THE EFFECTS OF SERVICE FAILURE AND SERVICE RECOVERY ON CUSTOMER LOYALTY IN E-SERVICES: AN EMPIRICAL INVESTIGATION

Rui Sousa[a] and Christopher A. Voss[b]

[a] *Catholic University of Portugal (Porto), Rua Diogo Botelho, 1327, 4169 - 005 Porto, Portugal, Tel: +351-22 619 6200, Fax: +351-22 619 6291, E-mail: rsousa@porto.ucp.pt*
[b] *London Business School, Sussex Place, Regent's Park, London NW1 4SA, United Kingdom, Tel: +44 - (0)20 7262 5050, Fax: +44 - (0)20 7724 7875, E-mail: cvoss@london.edu*

ABSTRACT
Although there has been a lot of research examining the effect of service failures and recovery on customer loyalty in traditional (bricks-and-mortar) services, there is still little rigorous empirical research examining this in e-service settings. The objective of this study is to empirically examine the impact of service failures and service recovery on customer loyalty in e-services. The study is based on data from an online questionnaire of customers of an e-banking service. The results validated the relationships generally found in traditional services: i) service failures result in decreased customer loyalty intentions; ii) effective recovery from failures increases customer loyalty intentions; and iii) unless recovery efforts elicit very high levels of satisfaction from customers, they can, at best, restore loyalty to the levels existing prior to the failures. The implications for the management of e-services are discussed.

Keywords: E-Services, Service Recovery, Loyalty.

INTRODUCTION
There has been substantial research in traditional services on the effect of service failures and recovery on customer loyalty. This work has produced three main findings. First, service failures have a negative effect on loyalty and have been found to be a driving factor in customer switching behavior (e.g., McCollough et al, 2000; Zeithaml et al, 1996). Second, in the event of a service failure, customers expect effective recoveries (Bitner et al, 1990) and their satisfaction with the recovery increases customer loyalty (Zeithaml et al, 1996; Miller et al, 2000; Tax and Brown, 1998). Third, it is not clear whether a "recovery paradox" exists, i.e., whether customers who experience a service failure followed by superior recovery exhibit behavioral intentions towards the service provider which are more favorable than they would be had no failure occurred (service recovery seen as an opportunity). While several studies report results consistent with a recovery paradox effect (e.g., Hansen and Danaher, 1999; Smith and Bolton, 1998), some empirical studies have cast doubts over it (Zeithaml et al, 1996; McCollough et al, 2000; Andreassen, 2001).

Although these service failure and recovery issues have received considerable attention in the literature, these topics have received only limited attention in the context of online services. Specifically, we lack an understanding of whether the knowledge gained in traditional services translates to online environments, given the absence of human intervention in the service encounter and the fact that the reasons for dissatisfactory online service encounters have been found to be different from traditional offline service research (Holloway and Beatty, 2003).

1

As such, the main objective of this paper is to empirically examine the impact of service failures and service recovery on customer loyalty in e-services. The structure of the paper is as follows. First, we review the literature in traditional services and e-services and develop research hypotheses related to the study's research question. Second, we describe the employed methodology, a survey study of customers of an e-banking service. Third, we present the data analysis. Finally, we discuss the results, main conclusions, limitations and avenues for future research.

LITERATURE REVIEW AND RESEARCH HYPOTHESES

In developing the research hypotheses, our default stance is to apply theory from traditional services to e-service settings. A prominent study which has addressed most of the discussed failure/recovery relationships in traditional services has been Zeithaml et al (1996). The authors conducted a mail survey of actual customers of four different companies to find that customer loyalty was highest for customers experiencing no service problems; next highest for customers experiencing service problems that were resolved, and lowest for customers experiencing service problems that were not resolved. Taking the stance that theory from traditional services will find replication in e-services, we adapt Zeithaml et al's (1996) hypotheses by adding a variable representing the degree with which customers were satisfied with the way the problem was solved, thus putting forward the following hypotheses:

H1. Loyalty behavior will be higher for customers experiencing no service problem (A) than for those experiencing service problems (B).

H2. Within the customer group who have experienced service problems (B), loyalty behavior will be higher for those for whom the problem has been resolved (B2), than for those for whom the problem has not been resolved (B1).

H3. Within the customer group who have experienced service problems and have had the problems resolved (B2), loyalty behavior will be higher for customers who are more satisfied with the way the problem was resolved (B2.1) than for those who are less satisfied (B2.2).

In addition, we test the service recovery paradox in an e-service context by adding the following hypothesis:

H4. Loyalty behavior will be higher for customers who have experienced service problems and have had the problems resolved with a high degree of satisfaction (B2.1) than for those who have not experienced service problems (A).

Failure and recovery issues in e-services exhibit some differences relative to traditional services. It is largely an empirical question whether these differences will affect the verification of the proposed hypotheses. Concerning H1, it has been found that the types of service failures online are considerably different from traditional services (Holloway and Beatty, 2003). Whether this hypothesis holds in e-services may depend on the degree of severity that customers attach to service failures (Craighead et al, 2004) and the extent to which customers may be more or less forgiving towards online failures (McCollough et al, 2000).

The validity of H2-H4 largely hinges on the degree of effectiveness of the service recovery that can be achieved in e-service settings. If recovery effectiveness can reach high levels, then problems may be speedily and adequately resolved to the extent that problem resolution per se as well as the quality of the resolution may have an impact on customer loyalty (support for H2 and H3). Moreover, customer loyalty might even emerge as strengthened relative to the level existing prior to the occurrence of the service failure (support for H4). Some studies suggest that the effectiveness of

service recovery can be increased in e-services, for two main reasons. First, the multi-channel nature of many of these services may facilitate customer complaints by offering new channels that reduce the time and effort required in the process (Tax and Brown, 1998; Holloway and Beatty, 2003) or remove the embarrassment of the face-to-face complaining process (Tax and Brown, 1998). Thus, as barriers to complaints are lowered, the complaining frequency is expected to increase, providing more opportunities for service providers to salvage dissatisfied customers. Second, the use of technological support to handle complaints (e.g., FAQs and troubleshooting engines that can automatically walk customers through problem-identification and resolution processes) may lead to more efficient service recovery systems (Tax and Brown, 1998).

However, there is some evidence of overall customer dissatisfaction with the service recoveries provided by online retailers (Holloway and Beatty, 2003) and other studies have suggested that it may be more difficult to perform effective recovery in e-service contexts. This is due to two types of reasons. First, because of the different nature of service failures online. Meuter et al (2000) identified three types of failures in e-services - technology failures, process failures and customer-driven failures - and acknowledged the difficulty in performing service recovery in such a context. Technology failures are those that prevent the customer from engaging with the e-service (e.g., web site is down or not working properly). In these cases, the customer may resort to interpersonal service either to complain or to receive the desired service (Meuter et al, 2000). Because of the wide reach of the Internet, there are a potentially large number of users affected by the breakdown and it may be difficult for the service provider to enable convenient and rapid customer access to a support agent (Sousa and Voss, 2006). This may compromise the speed of response to a service failure, a key part of maintaining loyalty (Miller et al, 2000). Process failures are those that occur at some point after the customer's interaction with the web site, but preventing correct service fulfillment from occurring. In these cases, because the initial interaction has taken place as expected, the customer expects the service to be provided successfully, and the failure only becomes apparent at a later stage (e.g., the items ordered through the internet are never received). Because of this deferred nature of fulfillment, it is difficult to detect such failures early and proactively contacting the customer to overcome them (Sousa and Voss, 2006). Customer-driven failures are those that occur as a result of a customer mistake on the web site (e.g., not being able to understand and complete the steps required for placing an order online). Because of the absence of customer contact, it is difficult to detect customer difficulties or mistakes in a timely manner and perform on-the-spot customer support (Bitner et al, 2002).

The second type of reasons is the reduced degree of human interaction. Findings from traditional services have shown that the role of front-line employees is key for effective recovery. A customer's anger is abated when employees act in a polite and empathetic manner and demonstrate a strong effort to solve the problem (Tax et al, 1998). An existing rapport between the customer and the service provider (achieved through the face to face interactions with service employees) also increases the likelihood of a satisfactory recovery from a service failure, acting as a switching barrier (DeWitt and Brady, 2003). On the contrary, e-services encounters are depersonalized and the technology creates a distance between customers and service personnel (Walker et al, 2002). Although e-service providers may offer access to human customer support agents, this typically happens via low richness media, such as e-mail and the phone.

METHODOLOGY

The study consisted in the administration of an online questionnaire to a sample of the customers of a major retail e-banking service. The questionnaire measured customers' loyalty intentions through a 1 (Very Unlikely) – 5 (Very Likely) rating on two items: the intention to revisit the site and whether customers would recommend the site to friends or relatives. It also included three cascaded questions related to the research hypotheses: SFR1) Whether a customer had experienced a service

problem in the last 6 months; SFR2) If so, whether it had been resolved; and SFR3) If so, the customer was asked to rate the satisfaction with the way the problem had been resolved.

A random sample of 35781 customers were targeted. The questionnaire was posted on the e-banking web site, right past the login stage and was active on the site for one month, resulting in 5,942 responses (a 16.6% response rate).

The final sample is characterized in Table 1. The table shows that the predominant demographic profile in the sample was that of a male, young and educated customer. This pattern is in line with the patterns observed for general Internet users in the European Union and the US, as well as for e-banking users (e.g., Eurostat, 2003). We conducted a non-respondent bias analysis employing t-tests to compare the profiles of respondents and non-respondents in terms of age, gender and education level. The analysis showed no significant differences, indicating the absence of non-respondent bias.

Table 1 - Characterization of the final sample (n=5,942).

Variable	Total Sample n=5942
Age	
[18-25[	9.4 %
[25-35[	44.8 %
[35-45[	22.1 %
[45-55[	14.2 %
55+	9.5 %
Gender	
Male	72.0 %
Female	28.0 %
Education	
Primary education	3.4 %
Secondary education	37.9 %
Higher Education (Bachelor's degree and above)	58.7 %

ANALYSIS AND RESULTS

In order to test hypotheses H1-H4, the combined sample was classified into different customer groups obtained by splitting the sample according to the responses to variables SFR1 (Groups A and B), SFR2 (Groups B1 and B2) and SFR3 (Groups B2.1 and B2.2) – see Figure 1. The hypotheses tests consisted in comparing the means of the summated loyalty scales across the different groups. The summated scales were obtained by taking the average of the measurement items, which assumes that the individual item weights are equal (Hair et al, 1998, p.129). Analysis of variance revealed that loyalty scores differed across the mutually exclusive groups A, B1, B2.1 and B2.2 (F-value significant at $p < 0.000$). The results of the individual group comparisons are shown in Table 2. Because multiple comparisons were evaluated to test these hypotheses, the alpha level for testing the significance of individual comparisons was reduced by applying the Bonferroni correction to ensure that the overall probability of Type I error across all four comparisons did not exceed .05 (for details, see footnote b) in Table 2).
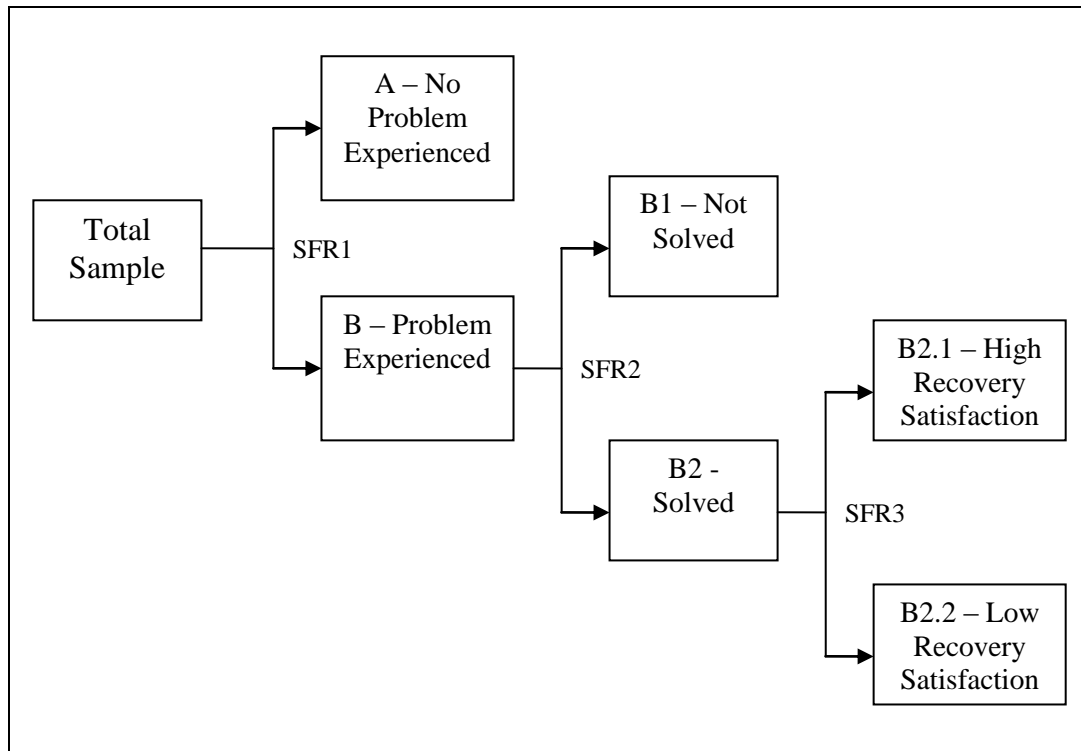
Figure 1 - Classification of total sample into different customer groups.

Table 2 - Comparison of the means of the summated loyalty scale across different customer groups (Hypotheses H1-H4).

Customer Group	N	% Within Groups	Mean	Standard Deviation	P-Values (b)	Conclusion
Total Sample	5942	--	4.12	0.79	--	--
SFR1	5942					
A – No problem experienced.	5030	84.7%	4.16	0.77	A vs. B:	H1
B – Problem experienced.	695	11.7%	3.94	0.75	<0.000	Supported
No Response to SFR1	217	3.6%	--	--		
SFR2	695					
B1 – Problem not solved	267	38.4%	3.82	0.95	B1 vs. B2:	H2
B2 – Problem solved	389	56.0%	4.01	0.80	0.005	Supported
No Response to SFR2	39	5.6%	--	--		
SFR3	389					
B2.1 High satisfaction with recovery (a)	145	37.3%	4.21	0.64	B2.1 vs. B2.2: <0.000	H3 Supported
B2.2 Low satisfaction with recovery (a)	219	56.3%	3.85	0.85	B2.1 vs. A: 0.147	H4 Not Supported
No Response to SFR3	25	6.4%				

(a) Groups B2.1 and B2.2 were obtained by splitting group B2 by the mean (3.28).

(b) P-values correspond to one-tailed t-tests for group means. Because multiple comparisons were evaluated to test the hypotheses, the Bonferroni correction was applied to the customary alpha level of .05 to control the Type I error rate. Specifically, the alpha level was lowered by a factor of four (the total number of comparisons) to yield a critical alpha level of .0125 for testing the significance of each comparison (Hair et al, p. 328).

Table 2 shows that there is support for H1, H2 and H3. Although loyalty was higher for group B2.1 than group A, the difference was not significant, and the results do not provide definite support for H4. The non-dichotomous nature of variable SFR3 allowed for additional more detailed analyses concerning H4. We compared the means of the summated loyalty scales between each of five individual customer groups corresponding to the five possible reported levels of satisfaction with recovery (variable SFR3) and Group A (see Table 3). Table 3 shows that whether loyalty levels after recovery are higher or lower relative to error-free service levels, depends on what the level of satisfaction with recovery is. Specifically, loyalty levels after recovery only surpass error-free levels for the maximum possible level of satisfaction with recovery (5). For a satisfaction level of 4, loyalty is slightly lower than for Group A, although this difference is not statistically significant. For satisfaction levels of 3 or below, customer loyalty is significantly lower relative to Group A. Therefore, there is only partial support for H4; that is, there is only evidence of the "recovery paradox" for the maximum possible level of satisfaction with recovery, but not for other levels.

Table 3 – Differences in the means of the summated loyalty scale between groups corresponding to different levels of satisfaction with recovery and Group A (Hypothesis H4).

Level of Satisfaction	Number of Customers	% Within Group B2	Mean	Std Dev	Difference in Means Relative to Group A	P-Value (a)	Loyalty Level vs. Group A (b)
5	29	8.0%	4.48	0.54	+0.32	0.001	Significantly Higher
4	116	31.9%	4.14	0.65	-0.02	0.433	Non-Significantly Lower
3	162	44.5%	3.91	0.85	-0.24	<0.000	Significantly Lower
2	43	11.8%	3.82	0.74	-0.33	0.003	Significantly Lower
1	14	3.8%	3.29	0.99	-0.87	0.003	Significantly Lower

(a) One-tailed t-tests.
(b) The Bonferroni correction was applied by lowering the alpha level by a factor of five (the total number of comparisons) to yield a critical alpha level of .01 for testing the significance of each comparison (Hair et al, p. 328).

CONCLUSIONS

These results have several implications for e-service settings. First, support for H1 shows that the detrimental effects of service failures are also present online. This may mean that customers do not attach a less degree of severity to failures nor seem to be more forgiving when interacting with e-services when compared to traditional services. Because failures in e-services have the potential to affect a large number of customers, this highlights the importance of investing in the prevention of failures in these settings.

Second, the support received by hypotheses H2 and H3 suggests that it is possible to achieve effective recovery in e-services. Thus, investing in service recovery efforts is a good investment from a customer loyalty perspective. Inferior recovery performance can lead to what Bitner et al (1990) termed a *double deviation* from customer expectations: the service provider fails to deliver on the initial service and the recovery service. Such recovery efforts must deal not only with the actual service failure (H2), but also to do so in such a way that the customer is satisfied with the way in which the problem is resolved (H3). This raises the issue of how to deliver good service recovery. What differentiates recovery that gives high satisfaction from that that has low

satisfaction? Our data cannot address this problem, but this is clearly an important area for future research.

Finally, the recovery paradox (H4) only manifested itself for the maximum level of satisfaction with the recovery effort. This reinforces the detrimental effect of failures in e-services: for levels of recovery satisfaction below the very highest, at best, we may be able to restore loyalty to levels existing prior to the failures. Thus, there is only weak support for considering recovery in e-services as an "opportunity" when compared to the loyalty level resulting from error-free service. Thus, in general, e-service providers may be better off avoiding service failures than responding to failure with superior recovery.

The fact that the research hypotheses, largely based on theory from traditional services, were broadly supported, further extends the generalizeability of research in traditional services into the context of online service. Thus, the article provides general support in defense of theories that draw parallels between online and offline services.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

The study was based on the investigation of actual customers in a real business setting representative of a pure information-based service, an approach that had been lacking in existing empirical research. The empirical conclusions in this study are based on findings from one service industry, e-banking, a very important type of service in today's e-service landscape. We believe that the findings can be generalized to other task-oriented e-services, but caution must be exercised in extending the conclusions of this study to other services. It may be important for future research to test the developed hypotheses in other types of e-services, in particular, types of e-services which are more strongly associated with experiential/hedonic use (e.g., entertainment services).

In our study we have focused on loyalty behavioral intentions, what Oliver (1999) called conative loyalty. Although previous research has provided empirical support for the causal link between intentions and actual actions (Venkatesh and Davis, 2000), future research may examine if loyalty behavior is linked to loyalty actions (what Oliver (1999) calls action loyalty).

The findings stressed the importance of recovery strategies for customer loyalty. As discussed earlier, setting up such strategies in e-services may be a challenge (Meuter et al, 2000) and, if they require human intervention, may decrease the level of scalability of the service and increase costs. Our study does not address the issue of whether it is cost-effective to employ these strategies. Future research should investigate the costs and benefits associated with recovery strategies. We also need to know more about which online recovery strategies (e-mail apologies, discounts on future transactions, personal phone calls and letters) are best suited for which online service failures.

REFERENCES

Andreassen, T. (2001), "From disgust to delight", *Journal of Service Research*, Vol. 4, No. 1, pp. 39-49.

Bitner, M., Ostrom, A., Meuter, M. (2002), "Implementing successful self-service technologies", *Academy of Management Executive*, Vol. 16, No. 4, pp. 96-108.

Bitner, M., Booms, B., Tetreault, M. (1990), "The service encounter: Diagnosing favorable and unfavorable incidents", *Journal of Marketing*, Vol. 54, No. 1, pp. 71-84.

Craighead, C., Karwan, K., Miller, J. (2004), "The effects of severity of failure and customer loyalty on service recovery strategies", *Production and Operations Management*, Vol. 13, No. 4, pp. 307-321.

DeWitt, T., Brady, M. (2003), "Rethinking service recovery strategies: The effect of rapport on consumer responses to service failure", *Journal of Service Research*, Vol. 6, No. 2, pp. 193-207.

Eurostat (2003), *Statistics on the information society in Europe*, report available at http://europa.eu.int/comm/enterprise/ict/studies/is-stat-96-02.pdf.

Hair, J., Anderson, R., Tatham, R., Black, W. (1998), *Multivariate Data Analysis*, 5th Edition, Prentice-Hall, Inc., New Jersey.

Hansen, D., Danaher, P. (1999), "Inconsistent performance during the service encounter: What's a good start worth?", *Journal of Service Research*, Vol. 1, February, pp. 227-235.

Holloway, B., Beatty, S. (2003), "Service failure in online retailing: A recovery opportunity", *Journal of Service Research*, Vol. 6, No. 1, pp. 92-105.

McCollough, M., Berry, L., Yadav, M. (2000), "An empirical investigation of customer satisfaction after service failure and recovery", *Journal of Service Research*, Vol. 2, November, pp. 121-137.

Meuter, M., Ostrom, A., Roundtree, R., Bitner, M. (2000), "Self-service technologies: Understanding customer satisfaction with technology-based service encounters", *Journal of Marketing*, Vol. 64, No. 3, pp. 50-64.

Miller, J., Craighead, C., Karwan, K. (2000), "Service recovery: A framework and empirical investigation", *Journal of Operations Management*, Vol. 18, No. 4, pp. 387-400.

Oliver, R. (1999), "Whence consumer loyalty?", *Journal of Marketing*, Vol. 63, No. 4, pp. 33-44.

Smith, A., Bolton, R. (1998), "An experimental investigation of customer reactions to service failure", *Journal of Service Research*, Vol. 1, No. 1, pp. 65-81.

Sousa, R., Voss, C. (2006), "Service quality in multi-channel services employing virtual channels", *Journal of Service Research*, Vol. 8, No. 4, pp. 356-371.

Tax, S., Brown, S. (1998), "Recovering and learning from service failure", *Sloan Management Review*, Fall, pp. 75-88.

Venkatesh, V., Davis, F. (2000), "A theoretical extension of the technology acceptance model: Four longitudinal field studies", *Management Science*, Vol. 46, No. 2, pp. 106-204.

Walker, R., Carig-Lees, M., Hecker, R., Francis, H. (2002), "Technology-enabled service delivery, an investigation of reasons affecting customer adoption and rejection", *International Journal of Service Industry Management*, Vol. 13, No. 1, pp. 91-106.

Zeithaml, V., Berry, L., Parasuraman, A. (1996), "The behavioral consequences of service quality", *Journal of Marketing*, Vol. 60, No. 2, pp. 31-46.