27 failure categories: service delivery system failures (14 categories), customer needs and requests (4 categories) and unprompted/unsolicited employee action (9 categories).

3.2 Questionnaire Design, Sampling and Survey

The questionnaire design, sampling and survey were divided into two stages. In the first stage, an expert questionnaire was designed and investigated, and subsequently, the fuzzy Delphi method was used to screen the initial restaurant-related service failure categories. In the second stage, a consumer questionnaire was designed and investigated according to a formal restaurant-related failure classification model in the first stage. They were presented as below.

3.2.1 Expert Questionnaire Design and Survey

Based on the “initial restaurant-related service failure classification model (Table 2),” a draft questionnaire with 27 failure categories (or items) was designed. The questionnaire includes two parts. The first part comprises the demographic information (gender, age, profession, ranking and the years of service), including both nominal and ordinal values. The second part requests the likelihood of occurrence of all service failure categories. These items were measured on a 7-point Likert scale (1=“Extremely impossible”~7=“Extremely possible”). We invited two scholars with over ten years of restaurant-related working experience to proof the descriptions and scenarios of the various failure categories in draft questionnaire, and provide suggestions for modification of ambiguous or hard-to-understand questions. To avoid missing critical restaurant-related failure categories, the questionnaires were semi-open-ended. At the end of questionnaires, two spaces were provided for respondents (i.e., experts) to fill in additional categories they believed to be necessary and critical. An expert questionnaire was then obtained.

Tsaur, Lin and Lin (2006) stated that the most crucial step in the Delphi method is screening respondents. The number of samples for the Delphi method should be between 15 and 30 (Delbecq, Van de Ven, & Gustafson, 1975). Along the similar line, in order to ensure the representativeness of expert decision-making group, we invited 10 professional restaurant managers and 10 professional university teachers in the restaurant field to complete the expert questionnaire. All of them possess over six years of restaurant-related managing or teaching experiences. These experts provided their impression of the likelihood of occurrence of all service failure categories using a 7-point Likert scale (Table 1). In order to construct a formal failure classification framework, fuzzy Delphi method was used to screen the initial service failure categories in Section 3.3.

3.2.2 Consumer Questionnaire Design and Survey

Next, we designed a questionnaire to investigate restaurant consumer perceptions on the severity of service failure based on the formal restaurant-related failure classification model. The questionnaire was divided into three parts. The first part is the consumer behavior survey (i.e., *level of satisfaction, repurchase intention, and willingness to recommend*) using a 7-point Likert scale. The second part comprised the service failure categories selection (i.e., tick) and corresponding severity cognition survey. Because this study assumes each valid respondent only encounter an incident (or failure category), the respondents can only select one most impressive incident (or failure category). We provided a space for the respondents to fill in a service failure category that was not included in the formal restaurant-related failure classification model and to tick its severity on a 7-point Likert scale (1=“Extremely non-severe”~7=“Extremely severe”). The third part requested demographic information, including gender, age, education level and profession.

Table 2. Definition and measurement of initial restaurant-related service failure classification model

Failure groups	Definition	Failure categories	References
Group 1: Service delivery system failures	Service failures caused by service staff's inability to respond immediately to customer complaints or dissatisfactions caused by service delivery system failure.	1.Food and beverage quality problems	2, 3
		2.Fail to receive proper services	1, 2
		3.Slow service	1, 2, 3
		4.Missing food or beverage items	3
		5.Equipment failure	2, 3
		6.Tableware problems	3
		7.Environmental problems	3
		8.Sold out or could not provide	1, 2, 3

		food or beverage items	
		9. Policy cognition differences	2, 3
		10. Failure to provide services in the order of customer arrival	3
		11. Booking mistakes	3
		12. Pre-service preparation and meal ordering failures	3
		13. Service failure during meals	3
		14. Billing mistakes	3
		15. Seating arrangement or replacement mistakes	2, 3
Group 2: Customer needs and requests	Service failures caused by the inability for the service delivery system to immediately respond and meet a customer's special needs or to provide customized services as requested by the customer.	16. Cooking procedure errors	1, 2
		17. Request to change dish varieties	1, 3
		18. Failure to comply with special requests from the customer	1
		19. Unexpected employee behavior	2
		20. Ordered the wrong dish	2
		21. Delivered the wrong dish	2, 3
Group 3: Unprompted/unsolicited employee action	Service failures caused by staff behavior that was not expected by the customers	22. Billed for the wrong dish	2
		23. Food and beverage spillage	3
		24. Lack of professional knowledge and technique	3
		25. Poor staff hygiene	3
		26. Poor service attitude	3
		27. Deceiving customers	3

1= Bitner et al. (1990); 2=Hoffman et al. (1995); 3= Lin et al. (2003)

Because respondents' recollection abilities vary, we limited the respondents to consumers over the age of 15 who had experienced service failures at a restaurant within the past year. The official survey lasted for approximately 2 months. Because random sampling of valid respondents was difficult, we referenced similar research, such as Hoffman et al. (1995), and utilized the purposive and snowball sampling method to reduce the possibility of research deviation. Trained interviewers were positioned at the entrance or parking lot of the case restaurants to search for valid respondents, or valid respondents were introduced by the restaurant staff or drawn from respondents who completed questionnaires. Small gifts were used as incentives to enhance respondents' willingness to participate. If a valid customer was willing to help, the interviewers gave the customer a questionnaire and explained the fill-in method.

3.3 Construction of a Formal Restaurant-related Service Failure Classification Model

To construct a formal restaurant-related service failure classification model, this study adopted the fuzzy Delphi method with triangular fuzzy numbers (Hsu, 1998) to screen the initial restaurant-related service failure categories. The implementation steps of fuzzy Delphi method are presented as follows.

- (1) Collect opinions from an expert decision-making group: This study conducted an expert questionnaire with a 7-point Likert scale to assess the likelihood of occurrence of all service failure categories in Table 2.
- (2) Transform the expert opinions into triangular fuzzy numbers: Table 1 shows how to transform the expert opinions into triangular fuzzy numbers (Chen, 2000). Assuming the triangular fuzzy number of j th failure category for i th expert is $\tilde{w}_j = (L_{ij}, M_{ij}, U_{ij})$. Integrating the triangular fuzzy numbers of j th failure category of l experts, we have

triangular fuzzy number of j th failure category, $\tilde{w}_j = (L_j, M_j, U_j)$, where

$$L_j = \text{Min}_i \{L_{ij}\}, \quad M_j = \left(\prod_{i=1}^l M_{ij} \right)^{1/l}, \quad U_j = \text{Max}_i \{U_{ij}\}, \quad i = 1, 2, \dots, l; j = 1, 2, \dots, k.$$

- (3) Conduct a defuzzification process: Fuzzy numbers cannot be ranked since they are not crisp values. This study used a defuzzification method, the graded mean integration method proposed by Chen and Hsieh (2000), to calculate the non-fuzzy likelihood of occurrence for each failure category. The defuzzification function is $w_j = (L_j + 4M_j + U_j)/6, j = 1, 2, \dots, k$.
- (4) Screen service failure categories: A screening threshold should be set up. Since there is no standard for setting the threshold, it is set based on the needs of the study (Hsu, 1998). Based on expert suggestions and the desire to identify as many service failure categories as possible, this study established a screening threshold as 0.4. When the non-fuzzy occurrence possibility of a failure category is larger than 0.4, the failure category is retained; otherwise, it is abandoned.
- (5) A formal restaurant-related service failure classification model was constructed based on the retained failure categories.

3.4 Measurement of Service Failure Severity

A total of n valid respondents (or consumers) were asked to select (or tick) the most impressive service failure category (or incident) that they had encountered and rate the corresponding severity in the consumer questionnaire. Next, Table 1 was used to convert the failure severities to triangular fuzzy numbers. The evaluations and opinions of all respondents were integrated into triangular fuzzy weights using the arithmetic mean. Finally, the graded mean integration method proposed by Chen and Hsieh (2000) was utilized to calculate the non-fuzzy weights (or severities) of the various service failure categories.

3.5 Construction, Estimation and Testing of Service Failure Indices

3.5.1 Statistical Theories of Service Failure Frequency

Let F_j represent the frequency of service failures for j th failure category over a period of time. Thus, F_j conforms to a *Non-homogeneous Poisson Distribution* (Ross, 2010). The probability function of F_j can be defined as below.

$$P(F_j = f_j) = \frac{e^{-\lambda_j} \lambda_j^{f_j}}{f_j!}, f_j = 0, 1, 2, 3, \dots; j = 1, 2, \dots, k, \quad (1)$$

where j denotes j th failure category, λ_j is the mean and variance of failure frequencies for j th failure category (i.e., $\lambda_j = E(F_j) = Var(F_j)$). Obviously, the smaller the frequency (F_j) of service failures, the better the service performance of j th failure category.

3.5.2 Tolerable Upper Limit of Failure Frequency for Failure Categories

The higher the severity of service failures, the larger the enterprises' performance loss is, and meanwhile, the lower the tolerable frequency of service failures is. We assume that each customer encounters at most one impressive service failure per transaction in this paper. Because the estimation of total service frequency (N) is not easy in practice, the number of customers is used to estimate service frequency in this paper. Moreover, the tolerable upper limit (V_j) of service failure frequency of j th failure category can be defined according the corresponding severity of j th failure category. If the severity of service failure is larger, then V_j will be set to be smaller. Under the assumption of this study (i.e., using the number of customers to estimate service frequency), V_j also varies with the sample number of respondents (i.e., sample size, n). Because each valid respondent must encountered an impressive service failure, the sample size (n) can also be regarded as the total sample number of service failures over a period of time. Therefore, V_j can be defined as below.

$$V_j = \frac{1/w_j}{\sum_{j=1}^k 1/w_j} \times n, \quad (2)$$

where w_j is the severity (or weight) of j th failure category and n denotes the sample size.

3.5.3 Definition, Estimation and Testing of Service Failure Indices

A service failure index (SFI_j) of the j th failure category is defined as follows:

$$SFI_j = \frac{V_j - \lambda_j}{V_j}, j = 1, 2, \dots, k. \quad (3)$$

Clearly, if $\lambda_j = V_j$, the $SFI_j = 0$. This represents that the number (or frequency) of service failures for the j th failure

category is at the maximum tolerable upper limit. When $\lambda_j > V_j$, $SFI_j < 0$. This represents that the number of service failures for the j th failure category has exceeded the tolerable upper limit. When $\lambda_j < V_j$, $SFI_j > 0$. This represents that the number of service failures for the j th failure category is tolerable. Hence, the larger the SFI_j value, the smaller the performance loss for the j th failure category is. Otherwise, the larger the service performance loss is.

Assume a restaurant states that its SFI_j value exceeds a required level (s_j). A hypothesis test (*i.e.*, $H_0: SFI_j \geq s_j$ versus $H_1: SFI_j < s_j$) will be used to judge whether the frequency of service failures for the j th failure category meets requirements. Generally speaking, s_j is set as 0, but s_j can also be set at a value larger than 0 to meet stricter requirements. In general, λ_j are unknown, and must be estimated. Let F_{ji} represents the number of the i th sample of service failures for j th failure category, then $(F_{j1}, F_{j2}, \dots, F_{jn})$ is a random sample with mean and variance of λ_j . If the sample mean of the number of service failures for the j th failure category (\bar{F}_j) is used to estimate λ_j , then the estimator of SFI_j can be intuitively defined as

$$\hat{SFI}_j = \frac{V_j - \bar{F}_j}{V_j}, \quad (4)$$

where $\bar{F}_j = \sum_{i=1}^n F_{ji}/n$, $j = 1, 2, \dots, k$. Because $E(\hat{SFI}_j) = SFI_j$, \hat{SFI}_j is an unbiased estimator of SFI_j , and it depends on complete and sufficient statistics \bar{F}_j only. Consequently, \hat{SFI}_j is a uniformly minimum variance unbiased estimator (UMVUE) of SFI_j . We have $Var(\hat{SFI}_j) = (1 - SFI_j)/n^2 V_j$.

By using \hat{SFI}_j as a test statistic of SFI_j , the rejection region can be expressed as $\{\hat{SFI}_j | \hat{SFI}_j < c_0\}$. Under the specified significance level α , the critical value (c_0) can be calculated as

$$c_0 = s_j - Z_\alpha \sqrt{(1 - s_j)/n^2 V_j}, \quad (5)$$

where Z_α is the upper α th quantile of the standard normal distribution.

If $\hat{SFI}_j < c_0$, the frequency of service failures of the j th failure category for the restaurant does not reach the stated performance level and is unqualified (or out of control limit). Improvement actions such as a six-sigma improvement project should be undertaken to improve service quality. If $\hat{SFI}_j > c_0$, the frequency of service failures of the j th failure category is within acceptable limit (or within control limit).

4. Data Analysis and Discussion

This study used a chain restaurant with three stores in Taiwan as an illustrative example. In the past few years, this restaurant has become well known for its delicious and economically priced meal sets, attracting the patronage of numerous students and younger consumers. However, with the growth of customer numbers and business, customer complaint incidents have increased substantially and restaurant reputation was damaged. The restaurant managers want to identify the unqualified failure categories to control the frequency of severe service failure and finally achieve customer satisfaction.

4.1 Analysis of Sample Structure

4.1.1 Expert Sample Structure

A total of 20 expert questionnaires were distributed, and 18 valid questionnaires were returned for a valid questionnaire return rate of 90%. Regarding gender, the proportion of male experts (10 experts, 55.56%) was higher than that of female experts (8 experts, 44.44%). Regarding age, those experts aged 41-50 years (11 experts, 61.11%)

accounted for the majority, followed by 51-60 years (4 experts, 22.22%), 31-40 years (2 experts, 11.11%) and those aged 60 years or above (1 expert, 5.56%). Regarding profession and ranking, ten experts (including 1 professor, 4 associate professors and 5 assistance professors) were university full-time teachers in restaurant field (55.56%) and eight experts (including 2 assistant managers and 6 restaurant managers or above) held positions in restaurant industry (44.44%). Regarding the years of service, those experts with 11-15 years (7 experts, 38.89%) and 16 years or above (7 experts, 38.89%) in restaurant industry (including school) accounted for the majority, four experts with 6-10 years (22.22%). They all had extensive and rich experience in restaurant industry. The results showed that the expert opinions were representative.

4.1.2 Consumer Sample Structure and Behavior Analysis

A total of 250 questionnaires were distributed, and 229 valid questionnaires were returned for a valid questionnaire return rate of 91.6%. The demographic information collected was as follows. Of the respondents, 50.88% were men and 49.12% were women. The leading consumer group comprised young adults under 30 years old (63.45%), successively followed by those between 31 and 40 years of age (18.71%), between 41 and 50 years of age (14.04%), and over 50 years of age (3.80%). The majority of consumers reported that they had college/university degrees (63.45%), successively followed by those with less than a high school education (30.99%), and people with master's degrees or greater (5.56%). Regarding profession, the majority of consumers were students (40.94%), successively followed by those working in the service industry (26.32%), manufacturing industry (18.42%), military and government department (6.43%), homemakers (5.85%), and others (2.04%). Regarding customers' behavior after encountering service failures, the average satisfaction level, repurchase intention, and willingness to recommend for the respondents who had experienced service failures were approximately 4 points or below using Likert 7-points scale. The results showed the service failures seem to have a negative effect on customers' subsequent consumption behaviors.

4.2 Analyzed Results of Fuzzy Delphi Method

A total of 18 effective expert questionnaires were returned (8 from restaurant managers and 10 from university teachers). No new service failure categories were generated. After screening with the fuzzy Delphi method, the initial restaurant-related service failure categories including *environmental problems*, *sold out or could not provide food or beverage items*, *booking mistakes*, *billing mistakes*, *request to change dish varieties*, *billed for the wrong dish*, and *deceiving customers* that did not attain the threshold value of 0.4 were deleted. As a result, 7 categories were deleted. Thus, we established a "formal restaurant-related service failure classification model" comprising three groups and 20 failure categories.

4.3 Descriptive Statistics for Restaurant Service Failure Frequency/Severity

Because all respondents neither reported experiencing two of the categories (*unexpected employee behavior*, or *failure to comply with special requests from customers*) nor filled in the empty space in the questionnaire, the above two categories were omitted and 18 service failure categories were finally re-numbered and analyzed, as shown in Table 3. Regarding the frequency, the top-three-ranking failure categories were sequentially presented as follows: *slow service* (20.09%), *fail to receive proper services* (13.97%), and *failure to provide services in the order of customer arrival* (11.35%). Restaurant managers should pay attention to these service failure categories, especially when 20.09% of the respondents believed that the restaurant had *slow service*, indicating an urgent need for improvement. The severities of all failure categories are larger than 0.5. The three most severe failure categories were: *poor staff hygiene* (0.921), *lack of professional knowledge and technique* (0.792), and *poor service attitude* (0.765). The severity of these service failures exceeded 0.7, indicating the importance of controlling these service failure categories.

4.4 Estimation and Testing of Service Failure Categories

Based on Equations (2)-(5) and the results in Table 3, the tolerable upper limits of failure frequency, estimation and testing of service failure indices were calculated and summarized in Table 4. For visualisation and convenience, the

\hat{SFI}_j and c_0 values can be used to construct a control limit of service failure indices shown as Figure 1. The

estimation and testing processes of failure indices for various failure categories were presented as follows:

Step 1: Assume that the failure index thresholds SFI_j of j th failure category (s_j) was set to 0.1 and the significance level (α) was set as 0.05.

Step 2: Based on V_j and \bar{F}_j values, \hat{SFI}_j were estimated. The results were shown in Table 4 and control chart (see Figure 1 for details).

Step 3: Critical values c_0 (Control limit in Figure 1) for the various failure categories were calculated. The results were shown in Table 4 and Figure 1.

Step 4: \hat{SFI}_j and c_0 were compared to test whether the service failure of the j th failure category was over the acceptable limit. If $\hat{SFI}_j < c_0$ (or \hat{SFI}_j value under control limit), the j th failure category was unqualified (or unacceptable).

Step 5: The results indicated that the restaurant had 7 unqualified (or unacceptable) failure categories as shown in Table 4 by the “*” mark. The other failure categories were qualified (or acceptable). The smaller the \hat{SFI}_j , the higher the priority for improvement of that failure category. Hence, the three failure categories most in need of improvement were *slow service* (i.e., $\hat{SFI}_3 = -2.63 < c_0 = -0.0024$), *fail to receive proper services* (i.e., $\hat{SFI}_2 = -1.33 < c_0 = -0.00231$), and *failure to provide services in the order of customer arrival* (i.e., $\hat{SFI}_8 = -1.03 < c_0 = -0.00239$) as circles in Figure 1.

The Pareto’s 80/20 rule indicated that the majority of quality problems were caused by failure in just a few categories. Using this assessment and testing process, restaurant managers can discover key areas of service quality problems and develop effective improvement plans and management strategies to enhance customer satisfaction and loyalty and gain a competitive advantage.

Table 3. Descriptive statistics of restaurant service failure frequency/severity

Failure groups	Failure categories	Frequency (\bar{F}_j)	Frequency Percentage	Rank	Severity (w_j)	Rank
Group 1: Service delivery system failures	1.Food and beverage quality problems	8	3.49%	10	0.643	6
	2.Fail to receive proper services	32	13.97%	2	0.578	15
	3.Slow service	46	20.09%	1	0.627	7
	4.Missing food or beverage items	12	5.24%	7	0.608	10
	5.Equipment failure	14	6.11%	6	0.61	9
	6.Tableware problems	7	3.06%	12	0.526	17
	7.Policy cognition differences	2	0.87%	16	0.6	11
	8.Failure to provide services in the order of customer arrival	26	11.35%	3	0.622	8
	9.Pre-service preparation and meal ordering failures	5	2.18%	14	0.591	13
	10.Service failure during meals	18	7.86%	4	0.684	4
	<i>Group 1 total</i>	<i>170</i>	<i>74.24%</i>	<i>1</i>	<i>0.609</i>	<i>2</i>
Group 2: Customer needs and requests	11.Seating arrangement or replacement mistakes	2	0.87%	16	0.6	11
	12.Cooking procedure errors	1	0.44%	18	0.5	18
	<i>Group 2 total</i>	<i>3</i>	<i>1.31%</i>	<i>3</i>	<i>0.550</i>	<i>3</i>
Group 3: Unprompted/u nsolicited employee action	13.Ordered the wrong dish	5	2.18%	14	0.58	14
	14.Delivered the wrong dish	12	5.24%	7	0.66	5
	15.Food and beverage spillage	8	3.49%	10	0.573	16
	16.Lack of professional knowledge and technique	6	2.62%	13	0.792	2
	17.Poor staff hygiene	10	4.37%	9	0.921	1
	18.Poor service attitude	15	6.55%	5	0.765	3
	<i>Group 3 total</i>	<i>56</i>	<i>24.45%</i>	<i>2</i>	<i>0.715</i>	<i>1</i>
	<i>Grand total</i>	<i>229</i>	<i>100.00%</i>		<i>0.638</i>	

The above results showed the failure frequency and severity, unqualified service failure categories, and the priority of service improvement based on primary service failure categories at the case restaurant. The results can help to establish service quality-improvement strategies and resource allocation plans. Finally, if test results indicate that certain failure categories do not meet requirements, quality improvement methods, such as the six-sigma method, can be used to analyze and identify the failure causes and adopt effective improvement strategies to meet performance requirements.

Table 4. Estimation and testing of service failure indices

Failure groups	Failure categories	w_j	V_j	\bar{F}_j	$S\hat{F}T_j$	c_0
Group 1: Service delivery system failures	1.Food and beverage quality problems	0.643	12.37	8	0.35	-0.00243
	2.Fail to receive proper services	0.578	13.76	32	-1.33*	-0.00231
	3.Slow service	0.627	12.68	46	-2.63*	-0.00240
	4.Missing food or beverage items	0.608	13.08	12	0.08	-0.00237
	5.Equipment failure	0.610	13.03	14	-0.07*	-0.00237
	6.Tableware problems	0.526	15.12	7	0.54	-0.00220
	7.Policy cognition differences	0.600	13.25	2	0.85	-0.00235
	8.Failure to provide services in the order of customer arrival	0.622	12.78	26	-1.03*	-0.00239
Group 2: Customer needs and requests	9.Pre-service preparation and meal ordering failures	0.591	13.45	5	0.63	-0.00233
	10.Service failure during meals	0.684	11.62	18	-0.55*	-0.00251
	11.Seating arrangement or replacement mistakes	0.600	13.25	2	0.85	-0.00235
	12.Cooking procedure errors	0.500	15.90	1	0.94	-0.00215
Group 3: Unprompted/un solicited employee action	13.Ordered the wrong dish	0.580	13.71	5	0.64	-0.00231
	14.Delivered the wrong dish	0.660	12.05	12	0.00	-0.00247
	15.Food and beverage spillage	0.573	13.88	8	0.42	-0.00230
	16.Lack of professional knowledge and technique	0.792	10.04	6	0.40	-0.00270
	17.Poor staff hygiene	0.921	8.63	10	-0.16*	-0.00291
	18.Poor service attitude	0.765	10.39	15	-0.44*	-0.00265
Total		0.638		229		

* denotes “unqualified”

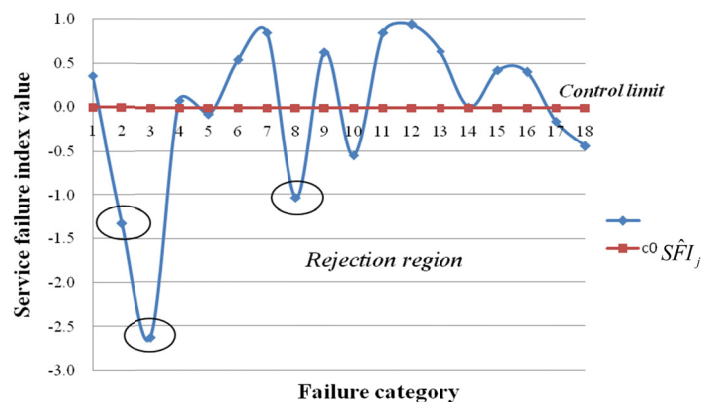


Figure 1. Control chart of service failure indices

5. Conclusion

Prior research has paid little attention to the construction of a service failure control process. This study proposed a restaurant-related service failure control method and used a chain restaurant as an illustrative example to develop tolerable upper limits for the frequency of service failure, to construct a service failure index, and to verify whether the frequency of service failure for the various service categories were acceptable using statistical testing methods (or control chart). The results indicate that this control process and method can effectively monitor occurrence frequency of service failure of various failure categories for restaurants, and assist restaurant managers in assessing and examining key service failures and in developing improvement plans and management strategies. In addition, this control process is conducive to reducing customer dissatisfaction and complaint, ensuring improvement of service quality, and achieving the purpose of service failure control and monitor.

Kim and Mauborgne (2005) in the *Blue Ocean Strategy* proposed a four-action framework including *eliminate*, *reduce*, *raise*, and *create*. According to this process, enterprises can eliminate unnecessary waste and processes,

reduce costs and customer complaints, raise customer satisfaction and service performances, and ultimately create differentiated service processes and characteristics using various methods and strategies. The restaurant-related service failure control process and method proposed in this study provides an effective control method for service failure and customer complaint management (reduction and enhancement within the four-action framework) in the restaurant industry. This approach surpassed the general restrictions of service failure and recovery research using the CIT method and provided an alternative perspective on service quality control and satisfaction assessment for both academia and the restaurant industry. It can also be referenced and revised, and even further applied in other industries.

The limitations of this study are that we did not use the CIT method to collect data from respondents. Rather, we compiled service failure classification research from restaurant-related literature to conduct the observations, inductions, and assessments. After review and correction by two expert scholars, we established an initial restaurant-related service failure classification model and used it to design the expert and consumer questionnaires. Therefore, collection of data did not take advantage of the strengths of the CIT method. However, this method required less time and had lower cost, effectively enhanced the respondents' willingness to participate in the survey, and more easily collected sample data that verified the proposed approach. The second limitation is that the tolerable upper limit of the number of service failures must be formulated based on sample size, which may not be very precise. This approach could be used in "define" or "measure" phase of six-sigma project in future studies.

References

- Bitner, M. J., Booms, B. H., & Tetreault, M. S. (1990). The service encounter: Diagnosing favorable and unfavorable incidents. *Journal of Marketing*, 54(1), 71-84. <http://dx.doi.org/10.2307/1252174>
- Chen, C. T. (2000). Extensions of the TOPSIS for group decision making under fuzzy environment. *Fuzzy Sets and Systems*, 114(1), 1-9. [http://dx.doi.org/10.1016/s0165-0114\(97\)00377-1](http://dx.doi.org/10.1016/s0165-0114(97)00377-1)
- Chen, H. T. (2013). Assessing the Performance of Service Failure Categories from Employee Perspective. *Journal of Global Business Management*, 9(3), 5-15. Retrieved from <http://connection.ebscohost.com/c/articles/94883619/assessing-performance-service-failure-categories-from-employee-perspective>
- Chen, S. H., & Hsieh, C. H. (2000). Representation Ranking, Distance, and Similarity of LR Type Fuzzy Number and Application. *Australian Journal of Intelligent Processing System*, 6(4), 217-229.
- Dalkey, N., & Helmer, O. (1963). An experimental application of the DELPHI method to the use of experts. *Management Science*, 9(3), 458-467. <http://dx.doi.org/10.1287/mnsc.9.3.458>
- Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes*. Glenview, IL: Scott, Foresman. <http://dx.doi.org/10.1177/105960117600100220>
- Dickens, J. (1987). The fresh cream cakes market: The use of qualitative research as part of a consumer research programme. In U. Bradley (Ed.), *Applied Marketing and Social Research* (pp. 23-68). New York, NY: John Wiley.
- Fisk, R. P., Brown, S. W., & Bitner, M. J. (1993). Tracking the evolution of the service marketing literature. *Journal of Retailing*, 69(1), 61-103. [http://dx.doi.org/10.1016/s0022-4359\(05\)80004-1](http://dx.doi.org/10.1016/s0022-4359(05)80004-1)
- Fornell, C. (1992). A National Customer Satisfaction Barometer: The Swedish Experience. *Journal of Marketing*, 56(1), 6-21. <http://dx.doi.org/10.2307/1252129>
- Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. London: Weidenfeld and Nicolson.
- Ha, J., & Jang, S. C. S. (2009). Perceived justice in service recovery and behavioral intentions: The role of relationship quality. *International Journal of Hospitality Management*, 28(3), 319-327. <http://dx.doi.org/10.1016/j.ijhm.2008.12.001>
- Hart, C. W. L., Heskett, J. L., & Sasser, W. E. Jr. (1990). The profitable art of service recovery. *Harvard Business Review*, 68(4), 148-156. Retrieved from <http://europemc.org/abstract/MED/10106796>
- Hoffman, K. D., Kelley, S. W., & Rotalsky, H. M. (1995). Tracking service failures and employee recovery efforts. *Journal of Service Marketing*, 9(2), 49-61. <http://dx.doi.org/10.1108/08876049510086017>
- Hong, J. S., Hsu, H., Liu, C. H., Lin, L., & Tsai, C. Y. (2011). Competency analysis of top managers in the Taiwanese hotel industry. *International Journal of Hospitality Management*, 30(4), 1044-1054. <http://dx.doi.org/10.1016/j.ijhm.2011.03.012>
- Hsia, T. C., Chen, H. T., & Chen, W. H. (2008). Measuring the Readability Performance (RP) of Aircraft

- Maintenance Technical Orders by Fuzzy MCDM Approach and RP Index. *Quality and Quantity*, 42(6), 795-807. <http://dx.doi.org/10.1007/s11135-006-9068-9>
- Hsu, T. H. (1998). The fuzzy Delphi analytic hierarchy process. *Journal of Chinese Fuzzy Systems Association*, 4(1), 59-72.
- Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tarsuta, R., & Mieno, H. (1993). The max–min Delphi method and fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems*, 55(3), 241–253. [http://dx.doi.org/10.1016/0165-0114\(93\)90251-c](http://dx.doi.org/10.1016/0165-0114(93)90251-c)
- Johnston, T. C., & Hewa, M. A. (1997). Fixing service failures. *Industrial Marketing Management*, 26(5), 467-473. [http://dx.doi.org/10.1016/s0019-8501\(96\)00158-7](http://dx.doi.org/10.1016/s0019-8501(96)00158-7)
- Kelley, S. W., & Davis, M. A. (1994). Antecedents to Customer Expectations for Service Recovery. *Journal of Academy of Marketing Service*, 22(1), 52-61. <http://dx.doi.org/10.1177/0092070394221005>
- Kim, W. C., & Mauborgne, R. (2005). *Blue ocean strategy: How to create uncontested market space and make the competition irrelevant*. Boston, MA: Harvard Business School Publishing.
- Lin, Y. H., Huang, D., & Huang, Y. L. (2003). An Analysis of the Typology of Service Failures and Recoveries: A Study of Sit-Down Restaurants in Taiwan. *Journal of Tourism Studies*, 9(1), 39-59. <http://www.airitilibrary.com/Publication/alDetailedMesh?docid=10255273-200306-9-1-39-59-a>
- Linstone, H. A. (1985). The Delphi Technique. In V. T. Covello, J. L. Mumpower, P. J. M. Stallen, & V. R. R. Uppuluri (Eds.), *Environmental Impact Assessment, Technology Assessment, and Risk Analysis* (pp. 621-649). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-642-70634-9_22
- Maxham III, J. G., & Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing*, 66(4), 57-71. <http://dx.doi.org/10.1509/jmkg.66.4.57.18512>
- McCollough, M. A., Berry, L. L., & Yadav, M. S. (2000). An Empirical Investigation of Customer Satisfaction after Service Failure and Recovery. *Journal of Service Research*, 3(2), 121-137. <http://dx.doi.org/10.1177/109467050032002>
- Mueller, R. D., Palmer, A., Mack, R., & McMullan, R. (2003). Service in the restaurant industry: an American and Irish comparison of service failures and recovery strategies. *International Journal of Hospitality Management*, 22(4), 395–418. [http://dx.doi.org/10.1016/s0278-4319\(03\)00072-0](http://dx.doi.org/10.1016/s0278-4319(03)00072-0)
- Murray, T. J., Pipino, L. L., & Gigch, J. P. (1985). A pilot study of fuzzy set modification of Delphi. *Human Systems Management*, 5(1), 76-80. <http://dx.doi.org/10.3233/HSM-1985-5111>
- Priluck, R. (2003). Relationship Marketing Can Mitigate Product and Service Failures. *Journal of Services Marketing*, 17(1), 37-48. <http://dx.doi.org/10.1108/08876040310461264>
- Ross, S. M. (2010). *Introduction to Probability Models* (10th ed.). San Diego, California: Academic Press. <http://dx.doi.org/10.1016/B978-0-12-375686-2.00007-8>
- Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research*, 36(3), 356-372. <http://dx.doi.org/10.2307/3152082>
- Spreng, R. A., Harrell, G. D., & Mackoy, R. D. (1995). Service recovery: Impact on satisfaction and intentions. *Journal of Service Marketing*, 9(1), 15-23. <http://dx.doi.org/10.1108/08876049510079853>
- Tang, M. T., Tzeng, G. H., & Wang, S. W. (1999). A hierarchy fuzzy MCDM method for studying electronic marketing strategies in the information service industry. *Journal of International Information Management*, 8(1), 1–22.
- Tsai, C. H., & Chen, C. W. (2011). The establishment of a rapid natural disaster risk assessment model for the tourism industry. *Tourism Management*, 32(1), 158-171. <http://dx.doi.org/10.1016/j.tourman.2010.05.015>
- Tsaur, S. H., Lin, Y. C., & Lin, J. H. (2006). Evaluating ecotourism sustainability from the integrated perspective of resource, community and tourism. *Tourism Management*, 27(4), 640–653. <http://dx.doi.org/10.1016/j.tourman.2005.02.006>
- Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8(3), 338-353. [http://dx.doi.org/10.1016/S0019-9958\(65\)90241-X](http://dx.doi.org/10.1016/S0019-9958(65)90241-X)
- Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning-1. *Information Sciences*, 8(3), 199-249. [http://dx.doi.org/10.1016/0020-0255\(75\)90036-5](http://dx.doi.org/10.1016/0020-0255(75)90036-5)