

THE SERVICE RECOVERY PARADOX: TRUE BUT OVERRATED?

Stefan Michel

Global Business Department

Thunderbird, The Garvin School of International Management

15249 N. 59th Avenue

Glendale, AZ 85306

(602) 978 72 96

FAX (602) 843 61 43

MichelS@t-bird.edu

Matthew L. Meuter

Department of Finance and Marketing

California State University, Chico

Chico, CA 95929-0051

(530) 898-5880

FAX (530) 898-6360

mmeuter@csuchico.edu

KEYWORDS: Service recovery, Recovery paradox, Banking, Satisfaction, Loyalty, Word of mouth

THE SERVICE RECOVERY PARADOX: TRUE BUT OVERRATED?

Empirical research paper

Purpose

To test the existence, frequency and magnitude of the service recovery paradox.

Design/methodology/approach

To date, much of the literature exploring the service recovery paradox has generated mixed results. We argue that a service recovery paradox is a rare event, which makes its measurement difficult, since the “treatment group” sample size is usually too small to produce significant results. For that reason, we test the existence, frequency and magnitude of the service recovery paradox in a banking context with more than 11,000 customer interviews based on actual customer encounters.

Findings

Overall, the survey findings support the argument that a service recovery paradox is a rare event, and the hypothesized mean differences are, albeit significant, not very large, which diminishes their managerial relevance to some degree.

Research limitations/implications

Because of the required extremely large sample size, no multi-item measures were collected. Furthermore, privacy concerns restricted us from a longitudinal study and from linking the survey results to behavioural data. Both limitations are inherent in the chosen setting.

Practical implications

While a service failure offers an opportunity to create an excellent recovery, the likelihood of a service paradox is very low. The implications of verifying a service recovery paradox do not suggest that ineffective service followed by an outstanding service recovery is a viable strategy.

Originality/value

To our knowledge, this is the first empirical study testing not only the existence and magnitude, but also the frequency of a service recovery paradox. This is crucial because the paradox is a very rare event, which, in turn, limits its managerial relevance.

THE SERVICE RECOVERY PARADOX: TRUE BUT OVERRATED?

ABSTRACT

The service recovery paradox refers to situations in which the overall satisfaction levels of recovered customers exceed those of customers who did not encounter any problems with the initial service. Our literature review shows that the paradox is found only under very special circumstances, indicating that it is a rare or even extreme incident. In this study using surveys of more than 11,000 customers of a retail bank, a very satisfying initial service is what is most preferred. Nevertheless, a service recovery paradox appears when customers compare a recovery that is “much better than expected” with an error-free initial service that is just “satisfying.” The paradox is present for both “overall satisfaction” and “word-of-mouth intentions.” Overall, the survey findings support the argument that a service recovery paradox is a rare event, and the hypothesized mean differences are, albeit significant, not very large, which diminishes their managerial relevance to some degree. This must be taken into account when managers design service recovery strategies or calculate the impact of service recovery improvements on customer equity.

INTRODUCTION

Increasing competitive pressures in many service industries, coupled with declining perceptions of customer service, have led to increased attention to service recovery in recent years (Andreassen, 2001, Kelley *et al.*, 1993, Maxham, 2001, Maxham and Netemeyer, 2002, Maxham and Netemeyer, 2002, Maxham and Netemeyer, 2003, McCollough *et al.*, 2000, McCollough and Bharadwaj, 1992, Smith *et al.*, 1999, Swanson and Kelley, 2001, Tax *et al.*, 1998). Service failures can lead to negative disconfirmation and ultimately dissatisfaction, though appropriate service recovery efforts may restore a dissatisfied customer to a state of satisfaction (Bitner *et al.*, 1990). Although some researchers have argued that the best strategy is to fail-safe the original service delivery, it is nearly impossible to eliminate all failures. Thus, firms with the ability to react to service failures effectively and implement some form of service recovery will be in a much better position to retain profitable customers.

A service failure is defined as “any service-related mishaps or problems (real and/or perceived) that occur during a consumer’s experience with the firm” (Maxham, 2001). In line with this wide definition of a service failure, service recovery can be defined as the service provider’s action when something goes wrong (Grönroos, 1988). More recently, Smith, Bolton and Wagner (1999, p. 357) have treated service recovery as “a ‘bundle of resources’ that an organization can employ in response to a failure.” Both complaint management and service recovery are based on service encounter failures. However, complaint management is based on the firm’s reaction to a customer complaint, whereas service recovery also addresses the firm’s ability to react immediately to a failed service encounter, pleasing the customer before he or she finds it necessary to complain. Because many customers dissatisfied with a service encounter are reluctant to complain (Andreassen and Best, 1977, Singh, 1990), proactive service recovery efforts—that is, those that attempt to solve problems at the point of the encounter—are the most effective way to minimize negative outcomes of a service failure (Lewis, 1996). Finally, the term “recovery paradox” refers to situations in which the

satisfaction, word-of-mouth intentions, and repurchase rates of recovered customers exceed those of customers who have not encountered any problems with the initial service (McCollough and Bharadwaj, 1992).

The purpose of this study is twofold: we first want to determine if, and under what conditions, the service recovery paradox occurs. Secondly, if it exists, we want to assess its frequency in a real-world setting. We rigorously test this phenomenon in a banking context with more than 11,000 customer interviews based on actual customer encounters. The remainder of the article is organized as follows: first, we discuss prior literature relating to service recovery and the recovery paradox. Second, we establish hypotheses to be tested, discuss our methodology, and outline our data collection efforts. Third, we present the results of the study, and, finally, we discuss the managerial implications, further research directions, and limitations of our research.

CONCEPTUAL FOUNDATIONS

Based on personal experiences and anecdotal evidence, the idea of a recovery paradox was developed more than 20 years ago by Etzel and Silverman (1981, p. 128), who stated that “it may be those who experience the gracious and efficient handling of a complaint who become a company’s best customer.” Other research has suggested that “a good recovery can turn angry, frustrated customers into loyal ones. It can, in fact, create more goodwill than if things had gone smoothly in the first place” (Hart *et al.*, 1990, p. 148).

Since then, a wide range of empirical studies have explored the service recovery paradox, as summarized in Table 1.

INSERT TABLE 1 ABOUT HERE

As Table 1 shows, eight studies refute the existence of a service recovery paradox (Andreassen, 2001, Berry *et al.*, 1990, Bolton, 1998, Brown *et al.*, 1996, Halstead and Page, 1992, Maxham, 2001, McCollough *et al.*, 2000, Zeithaml *et al.*, 1996). The consensus of

these studies indicates that there is no way to please customers more than with a reliable, first-time, error-free service. According to these findings, service recovery is a strategy to limit the harm caused by a service failure rather than a means to impress the customer with a special effort when something goes wrong. However, six other studies, some of them by the same authors mentioned above, indicate that the service recovery paradox is a real phenomenon (Bolton and Drew, 1992, Boshoff, 1997, Hocutt *et al.*, 2006, Hocutt *et al.*, 1997, Maxham and Netemeyer, 2002, McCollough, 2000, Smith and Bolton, 1998). Based on this literature review, it is fair to conclude that the evidence of the existence of a service recovery paradox is mixed, at best. The chosen methodology, e.g., using surveys versus experiments, or using cross-sectional versus longitudinal data, and the industry context seem not to have an impact of whether a service recovery paradox is found. The question then becomes: is there a simple, yet powerful explanation for this contradictory result? We see two possible reasons for this divergence. First, the service recovery paradox is not defined uniformly in the studies mentioned above: some authors test for a between-subject effect (comparing a recovery/complaining group with a control group) while others test for a within-subject effect (before a failure/complaint and after a failure/complaint). Moreover, the measurement of the recovery incidents is not defined uniformly nor is the dependent variable the same. In short, we have not found one single replication study on this subject but rather more than a dozen different approaches.

A second possible reason for the mixed finding is not related to a methodological problem but is, instead, based on the very nature of the paradox. It has been suggested that a service recovery paradox is a very rare event (Boshoff, 1997, Hart *et al.*, 1990), which means that it is not easy to detect even if it exists. To make things worse, it is further assumed that only a minority of dissatisfied customers complains (Andreasen and Best, 1977, Singh, 1990) and that most recoveries do not lead to customer satisfaction (Hoffman *et al.*, 1995). If this explanation is valid, one would need a very large sample to obtain a reasonably large group

of customers who received a very satisfactory recovery. This requirement may explain why some studies have failed to produce significant results.

Our study addresses these problems by using a narrowly defined comparison set, a survey-based approach with actual service encounters, and very large groups of customers as a treatment group (experiencing a failure) and as a control group (experiencing an error-free service). It is especially noteworthy that the treatment group consists of customers who experienced a service failure, regardless of whether they complained or not.

HYPOTHESES DEVELOPMENT

Satisfaction

Satisfaction is commonly defined using the expectancy disconfirmation paradigm (Oliver, 1980, Oliver, 1996), which claims that when consumers receive service that is better than expected, they will be satisfied (Oliver, 1980). Alternatively, service that is worse than expected leads to dissatisfaction. According to the disconfirmation theory, satisfaction results when the consumer has an encounter that is better than expected. In service recovery research, two evaluation phases occur. Service recovery starts, by definition, with initial customer dissatisfaction. After this first evaluation, when they determine the service was worse than they expected, customers may go through a recovery process with the firm, which leads to a second evaluation. In a study of 410 complaining customers of an interstate moving company, satisfaction with recovery had a greater impact on repurchase and word-of-mouth intentions than did satisfaction with the initial service (Spreng *et al.*, 1995). However, according to Andreassen (2000), disconfirmation, rather than expectations, has a dominant impact on satisfaction.

As the literature review in Table 1 suggests, a service recovery paradox is likely to occur when excellent recoveries are compared with mediocre, error-free service transactions

(Bolton and Drew, 1992; Boshoff, 1997; Hocutt et al., 1997, 2006; Smith and Bolton, 1998; McCollough, 2000), but not when excellent recoveries are compared with excellent, error-free service transactions (Berry *et al.*, 1990; Zeithaml *et al.*, 1996). Therefore, we hypothesize that:

H1a: Customers who report an error-free, *very* satisfying initial transaction have higher overall satisfaction rates than do customers who experience an initial service failure followed by a recovery that is much better than expected (no recovery paradox expected).

H1b: Customers who report an error-free, *just* satisfying initial transaction have lower overall satisfaction rates than do customers who experience an initial service failure followed by a recovery that is much better than expected (recovery paradox expected).

Recommendation Intention

Effective service recovery increases not only overall satisfaction, but also positive word of mouth (Spreng *et al.*, 1995, Swanson and Kelley, 2001). For example, one study shows strong links between satisfaction with complaint handling and word-of-mouth behaviour (Oliver and Swan, 1989). That is, successfully recovered customers recommend the company to others or “demonstrate a strong propensity to share positive information about the experience” (Swanson and Kelley, 2001). In Maxham’s (2001) study, a *very good* service recovery, compared with a *good* recovery, had a stronger impact on word of mouth than on satisfaction or repurchase intention. This impact is critical because many services possess credence qualities (Zeithaml and Bitner, 1996), for which word-of-mouth communication can

have an extremely powerful influence in terms of the consumer purchasing process (Richins, 1983). In line with the proposed effect on overall satisfaction, we hypothesize that:

H2a: Customers who report an error-free, *very* satisfying initial transaction have higher recommendation intentions than do customers who experience an initial service failure followed by a recovery that is much better than expected (no recovery paradox expected).

H2b: Customers who experience an error-free, *just* satisfying initial transaction have lower recommendation intentions than do customers who experience an initial service failure followed by a recovery that is much better than expected (recovery paradox expected).

METHODOLOGY

A central contribution of this study is the methodological approach we use to explore the *existence, frequency, and magnitude* of the service recovery paradox. In this research project, we compare satisfaction levels of actual customers in a real service encounter who experience no initial service failure, with those of customers who experience an initial service failure followed by a service recovery effort. We questioned those who experienced an initial failure further to explore their feelings and perceptions about the service recovery experience. This comparison enables us to determine more clearly the existence, frequency, and magnitude of the service recovery paradox.

One difficulty relating to this type of comparison is the cost. To find a sufficiently large number of failure encounters and service recovery efforts, an extraordinarily large number of respondents must be contacted. As a consequence, this design is efficient only as part of a larger, general service quality and customer satisfaction study. We therefore conducted

our study within the framework of a large-scale service quality study of a major Swiss bank that offers a full range of banking services to more than one million customers.

Survey

The respondents were sampled based on the fact that they had a recent interaction with the bank, and were asked about their encounter satisfaction. In addition, we assessed overall satisfaction with the bank and its service representatives, and the likelihood of recommending the bank.

If a service failure did not occur in the recent interaction, the respondent became part of the control group of consumers who experienced no service failure. However, if a service failure was experienced, whether or not the customer complained to the firm, the interviewer switched to a service recovery questionnaire. Because the context of the failure was known, we were able to assess the customer's disconfirmation with the service recovery. Single-item, five-point Likert scale measures were used because of survey length considerations and because this research was a part of a larger service quality study.

To assess the service encounter, customers were asked to report any deviation from their expectations, not just "critical incidents" (Bitner *et al.*, 1990) or significant failures. Previous research has shown that customers expect to see common mistakes and excuse them easily (Chung and Hoffman, 1998). To eliminate those situations in which a customer identifies a common, simplistic "failure" from our analysis, customers rated failures according to their magnitude, as has been done in previous research (Michel, 2004, Smith *et al.*, 1999, Webster and Sundaram, 1998). Only those failures rated as "not acceptable" or "absolutely not acceptable" were included.

To collect the data, we employed a telephone survey using trained interviewers who worked exclusively on this project. A computer-aided telephone interview (CATI) terminal was used, which prompts a specific question on the basis of each response. Interviewers

typed responses to each question when the interview was completed. One of the principle researchers attended the initial interviewer briefing and observed actual data collection twice during the process to ensure data were collected appropriately.

RESULTS

Sample

From a sample of 11,929 customers, 9,166 experienced no failure (76.8%), 2,638 experienced one failure (22.1%), and 125 experienced two failures (1.0%). Those who reported different types of failures during the initial service were dropped from further analysis. Of the 2,638 who experienced a single failure, 1,025 were not able to evaluate their recovery experience because it was still pending. This high ratio of pending service recovery efforts (39%) indicates the recency of the incidents. Another 424 customers did not evaluate the incident or did not finish the survey, resulting in final group of 1,189 valid cases for the recovery group. In the “no failure” group, 8,714 reported their satisfaction with the error-free initial service (452 did not complete the survey).

In total, more men (61.5%) participated in the study than did women (38.5%). A small group of respondents were under 30 years of age (12.4%); all others were well distributed across the other four age categories (24.8% aged 30–39; 21.2% aged 40–49; 20.6% aged 50–59; 21.1% aged over 60). The bank’s management perceives these distributions as representative of the bank’s customer base.

Test of Hypotheses

To test for the recovery paradox, both the failure and the no-failure groups were split into five subgroups. The failure groups range from group RECOVERY -- (service recovery was much worse than expected) to RECOVERY ++ (service recovery was much better than expected). The control groups are NOFAILURE -- (very dissatisfied with initial service) to

NOFAILURE ++ (very satisfied with initial service). List-wise deletion was used for all cases that were missing one of the dependent variables, resulting in a sample size of 9,474 customers for some of the analysis. In Table 2, we show the subsample group sizes, means, and standard deviations for both overall satisfaction and recommendation intention. Bar charts of means are shown in Figures 1 and 2.

INSERT TABLE 2 ABOUT HERE

Before proceeding with the testing of the hypotheses, it was important that we ensured that the anticipated patterns were observed. As we expected, overall satisfaction levels among those who encountered a failure were highest for the RECOVERY ++ group and lowest for the RECOVERY -- group. Overall satisfaction levels followed the same pattern for the no-failure groups (overall satisfaction highest for the NOFAILURE ++ group and lowest for the NOFAILURE -- group). The same patterns also hold for the recommendation intention data.

We controlled for gender, age, and relationship strength (a dichotomous item indicating whether the sponsoring firm is the respondent's main bank that measures relationship strength as either strong or weak). For each of the three control variables, two tests were conducted. First, we assessed recovery dissatisfaction between the groups using chi-square statistics. These results show that no such difference appears across the three control variables. Second, we investigated the mean values for each subgroup to determine if the same patterns occurred as those found for the overall sample. As we expected, age, gender, and relationship strength did not affect the mean values of any of the subgroups across overall satisfaction or recommendation intentions (i.e., overall satisfaction among those experiencing a failure was still greatest for RECOVERY ++ and lowest for RECOVERY --).

To test for the service recovery paradox, we compared those respondents who experienced a service failure (RECOVERY) with those who had an error-free first encounter (NOFAILURE). Specifically, we analyzed the mean differences of three groups. The

RECOVERY ++ group perceived the service recovery as “much better than expected,” whereas the NOFAILURE ++ group was “very satisfied” and the NOFAILURE + group was “satisfied” with the initial transaction. The Levene test showed that homogeneity of variances cannot be assumed ($F=156.933$, $p<0.001$ and for recommendation intention $F=23.471$, $p<0.001$). Since sample sizes are very unequal, ANOVA and t-test are not appropriate. Instead, the non-parametric Mann-Whitney U test was applied for significance testing (Glantz and Slinker, 2001, p. 326-327). Mean differences in overall satisfaction with the bank are significant between the RECOVERY ++ group and the NOFAILURE ++ group ($z = -3.498$, $p < 0.001$). Customers who did not experience a failure and were very satisfied with the transaction (NOFAILURE ++) show significantly higher overall satisfaction (4.56) than any other group. These results indicate that, as we hypothesized, very satisfying service is preferred to even the best recovery efforts. Therefore, with respect to very satisfying service in an initial encounter, we do not find a recovery paradox, in support of H1a. However, if customers were just satisfied with their initial transaction (NOFAILURE +), their overall satisfaction (4.13) was significantly lower than that of the RECOVERY ++ group (4.29) ($z = -2.838$, $p < 0.01$). These results indicate that a very satisfying service recovery effort is more satisfying overall than a just satisfactory initial service interaction, and, thus, suggest a service recovery paradox, in support of H1b. We illustrate these results in Figure 1.

INSERT FIGURE 1 ABOUT HERE

We found a slightly different pattern for recommendation intentions. Mean differences in the recommendation intentions between the NOFAILURE ++ (4.54) and RECOVERY ++ (4.46) groups are not significant ($z = -1.143$, $p > 0.1$), which indicates that a very good service recovery can lead to approximately the same recommendation intention as does a very satisfactory initial transaction. Therefore, H2a is not supported. Customers who

assessed the initial transaction as “just satisfying” (NOFAILURE +) report a significantly weaker recommendation intention (4.28) than do customers whose service recovery was much better than expected (RECOVERY ++). The mean differences are significant ($z = -1.709, p < 0.05$), in support of H2b. These results are illustrated in Figure 2.

INSERT FIGURE 2 ABOUT HERE

Since our results support the existence of a service recovery paradox, we are especially interested in its frequency and magnitude. Table 2 indicates that the service recovery paradox is a very rare event, since only 63 out of 1,189 respondents reported a service recovery episode that was “much better than expected.” Furthermore, the mean differences between the two subgroups RECOVERY ++ and NOFAILURE + are 0.16 (4.29–4.16, see Figure 1) for “overall satisfaction” and 0.18 (4.46–4.28, see Figure 2) for “recommendation intention.” Although the mean differences are significant, they are not very large.

DISCUSSION

We agree that “service recovery research is particularly challenging because the activities associated with recovery are triggered by a service failure, making systematic empirical studies ... very difficult to conduct” (Smith and Bolton, 1998, p. 70). However, by integrating service recovery research into a general service quality study, access to high-quality data for satisfied customers can provide a control group that gives unique insight into the service recovery paradox.

The findings of this study show that error-free, “very satisfying” initial transactions are the best way to drive customer satisfaction. These situations are more satisfying than even a service recovery effort that is “much better than expected.” However, we find evidence of a service recovery paradox when the initial service is just “satisfying.” In these situations, cus-

tomers who experience a recovery episode that is “much better than expected” are actually more satisfied than those who experience just a “satisfying” initial encounter with no failure. This finding supports the presence of a service recovery paradox in relation to satisfaction.

Although empirical support confirms that satisfaction levels are significantly higher after error-free, very satisfying initial encounters than after a strong service recovery, the same cannot be said for the likelihood of spreading positive word of mouth. There is no statistically significant difference in word-of-mouth intentions after initial error-free, “very satisfying” encounters and failures followed by a “much better than expected” service recovery. This evidence suggests that a recovery paradox is more likely to occur for word-of-mouth intentions than for overall satisfaction, as is supported by a previous study that finds that “change in repatronage intentions is consistently more positive than the change in cumulative satisfaction for all levels of satisfaction with the recovery effort” (Smith and Bolton, 1998, p. 75).

With no statistical difference in word-of-mouth intentions, it appears that either a strong initial encounter or a strong service recovery will generate the same level of positive word of mouth. However, similar to the satisfaction findings, we find evidence of a service recovery paradox related to positive word of mouth when the initial, error-free encounter is just “satisfying.” Word-of-mouth intentions are significantly higher when a customer experiences a failure followed by a service recovery that is “much better than expected.” Thus, the findings related to both satisfaction and word of mouth provide confirming evidence of a service recovery paradox.

However, our study also indicates that a service recovery paradox is very rare. Of the 1,189 customers who experienced a failure, only 63 (5.4%) evaluated the service recovery effort as “much better than expected” and the mean differences in the dependent variables “overall satisfaction” and “recommendation intentions” are significant but not very large (0.16 and 0.18, respectively).

MANAGERIAL IMPLICATIONS

Important managerial implications stem from these research findings. First and foremost, managers and employees must strive for a reliable, error-free service that is very satisfying. Both satisfaction levels and positive word of mouth decrease precipitously when just satisfying service is provided. Not only will customers value and appreciate the service, but it also is cost effective, in the sense that the service firm will have fewer encounters that need to be recovered. If a company creates a failure situation and then does not recover effectively, customers will be especially negative because of the “double deviation” concept of two failures in a row (Bitner *et al.*, 1990). By surveying all customers, not only complainers, the company can detect common failure points in their service delivery. More specifically, learning from all customers provides information about the frequency and magnitude of failures (Tax and Brown, 1998). Armed with this information, managers can prioritize their improvement efforts.

Furthermore, because service failures are inevitable in most settings, service recovery is important and can create a service recovery paradox. Not only must a firm have a response plan in place when consumers complain, but it also must foster an environment in which employees are empowered and trained to rectify service recovery problems quickly, even before customers generate complaints. Organizations, therefore, must embed a service recovery management system into their cultural context that champions the perspective that “errors are inevitable—but dissatisfied customers are not” (Hart *et al.*, 1990, p. 148).

The implications of verifying a service recovery paradox do not suggest that ineffective service followed by an outstanding service recovery is a viable strategy. First, recovery efforts that are perceived as very satisfying are expensive and difficult to manage. Second, it may be the uniqueness of a recovery that creates the “wow” effect. Customer delight is achieved “from having one’s expectations exceeded to a surprising degree” (Rust and Oliver,

2000, p. 86). Consequently, standardized recoveries can never create this uniqueness and surprise. Third, it is very difficult to create a culture of lax initial service delivery in which failures are accepted and yet develop a culture of fantastic service recovery efforts. Thus, a service recovery strategic focus should be both active, in the sense that employees aggressively implement service recovery efforts immediately, and passive, in the sense that service recovery opportunities should not be planned.

We have found, in executive seminars and focus group discussions with managers and employees, that effective salespeople use service recovery strategies even before customers are aware of a problem. For example, one sales manager of a construction supply retailer was an hour late with a delivery. When the sales manager realized that he would be late, he called the customer to explain that the warehouse had made a mistake and that he would return to the warehouse to ensure that the order was complete. Instead of being dissatisfied with the late delivery, the customer appreciated that the sales manager went the extra mile. In cases in which part of the service is produced behind “the line of visibility” (Zeithaml and Bitner, 1996, p. 280), service recovery management should include perception management as well.

Undeniably, there has been a trend calling for accountability of marketing actions (Berger *et al.*, 2002, Rust *et al.*, 2004) and service recovery efforts (Michel *et al.*, 2004, Zhu *et al.*, 2004). Assuming that firms strive to maximize customer equity, each service recovery program should be regarded as an investment that requires a positive return on investment. If we compare the frequency of a paradox (63 customers are affected) with the frequency of a service failure (2,753 customers out of 11,929), we conclude that investments in a better, more reliable error-free service may yield much higher returns than investments in a service recovery program. In an ideal world, companies achieve both objectives simultaneously. However, faced with the decision of whether to fire-fight a single service problem to retain a customer, or to spend the time and money for process improvement, the latter strategy is much more likely to have a higher return on investment.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Although this study provides a test of the service recovery paradox using actual satisfied and dissatisfied customers in real-life situations, some limitations remain. First, we consider only the retail banking industry. Although this industry is well suited for this research because of its long-lasting, formal relationships and many service transactions, the findings cannot be generalized to other industries until further studies are conducted.

Second, we use many single-item scales. Because sample size is crucial to detect rare service paradox incidents, our study was possible only because it was part of a larger commercial service quality study, for which the length of the survey questionnaire was a limitation. However, single-item measures have been used successfully in comparable studies (Bolton and Drew, 1992, Mittal *et al.*, 1999, Mittal *et al.*, 1998). Multi-item measures are preferred, but not always required, in service research (Drolet and Morrison's, 2000, Rossiter, 2002) if cost consideration and respondent fatigue are limiting factors.

In addition to addressing these limitations in future research projects, it would be insightful to pursue a replication study in service settings, as well as across cultures. Exploring whether customers of different nationalities have differing perspectives on failure and recovery would be a valuable addition. Linking our findings to additional outcomes, such as loyalty or switching behaviour, might also be interesting.

REFERENCES

- Andreasen A.R. and Best A. (1977) "Customers Complain-Does Business Respond?" *Harvard Business Review*, Vol 55 No July-August, pp. 93-101
- Andreassen T.W. (2000) "Antecedents of Satisfaction with Service Recovery", *European Journal of Marketing*, Vol 34 No 1/2, pp. 156-175
- Andreassen T.W. (2001) "From Disgust to Delight: Do Customers Hold a Grudge?" *Journal of Service Research*, Vol 4 No 1, pp. 39-49
- Berger P.D., Bolton R.N., Bowman D., et al. (2002) "Marketing Actions and the Value of Customer Assets: A Framework for Customer Asset Management", *Journal of Service Research*, Vol 5 No 1, pp. 39-54
- Berry L.L., Zeithaml V.A. and Parasuraman A. (1990) "Five Imperatives for Improving Service Quality", *Sloan Management Review*, Vol 31 No Summer, pp. 29-38
- Bitner M.J., Booms B.H. and Tetreault M.S. (1990) "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents", *Journal of Marketing*, Vol 54 No January, pp. 71-84
- Bolton R.N. (1998) "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction", *Marketing Science*, Vol 17 No 1, pp. 45-65
- Bolton R.N. and Drew J.H. (1992) "Mitigating the Effect of Service Encounters", *Marketing Letters*, Vol 3 No 1, pp. 57-70
- Boshoff C. (1997) "An Experimental Study of Service Recovery Options", *International Journal of Service Industry Management*, Vol 8 No 2, pp. 110-130
- Brown S.W., Cowles D.L. and Tuten T.L. (1996) "Service Recovery: Its Value and Limitations as a Retail Strategy", *International Journal of Service Industry Management*, Vol 7 No 5, pp. 32-46
- Chung B. and Hoffman K.D. (1998) "Critical Incidents. Service Failures That Matter Most", *Cornell Hotel and Restaurant Administration Quarterly*, Vol 39 No June, pp. 66-71
- Drolet A.L. and Morrison's D.G. (2000) "Do We Really Need Multiple-Item Measures in Service Research?" *Journal of Service Research*, Vol 3 No February, pp. 196-204
- Etzel M.J. and Silverman B.I. (1981) "A Managerial Perspective on Directions for Retail Customer Dissatisfaction Research", *Journal of Retailing*, Vol 57 No Fall, pp. 124-136
- Glantz S.A. and Slinker B.K. (2001), *A Primer of Applied Regression and Analysis of Variance*, 2. Edition, New York, McGraw Hill,
- Grönroos C. (1988) "Service Quality: The Six Criteria of Good Perceived Service Quality", *Review of Business*, Vol 9 No Winter, pp. 10-13
- Halstead D. and Page T.J.J. (1992) "The Effects of Satisfaction and Complaining Behavior on Consumer Repurchase Intentions", *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Vol 5 No 1, pp. 1-11
- Hart C.W.L., Heskett J.L. and Sasser W.E.J. (1990) "The Profitable Art of Service Recovery", *Harvard Business Review*, Vol 68 No July-August, pp. 148-156
- Hocutt M.A., Bowers M.R. and Donovan D.T. (2006) "The Art of Service Recovery: Fact or Fiction?" *Journal of Service Marketing*, Vol 20 No 3, pp. 199-207
- Hocutt M.A., Chakraborty G. and Mowen J.C. (1997), "The Impact of Perceived Justice on Customers Satisfaction and Intention to Complain in a Service Recovery", Brucks M. (Ed.), *Advances in Consumer Research*, Ann Arbor, Association for Consumer Research, pp. 457-463

- Hoffman K.D., Kelley S.W. and Rotalsky H.M. (1995) "Tracking service failures and employee recovery efforts", *Journal of Services Marketing*, Vol 9 No 2, pp. 49-61
- Kelley S.W., Hoffman K.D. and Davis M.A. (1993) "A Typology of Retail Failures and Recoveries", *Journal of Retailing*, Vol 69 No Winter, pp. 429-452
- Lewis B.R. (1996), "Service Promises, Problems and Retrieval. A Research Agenda", Edition, Manchester:
- Maxham J.G.I. (2001) "Service Recovery's Influence on Consumer Satisfaction, Word-of-Mouth, and Purchase Intentions", *Journal of Business Research*, Vol 54 No October, pp. 11-24
- Maxham J.G.I. and Netemeyer R.G. (2002) "A Longitudinal Study of Complaining Customers' Evaluations of Multiple Service Failures and Recovery Efforts", *Journal of Marketing*, Vol 66 No 4, pp. 57-71
- Maxham J.G.I. and Netemeyer R.G. (2002) "Modeling Customer Perceptions of Complaint Handling over Time: The Effects of Perceived Justice on Satisfaction and Intent", *Journal of Retailing*, Vol 78 No 4, pp. 239-252
- Maxham J.G.I. and Netemeyer R.G. (2003) "Firms Reap What They Sow: The Effects of Shared Values and Perceived Organizational Justice on Customers' Evaluation of Complaint Handling", *Journal of Marketing*, Vol 67 No 1, pp. 29-45
- McCollough M.A. (2000) "The Effect of Perceived Justice and Attribution Regarding Service Failure and Recovery on Post-Recovery Customer Satisfaction and Service Quality Attributes", *Journal of Hospitality & Tourism Research*, Vol 24 No 4, pp. 423-447
- McCollough M.A., Berry L.L. and Yadav M.S. (2000) "An Empirical Investigation of Customer Satisfaction after Service Failure and Recovery", *Journal of Service Research*, Vol 3 No 2, pp. 121-137
- McCollough M.A. and Bharadwaj S.G. (1992), "The Recovery Paradox: An Examination of Customer Satisfaction in Relation to Disconfirmation, Service Quality, and Attribution Based Theories", Allen C.T. (Ed.), *Marketing Theory and Applications*, Chicago, American Marketing Association, pp. 119
- Michel S. (2004) "Consequences of Perceived Acceptability of a Bank's Service Failures", *Journal of Financial Services Marketing*, Vol 8 No 4, pp. 388-400
- Michel S., Stringfellow A. and Sukumar R. (2004) "The Effects of Service Failure and Service Recovery on Customer Lifetime Value", MSI Conference at Yale University, New Haven
- Mittal V., Kumar P. and Tsiros M. (1999) "Attribute-Level Performance, Satisfaction, and Behavioral Intentions over Time: A Consumption-System Approach", *Journal of Marketing*, Vol 63 No April, pp. 88-101
- Mittal V., Ross W.T.J. and Baldasare P.M. (1998) "The Asymmetric Impact on Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions", *Journal of Marketing*, Vol 62 No 1, pp. 33-47
- Oliver R.L. (1980) "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions", *Journal of Marketing Research*, Vol 17 No November, pp. 460-469
- Oliver R.L. (1996), *Satisfaction-A Behavioral Perspective on the Consumer*, New York et al., McGraw Hill,
- Richins M.L. (1983) "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study", *Journal of Marketing*, Vol 52 No January, pp. 93-107
- Rossiter J.R. (2002) "The C-OAR-SE Procedure for Scale Development in Marketing", *International Journal of Research in Marketing*, Vol 19 No 4, pp. 305-335

- Rust R.T., Lemon K.N. and Zeithaml V.A. (2004) "Return on Marketing: Using Customer Equity to Focus Marketing Strategy", *Journal of Marketing*, Vol 68 No 1, pp. 109-127
- Rust R.T. and Oliver R.L. (2000) "Should We Delight the Customer?" *Journal of the Academy of Marketing Science*, Vol 28 No 1, pp. 86-94
- Singh J. (1990) "Voice, Exit and Negative Word-of-Mouth Behaviors: An Investigation Across Three Service Categories", *Journal of the Academy of Marketing Science*, Vol 18 No 1, pp. 1-15
- Smith A.K. and Bolton R.N. (1998) "An Experimental Investigation of Customer Reactions to Service Failures: Paradox or Peril?" *Journal of Service Research*, Vol 1 No 1, pp. 65-81
- Smith A.K., Bolton R.N. and Wagner J. (1999) "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery", *Journal of Marketing Research*, Vol 36 No August, pp. 356-372
- Spreng R.A., Harrell G.D. and Mackoy R.D. (1995) "Service Recovery: Impact on Satisfaction and Intentions", *Journal of Services Marketing*, Vol 9 No 1, pp. 15-23
- Swanson S.R. and Kelley S.W. (2001) "Service Recovery Attributions and Word-of-Mouth Intentions", *European Journal of Marketing*, Vol 35 No 1/2, pp. 194-211
- Tax S.S. and Brown S.W. (1998) "Recovering and Learning from Service Failures", *Sloan Management Review*, Vol No Fall, pp. 75-88
- Tax S.S., Brown S.W. and Chandrashekar M. (1998) "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing", *Journal of Marketing*, Vol 62 No April, pp. 60-76
- Webster C. and Sundaram D.S. (1998) "Service Consumption Criticality in Failure Recovery", *Journal of Business Research*, Vol 41 No 2, pp. 153-159
- Zeithaml V.A., Berry L.L. and Parasuraman A. (1996) "The Behavioral Consequences of Service Quality", *Journal of Marketing*, Vol 60 No April, pp. 31-46
- Zeithaml V.A. and Bitner M.J. (1996), *Services Marketing*, New York et al., McGraw Hill,
- Zhu Z., Sivakumar K. and Parasuraman A. (2004) "A Mathematical Model of Service Failure and Recovery Strategies", *Decision Sciences*, Vol 35 No 3, pp. 493-525

TABLE 1

LITERATURE REVIEW: EMPIRICAL STUDIES TESTING THE SERVICE RECOVERY PARADOX

Author(s)	Methodology, Sampling, Statistics	Main Results	Paradox
Bolton and Drew (1992)	Telephone survey of 1,064 small-business customers of a telecommunications service. Between-subject. Regression analysis.	A repair incident that is rated as “excellent” causes a recovery paradox.	Yes
Boshoff (1997)	Scenario-based experiment in the airline industry, 540 international tourists. Between-subject. ANOVA.	Service recovery paradox was found when the supervisor immediately offered the customer a full refund and an additional free airline ticket. Recovery paradox found.	Yes
Hocutt, Chakraborty and Mowen (1997)	2 × 2 × 2 factorial design experiment with 251 students in a restaurant setting. Between-subject. MANOVA.	Paradox not found when it was the provider’s fault, but the paradox was found when the mistake was customer’s fault.	Yes
Smith and Bolton (1998)	Written survey based on failure/recovery encounter scenarios in hotels (602 respondents) and restaurants (375 respondents). Within-subject. Mean analysis.	Cumulative satisfaction and repatronage intention after a very satisfactory service recovery is higher than prior cumulative satisfaction and repatronage intention.	Yes
McCollough (2000)	2 × 2 factorial design experiment with 128 students in a hotel setting. Between-subject. ANOVA and multiple linear regressions.	A recovery paradox with respect to transaction satisfaction is possible after a low-harm service failure where complete recovery is possible (e.g., room upgrade because of overbooking).	Yes
Maxham and Netemeyer (2002)	Longitudinal study with 255 complaining bank customers at four points in time. Within-subject. MANCOVA.	Recovery paradox found for one failure and recovery. No double deviation effect for one failure and dissatisfactory recovery, but strong effect after two failures.	Yes
Hocutt, Bowers and Donovan (2006)	2 × 2 × 2 factorial design experiment with 211 students in a restaurant setting. Between-subject. MANOVA.	Paradox was found only for best recovery scenario compared to no failure scenario.	Yes
Berry, Zeithaml and Parasuraman (1990)	Survey of 1,936 customers in different industries. Between-subject. Mean analysis.	“No service problem” is better than “service problem resolved satisfactorily.”	No
Halstead and Page (1992)	Survey of carpet buyers. Between-subject. ANOVA.	Repurchase intentions for noncomplaining satisfied customers is higher than for complaining customers who are satisfied with the complaint handling.	No
Brown, Cowles and Tuten (1996)	Experimental design in a retail setting with 424 students. Between-subject. ANOVA.	Service recovery has a positive impact on encounter satisfaction, but reliability is important for long-term success.	No
Zeithaml, Berry and Parasuraman (1996)	Customer surveys in four industries, n = 1009–3069. Between-subject. Regression, ANOVA.	No problem is better than good recovery, which is better than bad recovery.	No
Bolton (1998)	Longitudinal study of 599 cellular telephone customers. Proportional hazards regression. Within-subject.	Customers who experienced perceived gains during service encounters do not have longer duration times, even if customers perceived the encounter to have been handled in a “very satisfactory” manner.	No
McCollough, Berry and Yadav (2000)	Scenario-based experiments in an airline setting. Written survey in the airport with 615 airline travellers. Within-subject. LISREL, ANCOVA (GLM), ANOVA (GLM).	Recovery paradox for transaction-based satisfaction is rejected. “Harm” should be taken into account.	No
Maxham (2001)	Study 1: Experiment with 406 students in a haircut setting. Study 2: Survey of 116 complainers of an Internet service provider. Within-subject. MANOVA.	No significant differences on satisfaction and repurchase intention between “high” and “moderate” service recovery, but significant differences on word of mouth.	No
Andreassen (2001)	Telephone interviews in various industries (self-selected) based on the Norwegian Customer Satisfaction Barometer (NCSB). Between-subject. ANOVA.	Moderate degree of satisfaction with the recovery makes up for the service failure. Image is restored more easily than intent. Even with very high scores of satisfaction with the recovery, image and intent were not higher than for satisfied customers.	No

TABLE 2

IMPACT OF DISCONFIRMATION OF SERVICE RECOVERY ON OVERALL SATISFACTION WITH BANK AND RECOMMENDATION INTENTIONS

Groups		Overall Satisfaction (1 = very dissatisfied, 5 = very satisfied)	Recommendation Intention (1 = very unlikely to recommend, 5 = very likely to recommend)
RECOVERY -- (n = 60)	Mean	3.30	3.20
	SD	1.109	1.338
RECOVERY - (n = 174)	Mean	3.74	3.81
	SD	.871	.927
RECOVERY 0 (n = 718)	Mean	4.01	4.05
	SD	.737	.971
RECOVERY + (n = 141)	Mean	4.12	4.12
	SD	.722	1.038
RECOVERY ++ (n = 63)	Mean	4.29	4.46
	SD	.705	.714
NOFAILURE -- (n = 35)	Mean	3.51	3.54
	SD	1.121	1.421
NOFAILURE - (n = 86)	Mean	3.77	3.78
	SD	.807	1.162
NOFAILURE 0 (n = 275)	Mean	3.84	3.89
	SD	.710	.995
NOFAILURE + (n = 3415)	Mean	4.13	4.28
	SD	.557	.817
NOFAILURE ++ (n = 4507)	Mean	4.56	4.54
	SD	.575	.736
Total (n = 9474)	Mean	4.30	4.35
	SD	.663	.845

Notes: SD = standard deviation.

FIGURE 1

AVERAGE OVERALL SATISFACTION SCORE BY GROUPS

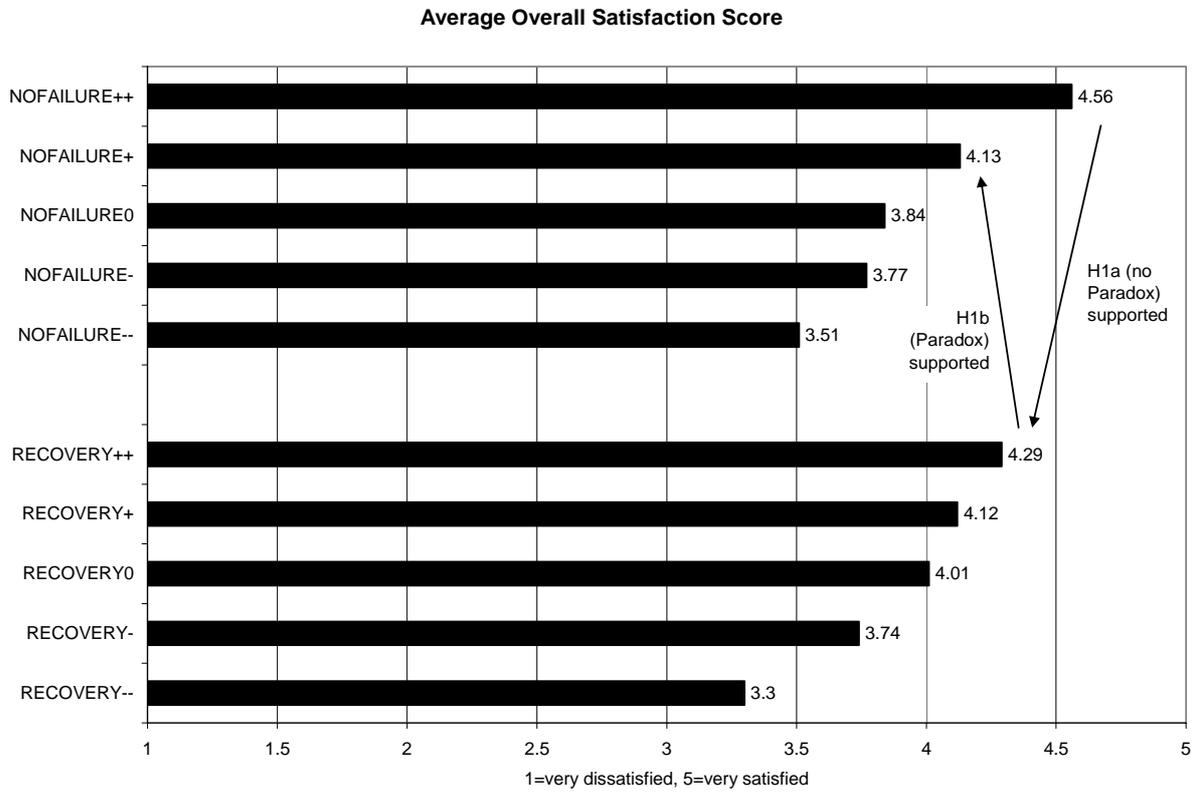


FIGURE 2

AVERAGE RECOMMENDATION INTENTION SCORE BY GROUPS

