Identifying Appropriate Compensation Types for Service Failures: A Meta-Analytic and Experimental Analysis

Holger Roschk[1] and Katja Gelbrich[1]

Abstract

This article examines how compensation type and failure type explain the recovery effect of compensation, using a meta-analysis (Study 1) and an experiment (Study 2). Drawing on resource exchange theory, we propose new classifications for both compensation and failure type and find three major results. First, consistent with our matching hypothesis, the strongest recovery effect is generally observed when compensation represents a resource similar to the failure it is supposed to offset, that is, immediate monetary compensation for a monetary failure, exchange for a flawed product, reperformance for a failed service, and psychological compensation for lack of attention. Surprisingly, lack of attention may also be rectified by the other compensation types. Second, consistent with our intertemporal choice hypothesis, immediate monetary compensation is generally more effective than delayed monetary compensation. Yet, this effect also varies with failure type. Third, resource-based classifications explain the recovery effect of compensation better than current classifications of compensation and failure type. As a theoretical contribution, the resource-based classifications help to explain the fluctuating effect sizes of compensation reported in prior research. From a managerial point of view, practitioners can choose the appropriate compensation type for a failure, one that repays in kind what customers have lost. As a result, companies achieve stronger recovery effects without additional costs.

Introduction

Compensating disgruntled customers is the most powerful way to offset a company's failure (Gelbrich and Roschk 2011a). Numerous studies report positive effects of compensation on customer reactions like satisfaction (SAT), loyalty (LOY), or positive word of mouth (WOM; e.g., Bonifield and Cole 2008; Mattila and Patterson 2004; Wirtz and Mattila 2004). However, the effect fluctuates from nonsignificant (Garrett 1999) to small/medium (Grewal, Roggeveen, and Tsiros 2008) or very large (Hess, Ganesan, and Klein 2003).

We argue that effect sizes mainly fluctuate because extant quantitative studies use different failure and compensation types identified in exploratory research (Bitner, Booms, and Tetreault 1990; Kelley, Hoffman, and Davis 1993). The failure types examined range from monetary harm like charges by the phone company for unsubscribed features (Liao 2007), to flawed goods like shoes falling apart (Blodgett, Hill, and Tax 1997), to a failed service like an overcooked steak (Hess, Ganesan, and Klein 2003), or inattentive employees giving undue preference to other customers (Bonifield and Cole 2008). Similarly, various compensation types have been examined. They may differ in kind, such as discount/money back (Webster and Sundaram 1998), exchange/reperformance (Blodgett, Hill, and Tax 1997), or apology (Liao 2007). Further, compensation types vary with

regard to the point in time when remuneration is provided. For example, a coupon toward the next flight (Grewal, Roggeveen, and Tsiros 2008) represents delayed compensation that is tied to future purchase, whereas a free dinner (Bonifield and Cole 2008) represents an immediate compensation that is tied to the current purchase. Both categories are known from the intertemporal choice literature (Pyone and Isen 2011).

In the light of this variety, we seek to provide an answer to the following question. Which compensation—both in kind and in point in time—is most effective to recover from a particular type of failure? Drawing on resource exchange theory (Foa and Foa 1974, 1976), we consider compensation as a resource that recompenses customers for a resource they lost due to the organization's failure (Orsingher, Valentini, and de Angelis 2010; Smith, Bolton, and Wagner 1999). We suggest new, resource-based classifications of compensation type and

---

[1] International Management at Catholic University Eichstaett–Ingolstadt, Ingolstadt, Germany

Corresponding Author:
Holger Roschk, International Management at Catholic University Eichstaett–Ingolstadt, Auf der Schanz 49, 85049 Ingolstadt, Germany.
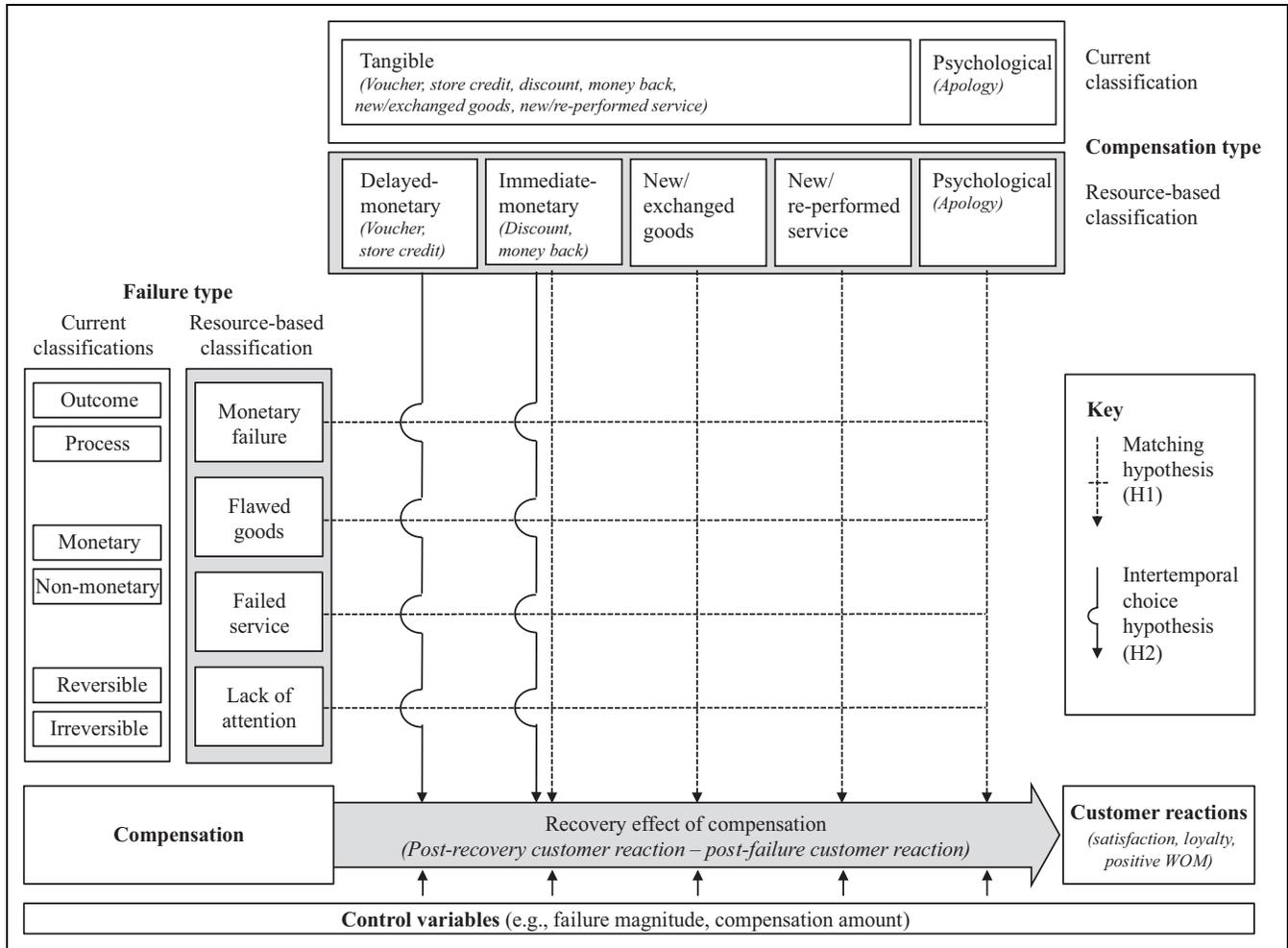Email: holger.roschk@ku.de

Figure 1. Theoretical framework.

failure type and examine in what way they influence the recovery effect of compensation through a meta-analysis and an experiment.

Hereby, we make three contributions to the literature. First, in matching type of compensation to type of failure based on exchanged resources, we may uncover an adequate form of remuneration for each failure situation. In particular, we propose new classifications of compensation type and failure type that match the resource categories used in resource exchange theory. Second, these resource-based classifications are compared to current classifications for compensation type, namely, tangible versus psychological, and failure type, namely, outcome versus process, monetary versus nonmonetary, and reversible versus irreversible failure. Third, we shed light on the fluctuating effect sizes of compensation on customer responses in previous studies. Based on these contributions, marketing managers learn that recovery from a particular failure may not only be achieved by increasing compensation amount, as suggested in prior research (Gelbrich and Roschk 2011b). Instead, choosing the right time and a kind of remuneration may serve the same purpose and cut recovery costs.

Theoretical Background

Overview

Figure 1 shows our theoretical framework, which suggests a temporal sequencing of failure and recovery (Maxham and Netemeyer 2002b). First, a failure deteriorates favorable customer reactions like SAT, LOY, and positive WOM to a post-failure level. Then, organizations provide compensation. This recovery effort mitigates the previous negative effect of the failure by increasing customer reactions to a postrecovery level, and the respective increase represents the recovery effect of compensation (at the bottom of Figure 1).

The matching hypothesis proposes that the effect is strongest when exchanged resources match in kind, advancing that compensation type and failure type interact on the recovery effect of compensation. The "intersections" in Figure 1 illustrate this hypothesis. The intertemporal choice hypothesis proposes that the recovery effect of compensation depends on the point in time for resource exchange, advancing different recovery effects for delayed versus immediate compensation. The two continuous vertical arrows in Figure 1 illustrate this hypothesis.

The bottom of Figure 1 shows the *recovery effect of compensation*, indicated by the relationship between compensation and customer reactions. Pertinent customer reactions reported in post-complaint research are customer SAT, LOY, and positive WOM (Maxham and Netemeyer 2002a; Orsingher, Valentini, and de Angelis 2010; Smith, Bolton, and Wagner 1999). The recovery effect of compensation is defined as the difference between post-recovery customer reactions and postfailure customer reactions, that is, customer reaction after and before compensation has been provided (Boshoff 1997; Maxham and Netemeyer 2002b). It can also be seen as a net effect explaining the increase in customer reactions caused by compensation, after taking into account the decrease in customer reactions caused by the failure.

The upper part of Figure 1 represents *compensation type*. We define compensation type as the form of benefit complainants receive from the company following a failure. The current literature distinguishes between tangible and psychological compensation (Gelbrich and Roschk 2011a). Tangible compensation represents a material benefit like a voucher/coupon (Weaver, Garcia, and Schwarz 2012), store credit (Lee and Park 2010), discount (Sparks and McColl-Kennedy 2001), money back (Estelami 2000), exchange/replacement (Kelley, Hoffman, and Davis 1993), or reperformance (Hess, Ganesan, and Klein 2003). Psychological compensation refers to an emotional benefit, which typically comes as an apology (Davidow 2003).

The left part of Figure 1 represents *failure type*. We define failure type as the form of a company's flaw when delivering a product or service. Current failure type classifications are outcome versus process failure (Smith, Bolton, and Wagner 1999; Zhu, Sivakumar, and Parasuraman 2004) and monetary versus nonmonetary failure (Gelbrich and Roschk 2011a; Gilly and Gelb 1982). Outcome failure like an overcooked steak refers to the result of product or service delivery; process failure like undue preference of other guests in a restaurant refers to the way of product or service delivery (Smith, Bolton, and Wagner 1999). Monetary failure, such as faulty tires, involves a financial loss, while a nonmonetary failure, such as a delayed flight, does not (Gilly and Gelb 1982). Although not explicitly stated, some studies present irreversible or nonrepairable failures like an unavailable meal (Smith, Bolton, and Wagner 1999). Other failures like incorrect restaurant order are reversible. Thus, we include irreversible versus reversible failure as a current classification.

Drawing on resource exchange theory (Foa and Foa 1974), this article provides new, resource-based classifications of compensation type and failure type and tests them as moderators on the recovery effect of compensation. These results will then be compared to the effect based on current classifications, including control variables like failure magnitude (Smith, Bolton, and Wagner 1999) and compensation amount (Gelbrich and Roschk 2011b).

Resource Classes According to Resource Exchange Theory

Resource exchange theory posits that social interactions may entail the exchange of the following resources (Foa and Foa 1974): money, goods, services, love, or status. Money is defined as a currency or token with an exchange value. Goods represent objects, products, or materials. Services refer to activities on the body or belongings of a person. Love means expressions of affectionate concern. Finally, status entails expressions of prestige and esteem.[1] These resource classes can be distinguished on two dimensions: particularism and concreteness (Foa and Foa 1974).

Concreteness is the resource's form of expression characteristic. It ranges from abstract forms, representing a symbolic value, to concrete forms, representing a tangible value (Brinberg and Wood 1983). Goods and services are most concrete. They require the exchange of a tangible product or an activity. Money has a medium concreteness because it represents an abstract figure and a coin or banknote as a physical object. Love also has a medium concreteness because it may entail tangible activities like touching and symbolic meaning like a smile. Status is rather symbolic. It is usually exchanged in abstract ways through language or facial expressions (Foa and Foa 1976).

Particularism is the degree to which the value of a resource depends on the persons involved in the exchange. High personal influence represents a particularistic resource, whereas low personal influence represents a universal resource (Brinberg and Wood 1983). Love is most particularistic. It matters a lot from whom we receive love. We highly value affectionate concern from a person we love, but we disapprove of such concern when expressed by a person we dislike. Status and service are also quite particularistic. For example, it may be more pleasant to be served by someone we like rather than by someone we dislike. Yet, both resources are less particularistic than love because they do not entail deep affection. Goods are rather universal because, unlike services, their exchange is not performed on the belongings or body of a person. Further, unlike status, their exchange does not necessarily interfere with self-esteem. Money is most universal because its value exists independent of the particular persons involved (Foa and Foa 1976).

Resource-Based Compensation Type Classification

We distinguish between monetary compensation, new/exchanged goods, new/reperformed service, and psychological compensation (see Figure 1). Monetary compensation equals "money" in resource exchange theory, that is, a currency or token representing a quantifiable amount of money (Donnenwerth and Foa 1974). Four compensation forms fall into this category: voucher (coupon for a price reduction on the next purchase), store credit (a loan disbursed with the next purchase), discount (a price reduction <100% of the purchase price), and money back (100% price reduction).

We classify new/exchanged goods as a distinct resource class, which relates to the resource "goods" in resource exchange theory. This resource is more concrete than "money" because it requires the exchange of a concrete object, such as a new pair of shoes (Foa and Foa 1974). Further, unlike money, new/exchanged goods have no exchange value per se. Instead, their financial value is materialized in a particular product.

New/reperformed service, such as serving a new meal, represents the resource class "services" in resource exchange theory. Just as new/exchanged goods, it is more concrete than money because a concrete activity is exchanged. Further, it has no exchange value per se because its financial value is materialized in a particular service.

Psychological compensation relates to "love" according to resource exchange theory because it describes some sort of affectionate concern. Psychological compensation comes in the form of an apology. It means that the company expresses regret and empathy for the customer's distress (Liao 2007). We also subsume "status" under this category because an apology reestablishes self-esteem, which may have suffered through a company's failure (Roschk and Kaiser 2013).

Further, resource exchange theory does not only specify different kinds of resources but also points out the role of time in resource exchange. In particular, "money" can be stored over time (Foa and Foa 1974). Hence, we refine our monetary compensation category. Research on monetary benefits suggests that incentives (Roehm and Roehm 2011), rebates (Pyone and Isen 2011), or compensation (Noone and Lee 2011) can be provided either in the future or in the present, that is, delayed or immediate. Of the monetary compensation forms listed earlier, voucher and store credit represent delayed monetary compensation. They are tied to a future purchase and can only be redeemed, when the consumer repurchases from the same company. In contrast, discount and money back represent immediate monetary compensation. They are tied to the present purchase and offset a failure immediately when it occurred (Kim and Ulgado 2012).

We do not consider temporal aspects for new/exchanged goods, new/reperformed service, and psychological compensation. Although goods and services theoretically can be stored, it usually makes no sense to tie the supply of a new product or service to a future transaction. Further, an apology is not stored. It acknowledges the distress complainants experience during the focal transaction (Davidow 2003).

Resource-Based Failure Type Classification

We distinguish between monetary failure, flawed goods, failed service, and lack of attention (see Figure 1). Monetary failure equals "money" in resource exchange theory and represents a financial loss, such as charging customers for unsubscribed features (Liao 2007). Flawed goods represent the resource class "goods" and refer to a defective, malfunctioning, or inferior product, such as tennis shoes falling apart (Blodgett, Hill, and Tax 1997). Although flawed goods may entail a financial loss, the monetary value of this loss is hard to quantify. To what extent does, for instance, a stain devaluate a cloak? Flawed goods differ from monetary failure in that the flaw is present in the product itself, that is, is more tangible. Service failure constitutes the resource "services" and describes a flaw in service delivery, such as an overcooked steak (Hess, Ganesan, and Klein 2003). Like goods, this failure type category entails a rather concrete loss that is hard to monetize. As opposed to

flawed goods, however, service failure describes a more particularistic resource, whose delivery requires interpersonal interactions (Zeithaml, Parasuraman, and Berry 1985). Lack of attention refers to the resource "love" and means that companies treat customers in an unfriendly, impatient, or noncaring manner (Bitner, Booms, and Tetreault 1990). An example is an undue preference of other customers (Bonifield and Cole 2008). We include "lack of status" in this category because a rude treatment is often perceived as a denigration, which damages a person's prestige (Cropanzano et al. 2001). Finally, temporal issues are irrelevant for failure type classification because failures occur in the present. They relate to the focal—rather than to a future—transaction.

Hypotheses

Type of Exchanged Resource. Resource exchange theory claims that people prefer to exchange the same or proximal resources (Foa and Foa 1974). People are more satisfied when a resource they give away matches the resource they receive in return (Donnenwerth and Foa 1974). The rationale for this claim is rooted in the concepts of justice (Adams 1965). An exchange is perceived as fair, if the input–output ratio between the two parties is balanced (Walster, Berscheid, and Walster 1973). This is because people strive for reciprocity (Gouldner 1960) and therefore seek to repay in kind "what another has done for us" (Whatley et al. 1999, p. 251). Empirical research on social exchange supports that people seek to exchange similar resources (Brinberg and Wood 1983) and that the perceived similarity of these resources increases SAT (Teichman and Foa 1975). In an economic context, product and service delivery represent particular forms of social interaction.

The postcomplaint literature has established that justice perceptions mediate the effect of organizational responses to a complaint (Gelbrich and Roschk 2011a). Following these arguments, we propose that the perceived similarity of compensation type and failure type reinforces perceived justice and thus fosters positive customer reactions. In particular, the recovery effect of compensation should be strongest in case of the following Compensation Type × Failure Type combinations: monetary compensation for a monetary failure, new/exchanged goods for flawed goods, new/reperformed service for failed service, and psychological compensation for lack of attention. Hence, we propose the following interaction effect between compensation type and failure type:

> **Hypothesis 1:** The recovery effect of compensation on customer reactions is strongest when compensation type and failure type represent the same resource.

Point in Time for Resource Exchange. Resource exchange theory suggests that storable resources like money can be exchanged at different points in time (Foa and Foa 1976). We propose that

the recovery effect of monetary compensation depends on whether it is an immediate compensation provided for the purchase at hand or a delayed compensation, which is tied to a future purchase from the same company.

A large body of research deals with temporal aspects of resource exchange, which is referred to as intertemporal choice or intertemporal evaluation (e.g., Chan and Mukhopadhyay 2010; Pyone and Isen 2011; Zaubermann et al. 2009). In a marketing context, intertemporal evaluation occurs when consumption may take place at different points in time (Zauberman et al. 2009). When evaluating different consumption times, people tend to be subject to a present bias. Their subjective value of present consumption is greater than that of future consumption (Pyone and Isen 2011). The present bias was first described in Samuelson's (1937) discounted utility model, which claims that people discount the utility of future benefits, which devaluates these benefits. A reason for this effect may be that people seek to avoid the uncertainty entailed in the future as well as negative emotions arising from deferred consumption, and they seek to enjoy a product or service immediately (Chan and Mukhopadhyay 2010).

Pertinent research largely confirms the present bias (e.g., Zauberman 2003). Further, it suggests that the present bias can be reversed when consumers are allowed to choose (Chan and Mukhopadhyay 2010) and/or when consumption is pleasurable (Loewenstein 1987). However, reverse effects can be excluded for compensation because customers are usually not allowed to choose between different compensation times, and recovery situations are not pleasurable. Some recovery studies provide initial evidence of the present bias, yet only for severe failures (Kim and Ulgado 2012), for an irreversible failure and overcompensation (Noone and Lee 2011), or for diverging compensation amounts (Sparks and McColl-Kennedy 2001). We propose that the present bias applies to a broader context:

> **Hypothesis 2:** The recovery effect of immediate monetary compensation on customer reactions is stronger than the recovery effect of delayed monetary compensation.

Overview of the Studies

We conduct two studies. Study 1 is a meta-analysis of studies examining the recovery effect of compensation. A meta-analysis is deemed appropriate because it consolidates the findings of extant studies, which, each on their own, comprise a small number of conceivable compensation types and failure types but taken together cover the whole spectrum. Further, the meta-analysis enables us to compare the explanatory power of current versus resource-based compensation and failure type classifications on an accumulated basis. Study 2 is an experiment conducted to enhance internal validity because the meta-analytic integration of independent single studies usually entails a large degree of heterogeneity (Glass, McGaw, and Smith 1981).

Study 1: Meta-Analysis
Method

Literature Search. A computer-aided bibliographic key word search was conducted to identify studies reporting on the recovery effect of compensation on any customer reaction, by using any compensation type for any failure specified in our theoretical framework depicted in Figure 1. Search words were "consumer complaint," "failure," "(service) recovery," "compensation," and "customer satisfaction." Further, we conducted an issue-by-issue search of leading pertinent journals as well as backward and forward reference analyses of extant reviews (e.g., Davidow 2003) and pertinent articles (e.g., Smith, Bolton, and Wagner 1999). The search covered the period 1980 to March 2013 and the following databases: Ebsco Business Source Complete, Science Direct, Emerald Management Xtra, ABI/Inform, PsycINFO, Google Scholar, and the Social Science Citation Index. Further, we directly e-mailed researchers of published studies, which did not report sufficient statistics for effect size calculation. We also obtained unpublished work by searching the Social Science Research Network (SSRN) database and by posting a note at the Electronic List Server for Marketing Academics. We filtered out studies that were duplicates or contained insufficient information to calculate effect size of a single compensation type on customer reactions.

This procedure generated 55 empirical articles. A complete list of studies included is available from the authors upon request. From these studies, a total of 92 effect sizes based on 81 independent samples with a total N of 21,474 were obtained. Of the 92 effect sizes, 96.7% were based on experimental data and 3.3% on survey data. Subjects were students in 51.1% of the cases and nonstudents in 48.9% of the cases; 20.7% of the effect sizes stemmed from publications in top-tier outlets and 79.3% from non-top-tier publications. With regard to the industry settings, effect size distribution was as follows: 26.6% restaurant; 18.5% consumer goods like shoes, books, or electronics; 14.7% hotel; 10.9% air travel; and 29.3% others such as telecommunication or car repair.

Overview of the Meta-Analytical Approach. Single experimental studies usually do not explicitly broach the issue of different compensation and failure types, but a prototypical study compares postfailure customer reactions in a no-compensation group to postrecovery customer reactions in a compensation group (e.g., Bonifield and Cole 2008). In order to examine the Compensation Type × Failure Type contingencies displayed in Figure 1, we proceeded as follows for each individual study. First, we coded compensation type and failure type, based on the description of the experimental design, and we coded the customer reaction based on the dependent variable used. For the respective failure type, we then compared the postfailure reaction to the postrecovery reaction for the respective compensation type. Based on this comparison, we calculated the effect size for the respective combination of compensation type and failure type, as described subsequently.[2]

Coding Procedure. Two independent judges coded the variables displayed in Figure 1 (Eisend 2009): compensation type, failure type, customer reactions, and control variables. For compensation type, we coded compensation as either tangible (e.g., 20% discount) or psychological (e.g., apology), which represents the current classification. Further, we coded compensation as delayed-monetary (e.g., US$175 discount on the next flight), immediate-monetary (e.g., 20% discount on the current bill), new/replaced goods (e.g., new tennis shoes), new/reperformed service (e.g., newly cooked steak), or psychological (e.g., apology), which represents the resource-based classification. In order to reduce the number of blank cells, we collapsed new/replaced goods and new/reperformed service into one category, which can be justified because both categories represent concrete resources with a monetary value.

With regard to failure type, we coded three variables for current classifications: outcome (e.g., overcooked steak) versus process (e.g., undue preference of other customers), monetary (e.g., faulty tires) versus nonmonetary (e.g., delayed flight), and reversible (e.g., incorrect restaurant order) versus irreversible (e.g., unavailable meal). We also coded one resource-based variable: monetary failure (e.g., charging customers for unsubscribed features) versus flawed goods (e.g., tennis shoes falling apart) versus failed service (e.g., overcooked steak) versus lack of attention (e.g., unavailable service personnel).

Further, any construct that measured SAT (e.g., satisfaction with the service recovery), LOY (e.g., repatronage intention), or positive WOM (e.g., recommending the company and its services), as well as rarely used constructs like desire for vengeance, was coded as a customer reaction. We subsumed all customer reactions under one category to ensure a meaningful number of studies for the relationships examined in our meta-analysis.

As control variables, we included failure magnitude, compensation amount, and study characteristics, which may bias the results. Failure magnitude, which is also referred to as failure severity, is the size of the loss entailed in a failure (Smith, Bolton, and Wagner 1999). Compensation amount refers to the scope of remuneration customers receive (Gelbrich and Roschk 2011b). Coders were asked to rate both variables on a 7-point rating scale ranging from 1 = *very low* to 7 = *very high*. The mean value of both coders was then entered into the analysis. Finally, we included study characteristics. These are subject (student or nonstudent), method (survey or experiment), publication outlet (non-top-tier or top-tier), and the number of examined compensation levels (two or more than two/continuous; Gelbrich and Roschk 2011a; Orsingher, Valentini, and de Angelis 2010).

Intercoder reliability (I_R) was assessed using Rust and Cooil's (1994) measure for the categorical variables. I_R ranged from .87 for outcome versus process failure to 1.00 for tangible versus psychological compensation, thus exceeding the required .7 threshold. Inconsistencies were resolved through discussion. For the continuous variables, we used the intraclass correlation coefficient, which was adequate both for failure

magnitude and for compensation amount with values of .87 and .88, respectively (Shrout and Fleiss 1979).

Calculation of Effect Size. We used the correlation coefficient as the effect size metric for the meta-analysis (Eisend 2009). For the prototypical study described earlier, the standardized mean difference between the postrecovery reaction and the postfailure reaction served as input for calculating the correlation coefficient, using the formula provided by Cohen (1988, p. 23, 82). For other studies, correlation coefficients were assessed directly from the studies or calculated or approximated through other statistical data, such as η^2, F ratios, β coefficients, and by means of formulas provided by Glass, McGaw, and Smith (1981, p. 149), Cohen (1988, p. 23), and Peterson and Brown (2005). Single correlation coefficients were adjusted for reliability in order to correct for measurement errors (Geyskens, Steenkamp, and Kumar 1999; Hunter and Schmidt 2004). The mean of these single reliability adjusted correlation coefficients yielded the pooled correlation coefficients r_p. Adjustment for sampling error was made by weighting the single reliability adjusted correlation coefficients by sample size (Hunter, Schmidt, and Jackson 1982). For the purpose of simplicity, we talk about "effect sizes" instead of "single reliability corrected effect sizes, weighted by sample size" in the following.

Course of Analysis. The analysis comprised three steps. First, univariate statistics were provided, including the pooled correlations coefficients for all possible pairs of failure and compensation types (r_p), 95% confidence intervals (Schmidt and Hunter 1999), total sample size, and the number of observations. The pooled correlations were tested for significance by z statistics. The second step covered hypotheses testing. The third step comprised comparisons of resource-based compensation type and failure type classifications with current classifications. In the second and third steps, we included two control variables, failure magnitude and student versus nonstudent subjects, since they turned out to be statistically significant in subsequent model estimations.

Results

Univariate Statistics. Table 1 depicts the pooled correlation coefficient r_p, which is represented by compensation's effect size on customer reactions, and univariate statistics for all pairs of resource-based compensation type and failure type. The overall pooled correlation coefficient is .39, ranging from .11 for Flawed Goods × Delayed Monetary compensation to .70 for Monetary Failure × Immediate Monetary compensation. All coefficients are significant at $p < .05$.

Test of Hypothesis 1. The test of the matching hypothesis only includes the cells encircled by the dotted line in Table 1. Delayed monetary compensation is excluded to rule out confounding effects of remuneration provided in future. We first estimate an analysis of covariance (ANCOVA)

Table 1. Univariate Statistics for Study 1.

Failure type		Delayed-monetary Voucher (88.2%), (store) credit (11.8%)	Immediate-monetary Discount on bill (41.0%), money back (35.9%), cash-refund (23.1%)	New / exchanged goods new / re-performed service Product exchange (21.4%), gift (7.1%), newly cooked meal (21.4%), service correction (e.g., law n repaired; 50.1%)	Psychological Apology (100.0%)	Row total
Monetary failure			.70		.13	.60
Account overcharge (57.1%), additional travel costs (28.6%), other monetary loss claimed (14.3%)	r_p					
	CI		.47 | .94		n.A.	.34 | .86
	n (k)		1779 (6)		395 (1)	2174 (7)
Flawed goods		.11	.34	.69	.26	.27
Inferior product quality (e.g., defective shoes; 50.0%), stained clothing (38.9%), unsatisfactory product performance (11.1%)	r_p					
	CI	.02 | .20	.27 | .42	.25 | 1.14	.15 | .38	.15 | .38
	n (k)	1637 (5)	972 (7)	416 (3)	359 (3)	3384 (18)
Failed service		.16	.36	.68	.38	.40
Inedible meal (e.g., overcooked steak; 19.3%), air travel delay (15.8%), overbooked hotel (8.8%), wrong haircut (7.0%), car not ready (7.0%), delayed delivery (7.0%), incorrect restaurant order (7.0%), others (28.1%).	r_p					
	CI	.10 | .23	.29 | .44	.59 | .78	.30 | .46	.34 | .46
	n (k)	2708 (11)	4402 (22)	2948 (10)	3880 (14)	13937 (57)
Lack of attention		.44	.28	.48	**.44**	.34
Undue preference of others (30.0%), bad attitude to customer needs (30%), revealing personal information (20.0%), staff unavailability (10%), others (10%).	r_p					
	CI	n.A.	.18 | .37	n.A.	.37 | .51	.27 | .41
	n (k)	25 (1)	1198 (4)	65 (1)	690 (4)	1978 (10)
Column total		.14	.42	.68	.36	.39
	r_p					
	CI	.09 | .20	.34 | .50	.58 | .78	.30 | .43	.34 | .44
	n (k)	4370 (17)	8351 (39)	3429 (14)	5324 (22)	21474 (92)

Note. r_p = mean recovery effect of compensation on customer reactions, represented by the sample-weighted reliability adjusted correlation coefficient (= pooled correlation coefficient); CI = 95% confidence interval; k = number of observations; N = total sample size; NA = not available. All correlations are significant at p = .05. In cases where only one effect is observed, significance level was assessed using the individual study's z-value. The average value of fail-safe N is 1,247. In the majority of cases the tolerance level of five times the number of observations plus 10 is met (Rosenthal 1979), except for the four cells where the total sample size is below 400.

Table 2. ANCOVA Results for Hypothesis Testing (η^2 Values) in Study 1.

DV: Single Reliability Corrected Effect Sizes of Compensation on Customer Reactions	Model 1: Matching Hypothesis		Model 2: Intertemporal Choice Hypothesis
	(a) Interaction Effect	(b) Matching Variable	
Resource-based compensation type	.08†	.00 *ns*	.21**
Resource-based failure type	.04*ns*	.04 *ns*	.03*ns*
Resource-based compensation type × failure type	.20*	—	—
Matching variable	—	.15**	—

Note. ANCOVA = analysis of covariance; DV = dependent variable; *ns* = not significant. η^2 values are depicted. All models include the following control variables: failure magnitude and subject (student vs. nonstudent).
$^\dagger p < .10.$ * $p < .05.$ ** $p < .01.$

with the effect sizes of compensation on customer reactions as the dependent variable, the resource-based compensation type classification and the resource-based failure type classification, and their interaction as independent variables. The results are depicted in Table 2, Model 1a. They show a significant Compensation Type × Failure Type interaction ($\eta^2 = .20$, $p = .015$), indicating that the effect size of compensation on customer reactions depends on specific combinations of compensation type and failure type, which provide preliminary support for Hypothesis 1.

For an exact test of Hypothesis 1, we generate a matching variable by coding the off-diagonal elements in the dotted area as "nonmatching resources" versus the boldfaced diagonal elements as "matching resources". This dummy variable then replaces the Compensation Type × Failure Type interaction in another ANCOVA estimation. The results are shown in Table 2, Model 1b. The significant effect of the matching variable ($\eta^2 = .15$, $p = .001$) indicates that the recovery effect of compensation on customer reactions depends on the match of resources. In particular, the results shows that, on average, the recovery effect on the diagonal elements ($r_p = .66$) is stronger than the recovery effect on the off-diagonal elements ($r_p = .35$). Hence, Hypothesis 1 is supported.

Test of Hypothesis 2. The test of the intertemporal choice hypothesis only includes the gray-shaded cells in Table 1, retaining delayed versus immediate monetary compensation and excluding all nonmonetary compensation types. The results of the respective ANCOVA are depicted in Table 2, Model 2. The effect size of compensation on customer reactions varies significantly by compensation type ($\eta^2 = .21$, $p = .001$). For the gray-shaded cells, immediate monetary compensation exhibits a stronger effect ($r_p = .35$) than delayed monetary compensation ($r_p = .14$). Hence, Hypothesis 2 is supported.

Comparison of Resource-Based Versus Current Classifications. By crossing the two different compensation type classifications with the four different failure type classifications in Figure 1, eight models are conceivable. In order to compare these models, we conduct eight ANCOVAs, using the respective compensation type classification, failure type classification, and their interaction as independent variables and the effect size of

compensation on customer reactions as the dependent variable. The first four ANCOVAs are displayed in the upper part of Table 3. They use the current compensation type classification, namely, tangible versus psychological, and one of the four different failure type classifications, these being outcome versus process, monetary versus nonmonetary, reversible versus irreversible, and resource based. The second four ANCOVAs are displayed in the lower part of Table 3. They use the resource-based compensation type classification, namely, delayed monetary compensation versus immediate monetary compensation versus new/exchanged goods and new/reperformed service versus psychological compensation, and one of the four different failure type classifications. Models are compared based on Akaike's information criterion (AIC), which is a goodness-of-fit measure for nonnested model comparison, with lower values indicating a better fit (Akaike 1987).

Is the resource-based compensation type classification better than the current compensation type classification? To answer this question, we compare the four models using the resource-based compensation type classification, depicted in the lower part of Table 3, to the four models using the current compensation type classification, depicted in the upper part of Table 3. The models in the lower part show lower AIC values, that is, a better fit, than those in the upper part, regardless of using outcome versus process (494 < 518), monetary versus nonmonetary (490 < 516), reversible versus irreversible (491 < 517), or resource-based failure type classifications (474 < 506).

Is the resource-based failure type classification better than the current failure type classification? To answer this question, we compare the models using the resource-based failure type classification, which are shown in the last column of Table 3, to the three models using the current failure type classifications, which are shown in the first three columns of Table 3. This comparison is done for the upper and lower part of Table 3. In the upper part, the model using the resource-based failure type classification has lower AIC values, that is, a better fit, than the current classification models (506 < 518, 516, 517); the same applies to the lower part of Table 3 (474 < 494, 490, 491).

Finally, the model that uses both the resource-based compensation and failure type classification, which is presented in the last column of Table 3's lower part, performs best. It has

Table 3. Model Comparisons for Study 1.

DV: Single Reliability Corrected Effect Sizes of Compensation on Customer Reactions	Failure Type Classification			
	Outcome Versus Process	Monetary Versus Nonmonetary	Reversible Versus Irreversible	Resource-Based Failure Type
Models using the current compensation type classification				
Tangible versus psychological compensation	.00 *ns*	.01 *ns*	.00 *ns*	.02 *ns*
Failure type	.00 *ns*	.01 *ns*	.03 *ns*	.06 ns
Tangible versus Psychological Compensation × Failure Type	.02 *ns*	.05 *	.00 *ns*	.11 *
AIC	518	516	517	506
Models using the resource-based compensation type classification				
Resource-based compensation type	.35 ***	.37 ***	.35 ***	.08 *ns*
Failure type	.00 *ns*	.00 *ns*	.04 †	.03 *ns*
Resource-based compensation type × failure type	.02 *ns*	.06 *ns*	.01 *ns*	.19 *
AIC	494	490	491	474

Note. AIC = Akaike's information criterion; DV = dependent varaiable; *ns* = not significant. η^2 values are depicted. All models include the following control variables: failure magnitude and subject (student vs. nonstudent).
†*p* < .10. **p* < .05. ***p* < .01. ****p* < .001.

the lowest AIC value of 474. In summary, these results support the postulate that resource-based classifications outperform current classifications.

Discussion of Study 1

Study 1 is a cross-sectional meta-analysis, finding that the recovery effect of compensation on customer reactions is largest when compensation type and failure type represent the same resource and that the recovery effect of immediate monetary compensation on customer reactions is stronger than that of delayed monetary compensation. Further, the resource-based classifications of compensation type and failure type better explain the recovery effect of compensation than traditional classifications.

However, one limitation of Study 1 is the small number of studies, given the number of conceivable Failure Type × Compensation Type combinations. This limitation leads to overlapping confidence intervals, even for notably diverging mean effect sizes of .34 versus .69 for Flawed Goods × Immediate Monetary compensation versus Flawed Goods × New/Exchanged Goods/Reperformed Service. Further, we were forced to consolidate. In particular, we had to leave cells blank, for example, Monetary Failure × Delayed Monetary compensation, and we had to subsume different customer reactions and new/exchanged goods–new/reperformed service under one category each.

The second limitation is a decrease in internal validity, which naturally occurs when integrating heterogeneous studies with different variable operationalizations or cross-sectional data. Hence, we conduct Study 2, which is an experiment that entails a full factorial design of our Resource-Based Compensation Type × Failure Type classifications and controls for heterogeneity. Importantly, we distinguish between the three customer reactions: customer SAT, LOY, and positive WOM.

Study 2: Experiment

Method

Experimental Design. We used a 4 × 5 between-subject, scenario-based experiment, crossing the four failure types monetary failure, flawed goods, failed service, and lack of attention with the five compensation types delayed monetary compensation, immediate monetary compensation, new/exchanged goods, new/reperformed service, and psychological compensation. The core scenario describes a customer who has the stained passenger seat of his car cleaned at an auto repair service for the price of 40€.[3]

Service failure type was manipulated in the following way. In the monetary failure condition, the customer discovers that she or he is charged the cleaning of two seats instead of one at 40€ each. In the flawed goods condition, the car mats in the front, which the customer had recently bought for 40€, are now oil stained and soaked with water. In the failed service condition, the stains on the passenger seat are still clearly visible, although the customer has just paid 40€ for the cleaning. In the lack of attention condition, the service person receives a mobile phone call, returns to his office without an explanation, and returns 15 min later, after having had a coffee with his colleague. Compensation type was manipulated by describing the customer as complaining to the service person and receiving one of the following compensation forms. Delayed monetary compensation is given in the form of a voucher of 40€ for a future service from the provider and immediate monetary compensation as a refund of 40€. New/exchanged goods come as two new car mats or an exchange of two car mats worth 40€ altogether.[4] New/reperformed service is operationalized as cleaning the driver's seat or the passenger's seat again, worth 40€.[5] In the psychological compensation condition, the customer receives a sincere apology.

In order to rule out confounding effects, we conducted a series of pretests, which ensured that failure magnitudes and compensation magnitudes were approximately the same across

the different failure types and compensation types, respectively. Then, participants were randomly assigned to one of the 20 scenarios and were asked to answer a series of questions.

Data Collection and Sample. Data were collected online, from a national sample of adult German consumers based on quotas for gender and age. We received 493 usable questionnaires. Cell sizes ranged from 23 to 25 persons across the 20 scenarios. In this sample, 51.7% of the respondents were women, and their ages ranged from 18 to 90 years ($M = 47.3$). A chi-square test was nonsignificant, indicating that the distributions of gender and age did not significantly differ from the German adult population, $\chi^2(2) = 0.423$, $p = .810$.

Measures. We measured the two manipulation checks by one categorical variable each. For failure type, respondents indicated whether the failure in the scenario describes a wrong billing, a damage of a car accessory, a failed cleaning of the seat, or an inattentive treatment. For compensation type, respondents indicated whether compensation comes in the form of a voucher worth 40€ for a future service from the provider, a refund of 40€, two new/exchanged car mats worth 40€ altogether, a/another cleaning of the driver's/passenger seat worth 40 €, or a sincere apology.

We captured three customer reactions, customer SAT, LOY, and positive WOM, both after the failure occurred and after compensation was provided measuring postfailure and postrecovery customer reactions. Customer SAT was captured by 4 items adapted from Hess, Ganesan, and Klein (2003): *I am pleased with the service*, *The situation is satisfying*, *I am happy*, and *My expectations have been fulfilled*, with Cronbach's α = .85 for postfailure and .96 for postrecovery. LOY and positive WOM were captured by 1 item each adapted from Maxham and Netemeyer (2002a): *I would use services from this auto repair service provider again* and *I would tell other positive things about this auto repair service provider.*

Control variables concerning the failure were failure magnitude and failure attributions. The measure of failure magnitude was adapted from Hess, Ganesan, and Klein (2003). We asked the respondents to rate 3 items: *For me as a customer, the failure is ... significant, major, severe* (Cronbach's α = .93). Failure attributions were captured by two items adapted from Folkes, Koletsky, and Graham (1987): *The auto repair service provider had control over the reason for the failure* and *There are actions the auto repair service provider could take but did not to prevent the failure* (Cronbach's α = .84). Other control variables related to compensation as well as personal attitudes and experience. Compensation magnitude was measured by 4 items developed for the purpose of this study, namely, *The reaction of the service employee is ... significant, major, noteworthy, considerable* (Cronbach's α = .89). Attitude toward complaining was captured by 3 reverse-coded items adapted from Blodgett, Hill, and Tax (1997): *I am usually reluctant to complain*, *In general, I avoid filing complaints*, and *I am less likely to complain than most people I know* (Cronbach's α = .89). Finally, the respondents indicated their car ownership (yes

vs. no), and their experience with auto repair services. Experience was measured on two scales anchoring at 1 = *few experience / infrequent use* and 7 = *much experience/frequent use* (Cronbach's α = .93). Unless otherwise stated, we used 7-point Likert-type scales ranging from 1 = *do not agree at all* to 7 = *fully agree*.

Manipulation Checks. We conducted a chi-square test crossing the manipulation of failure type with the respective manipulation check measure. The chi-square test was significant, $\chi^2(9) = 1,394.656$, $p < .001$, with 97.8% correctly perceived failure types. Further, the respondents considered the failure description to be realistic ($M = 5.52$ on a 7-point scale). Another chi-square test, which crossed the manipulation of compensation type with the respective manipulation check measure, was also significant, $\chi^2(16) = 1,884.460$, $p < .001$, with 98.2% correctly perceived compensation types. Finally, respondents perceived the whole complaint situation as realistic ($M = 5.48$). Hence, the manipulations were successful.

Results

Descriptive Statistics and Course of Analysis. Table 4 shows the mean recovery effect of compensation on the three customer reactions for the 20 different combinations of failure type and compensation type. In accordance with Figure 1, the recovery effect of compensation is calculated as the difference between the respective postrecovery reaction and postfailure reaction and serves as the dependent variable in all subsequent analyses. Hypothesis testing, which is displayed in Table 5, is conducted in the same fashion as in Study 1. Again, we include control variables that turn out to be statistically significant in subsequent model estimations. These are experiences with car repair, car ownership, attitude toward complaining, attributions, failure magnitude, and compensation magnitude.

Test of Hypothesis 1. In order to test the matching hypothesis, we only include the cells encircled by the dotted line in Table 4. We estimate three ANCOVAs, each of which uses the recovery effect of compensation on SAT, LOY, or WOM as the dependent variable. The independent variables are compensation type classification, failure type classification, and their interaction. The results are shown in Table 5, Model 1a. The significant Compensation Type × Failure Type interactions for SAT ($\eta^2 = .11$, $p < .001$), LOY ($\eta^2 = .06$, $p = .004$), and WOM ($\eta^2 = .08$, $p < .001$) indicate that the recovery effect of compensation on customer reactions depends on specific combinations of compensation type and failure type, which provides preliminary support for Hypothesis 1.

For an exact test of Hypothesis 1, we again replace the Compensation Type × Failure Type interaction with a matching variable, comprising the off-diagonal versus the diagonal elements in the dotted area of Table 4. We then estimate individual ANCOVAs for each customer reaction. The results are displayed in Table 5, Model 1b. The significant effects of the

Table 4. Satisfaction, Loyalty, and Positive WOM Cell Means by Experimental Conditions for Study 2.

				Compensation type			
Failure type		Delayed-monetary	Immediate-monetary	New / exchanged goods	New / re-performed service	Psychological	Row total
Monetary failure	Satisfaction	.80	**3.24**	1.53	1.55	1.23	1.67
	Loyalty	.32	**2.21**	.63	1.54	.58	1.06
	Positive WOM	.60	**2.63**	1.41	1.33	1.05	1.40
Flawed goods	Satisfaction	1.65	2.24	**3.05**	1.68	1.28	1.98
	Loyalty	1.54	1.07	**1.76**	1.09	.78	1.25
	Positive WOM	1.39	1.28	**1.85**	1.47	1.13	1.42
Failed service	Satisfaction	1.62	2.62	1.14	**3.30**	1.82	2.10
	Loyalty	1.65	1.27	1.02	**2.28**	1.23	1.49
	Positive WOM	1.63	2.25	1.11	**2.65**	1.45	1.82
Lack of attention	Satisfaction	1.97	2.26	2.05	1.97	**1.75**	2.00
	Loyalty	1.69	.63	1.18	1.68	**1.12**	1.26
	Positive WOM	1.18	.71	1.23	1.51	**1.30**	1.19
Column total	Satisfaction	1.51	2.59	1.94	2.12	1.52	1.92 (1.55)
	Loyalty	1.30	1.29	1.15	1.65	.93	1.26 (1.62)
	Positive WOM	1.20	1.72	1.40	1.74	1.23	1.46 (1.44)

Note. ANCOVA = analysis of covariance; WOM = word of mouth. The table displays the mean recovery effect of compensation on customer reaction. This effect is an estimated value resulting from an ANCOVA that uses the difference between the respective postrecovery reaction and postfailure reaction as the dependent variable, failure type, and compensation type and their interaction as independent variables. The following control variables are included experiences with car repair, car ownership, attitude toward complaining, attributions, failure magnitude, and compensation magnitude. Standard deviations are depicted in parenthesis.

Table 5. ANCOVA Results for Hypothesis Testing (η^2 Values) in Study 2.

DV: Postrecovery Reaction–Postfailure Reaction		Model 1: Matching Hypothesis		Model 2: Intertemporal Choice Hypothesis
		(a) Interaction Effect	(b) Matching Variable	
Compensation type	Satisfaction	.06 ***	.06 ***	.10 ***
	Loyalty	.03 *	.03 *	.00 ns
	Positive WOM	.02 *	.02 *	.03 *
Failure type	Satisfaction	.01 ns	.00 ns	.00 ns
	Loyalty	.00 ns	.00 ns	.00 ns
	Positive WOM	.03 *	.02 *	.05 *
Compensation Type × Failure Type	Satisfaction	.11 ***	—	—
	Loyalty	.06 **	—	—
	Positive WOM	.08 ***	—	—
Matching variable	Satisfaction	—	.08 ***	—
	Loyalty	—	.04 ***	—
	Positive WOM	—	.05 ***	—

Note. ANCOVA = analysis of covariance; DV = dependent variable; WOM = word of mouth. η^2 values are depicted. All models include the following control variables: experiences with car repair, car ownership, attitude toward complaining, attributions, failure magnitude, and compensation magnitude.
*$p < .05$. **$p < .01$. ***$p < .001$.

matching variables for SAT ($\eta^2 = .08$, $p < .001$), LOY ($\eta^2 = .04$, $p < .001$), and WOM ($\eta^2 = .05$, $p < .001$) indicate that the recovery effect of compensation on these customer reactions depends on the match of resources. In particular, the results show that, on average, the recovery effect on the diagonal elements (SAT = 2.85, LOY = 1.86, and WOM = 2.13) is stronger than the recovery effect on the off-diagonal elements (SAT = 1.80, LOY = 1.08, and WOM = 1.34). Hence, Hypothesis 1 is supported.

Test of Hypothesis 2. The test of the intertemporal choice hypothesis only includes the gray-shaded cells in Table 4. The results of the ANCOVAs are depicted in Table 5, Model 2. They show that the recovery effect of compensation on customer reactions significantly varies by compensation type for SAT ($\eta^2 = .10$, $p < .001$) and WOM ($\eta^2 = .03$, $p = .020$). Immediate monetary compensation exhibits a stronger effect than delayed monetary compensation, with values of 2.59 > 1.51 for SAT and 1.72 > 1.20 for WOM. Yet, the result is nonsignificant for LOY ($\eta^2 = .00$, $p = .762$), with equally strong effects of 1.29 and 1.30 for immediate and delayed compensation. Hence, Hypothesis 2 is supported for SAT and positive WOM only.

Although not specifically hypothesized, we also examine whether the recovery effect of immediate versus delayed monetary compensation depends on failure type. For this purpose, we add the interaction term of Compensation Type × Failure Type into Model 2. The results of the respective ANCOVAs show significant interactions for SAT ($\eta^2 = .06$, $p = .009$), LOY ($\eta^2 = .11$, $p < .001$), and WOM ($\eta^2 = .09$, $p = .001$).

Discussion of Study 2

The results of Study 2 largely support the results of Study 1, yet with a full-factorial design that controls for potential confounds

and distinguishes between the three customer reactions, namely, SAT, LOY, and positive WOM. In particular, Study 2 replicates the support of the matching hypothesis and the intertemporal choice hypothesis. Yet, the intertemporal choice hypothesis only holds true for customer SAT and positive WOM. For customer LOY, immediate and delayed compensation are equally effective. Further, although not hypothesized, Study 2 reveals that the recovery effect of delayed versus immediate compensation on the customer reactions varies with failure type.

Overall Discussion

Theoretical Issues

Drawing on resource exchange theory (Foa and Foa 1974, 1976), the results of both a cross-sectional meta-analysis (Study 1) and a controlled experiment (Study 2) enhance extant knowledge of the recovery effect of compensation in three ways. First, we shed light on the adequate compensation type—both in kind and in point in time—for a particular failure type. Second, we show that resource-based classifications of compensation type and failure type outperform current classifications in predicting the recovery effect of compensation. Third, we offer new explanations of the fluctuating effect size of compensation on customer reactions observed in prior studies.

Type of Compensation. The cells encircled by the dotted line in Table 1 (Study 1) and in Table 4 (Study 2) illustrate the hypothesized match of resources. The recovery effect of compensation is greater for the boldfaced diagonal elements than for the off-diagonal elements. As the results in both studies show the same pattern, we discuss them for the recovery effect on customer SAT in Study 2, which is represented by the first number in each cell of Table 4.

We first answer the question whether the different compensation types are more or less adequate to rectify a certain failure type. Conducting horizontal comparisons, we see that a monetary failure is best rectified by immediate monetary compensation; the recovery effects of new/exchanged goods, new reperformed service, and psychological compensation are lower (3.24 > 1.53, 1.55, 1.23). Flawed goods are best rectified by an exchange; the effects of monetary compensation, new/reperformed service, and psychological compensation are lower (3.05 > 2.24, 1.68, 1.28). A failed service is best rectified by reperformance; the effects of monetary compensation, new/exchanged goods, and psychological compensation are lower (3.30 > 2.62, 1.14, 1.82).

However, lack of attention is a special case. In general, the matching resource of psychological compensation has a significant recovery effect of 1.75, but monetary compensation, new/exchanged goods, and new/reperformed service are about equally effective, with values of 2.26, 2.05, and 1.97, respectively. A reason for this unexpected finding may lie in the nature of this failure, which represents the loss of symbolic resources, namely, "love" and "status" (Foa and Foa 1974). Monetary compensation and a new product or service may, as well as psychological compensation, entail a symbolic meaning. They show that the company acknowledges the failure and makes an effort to rectify the customer's loss (Fehr and Gelfand 2010). In addition, they encompass a tangible value. Hence, these alternative compensation types may rectify lack of attention to the same extent as an apology.

We next discuss the results with regard to the question whether a certain compensation type is more or less adequate to rectify different failure types. Conducting vertical comparisons, we notice that immediate monetary compensation has the greatest recovery effect in case of a monetary failure; its recovery effect for flawed goods, failed service, and lack of attention is lower (3.24 > 2.24, 2.62, 2.26). New/exchanged goods have the greatest recovery effect in case of flawed goods; their effects for monetary failure, failed service, and lack of attention are lower (3.05 > 1.53, 1.14, 2.05). New/reperformed service has the greatest recovery effect in case of failed service; its effect for monetary failure, flawed goods, and lack of attention is lower (3.30 > 1.55, 1.68, 1.97).

Again, psychological compensation is a special case. In general, it has a recovery effect for lack of attention as the matching failure (1.75). Yet, this effect is only slightly stronger than that for monetary failure (1.23) and flawed goods (1.28) and about equally strong as for failed services (1.82). These unexpected results can be explained by the fact that all failures, not only lack of attention, may imply a social denigration, that is, a loss of "love" and "status." For example, customers who recognize that they were charged too much, that a product is inferior, or that the steak is overcooked, may perceive that the company is ignorant of their needs. Psychological compensation, which acknowledges the complainant's distress (Davidow 2003), directly aims to rectify such feelings of denigration.

Point in Time of Compensation. Study 1 and Study 2 provide general support for the intertemporal choice hypothesis: Immediate monetary compensation is more effective than delayed monetary compensation. The reason is that consumers are subject to a present bias known from research on intertemporal evaluation. People seek immediate rewards (Pyone and Isen 2011) and avoid the uncertainty entailed in the future (Chan and Mukhopadhyay 2010). However, Study 2 reveals that this only holds true for customer SAT and positive WOM, not for LOY. Along with the significant interactions between immediate versus delayed compensation and the three customer reactions, this finding suggests that the intertemporal choice hypothesis depends on the failure type.

In case of a monetary failure, immediate monetary compensation is clearly superior to delayed monetary compensation, as indicated by the consistently greater recovery effects for SAT (3.24 > 0.80), LOY (2.21 > 0.32), and positive WOM (2.63 > 0.60). For flawed goods and failed service, immediate compensation outperforms delayed compensation in terms of SAT (flawed goods: 2.24 > 1.65, failed service: 2.62 > 1.62), but the effect is reversed for LOY (flawed goods: 1.07 < 1.54, failed service: 1.27 < 1.65). The results are mixed for WOM (flawed goods: 1.28 < 1.39, failed service: 2.25 > 1.63). For lack of attention, immediate monetary compensation is inferior to delayed compensation: It leads to a similar increase in SAT (2.26 vs. 1.97), yet to lower increases in LOY (0.63 < 1.69) and WOM (0.71 < 1.18).

Obviously, the present bias remains valid for monetary failures. If customers have to wait for having a monetary loss rectified, they will remain upset and will not recommend the company to others; the voucher offered instead does not tempt them to reuse the company, although vouchers can only be redeemed through a repurchase. In case of flawed goods or a failed service, the present bias mainly affects SAT judgments. Customers feel bad about the delay of rewards comprised in a voucher. Yet, they may be tempted to repurchase from the company in order to redeem the voucher even though they would prefer to receive immediate compensation (Knox and van Oest 2012). For lack of attention, the present bias does not occur. This may be because delayed compensation in the form of a voucher represents a symbolic meaning that goes beyond a mere monetary value. It may be perceived as a present, which acknowledges the customer's denigration and mitigates the negative repercussions of the failure.

Resource-Based Classification. Current classifications of compensation type and failure type are helpful to consolidate the broad spectrum of conceivable failures and compensation types. However, Study 1 shows that the resource-based classifications explain the recovery effect of compensation for a particular failure better than current classifications. This conclusion can be drawn from the model comparisons depicted in Table 3. Models including the resource-based compensation type classifications displayed in the lower part of Table 3 have a better goodness of fit than models including the current compensation type classification displayed in the upper part of Table 3. Further, models including the resource-based failure type classifications displayed in the last column of Table 3 have a better

goodness of fit than models including the current failure type classifications displayed in the first three columns of Table 3. Finally, the best model is the one that combines the resource-based classification for both compensation and failure type shown in the last column in the lower part of Table 3.

Fluctuating Effect Sizes. Although the effect of compensation on customer reactions is well established in prior research (Davidow 2003), effect size fluctuates remarkably (e.g., Garrett 1999 vs. Grewal, Roggeveen, and Tsiros 2008 vs. Hess, Ganesan, and Klein 2003). Prior research already finds context variables such as culture, which may account for this fluctuation (Mattila and Patterson 2004). Our research results presented in Table 4 show that failure type and compensation type also have a large impact on the recovery effect of compensation. The recovery effect is very high when companies rectify a monetary failure by immediate monetary compensation (SAT = 3.24, LOY = 2.21, and WOM = 2.63), whereas it is almost negligible when they seek to rectify a monetary failure by delayed monetary compensation (SAT = 0.80, LOY = 0.32, WOM = 0.60).

In addition to their enormous impact, failure type and compensation type are important for recovery, because previous research shows that increasing compensation amount yields more positive customer reactions, though at diminishing returns for very high compensation (Gelbrich and Roschk 2011b). Our research shows that matching compensation type with failure type and providing immediate rather than delayed monetary compensation may have the same effect but does not incur additional costs.

Finally, our research suggests that effect sizes may fluctuate in previous studies because they draw on different customer reactions. In general, the results of Table 4 indicate that compensation is most appropriate to restore SAT after a failure occurred; the effects on behavioral intentions are smaller. Yet, LOY and positive WOM are still largely increased in the matching cells with recovery effects of around 2.00 and exhibiting the same pattern as SAT.

Managerial Implications

Our recommendations for complaint management evolve from our two major findings. The first recommendation is that organizations should match compensation type with failure type, which comprises two steps. In a first step, frontline employees need to identify the resource-based failure type. The prototypical failure types presented in the first column of Table 1 may facilitate this task. Monetary failures represent a mere financial loss for customers. Flawed goods represent a loss that derogates the product benefit, while a failed service occurs when customers cannot make full or any use of the aspired service benefit. Lack of attention represents a psychological loss, which denigrates customers and threatens their status and self-esteem. In the second step, companies determine the compensation type that repays in kind what customers lost, by drawing on the prototypical compensation types presented in the first row of Table 1. A monetary failure is best rectified by reimbursing the money. Flawed goods should be exchanged, while a failed service should be reperformed. Finally, lack of attention may be appropriately rectified by apologizing for the failure to reestablish self-esteem, but the other compensation forms serve this purpose as well. However, as none of them achieves very large recovery effects for lack of attention, companies may need to combine different compensation forms. For example, they can provide an apology and a new service for free.

The second recommendation evolves from the finding that the recovery effect of immediate monetary compensation is generally stronger than the effect of delayed monetary compensation, but this effect also depends on failure type. Monetary failures should generally be rectified immediately, by means of a discount on the current bill or a cash refund. In the case of flawed goods or failed service, companies that are unwilling or unable to provide an exchange or a reperformance may choose the following. They can either maximize customer SAT through immediate monetary compensation or maximize LOY through delayed compensation. However, the latter solution should be used carefully because it creates lock-in effects and thus spurious LOY. Finally, for lack of attention, delayed monetary compensation may be an alternative to immediate monetary compensation because it induces greater LOY.

Limitations and Future Research

The limitations of our studies provide opportunities for future research into three directions. First, the failure type classification may be refined and extended. For instance, exploratory research suggests that lack of attention comes in different forms. Service employees may neglect customer needs, which is reflected in statements like "On a hot day, the waitress refused to move me from a window table because there was nothing else free in her section" or they may behave in a dismissive way, which is reflected in statements like "The waiter at this expensive restaurant treated us like dirt because we were only high school kids" (Bitner, Booms, and Tetreault 1990, p. 77f.). Further, exploratory research reports failures related to lack of information, such as customers who were misled by wrong sales information (Kelley, Hoffman, and Davis 1993). Since "information" represents a distinct resource category in resource exchange theory (Foa and Foa 1974), lack of information may be treated as another failure type.

Second, the compensation type classification may be refined and extended. In particular, one may find different kinds of psychological compensation that account for various forms of lack of attention. Further, future research could treat the explanation for a failure, which is described as informational support in the recovery literature (Gelbrich 2010), as a compensation type, and examine whether it makes up for lack of information better than other compensation types.

who supported the meta-analysis by providing valuable additional statistics on their studies.

Notes

1. Foa and Foa (1974) originally present ''information,'' that is, advice and opinions, as a sixth resource class. We drop this category because recovery research treats information as a separate recovery effort that relates to explaining a problem (Davidow 2003; Mattila and Patterson 2004). Further, though lack of information, such as wrong instruction to product usage, is reported in critical incident research (Kelley, Hoffman, and Davis 1993), no recovery study examines such a failure.

2. Given that we used the effect size of compensation as the dependent variable, we melted down "Compensation Type × Failure Type" to a two-way interaction on effect size, whereas it actually represents a "Compensation (yes vs. no) × Compensation Type × Failure Type" three-way interaction on customer reactions.

3. In the case of the monetary failure, the price charged is 80€.

4. We had to adapt this compensation type to the failure type. For damaged car mats in the flawed goods condition, the service person offers an exchange of the mats. For all other failures, he offers two new car mats, along with the information that the mats look nagged for justifying the new mats.

5. Again, we adapted this compensation type to the respective failure type. For a failed cleaning of the passenger seat in the failed service condition, the provider offers another cleaning of the passenger seat. For all other failures, he offers the cleaning of the driver's seat, along with the information that the driver's seat is dirty as well for justifying the cleaning of the driver's seat.

References

Adams, Stacy J. (1965), "Inequity in Social Exchange," in *Advances in Experimental Social Psychology*, Vol. 2, Leonard, Berkowitz, ed. New York, NY: Academic Press, 267-299.

Akaike, Hirotugu (1987), "Factor Analysis and AIC," *Psychometrika*, 52 (3), 317-332.

Bitner, Mary J., Bernard H. Booms, and Mary S. Tetreault (1990), "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents," *Journal of Marketing*, 54 (January), 71-84.

Blodgett, Jeffrey G., Donna J. Hill, and Stephen S. Tax (1997), "The Effects of Distributive, Procedural, and Interactional Justice on Postcomplaint Behavior," *Journal of Retailing*, 73 (2), 185-210.

Bonifield, Carolyn and Catherine A. Cole (2008), "Better Him than Me: Social Comparison Theory and Service Recovery," *Journal of the Academy of Marketing Science*, 36 (Winter), 565-577.

Boshoff, Christo (1997), "An Experimental Study of Service Recovery Options," *International Journal of Service Industry Management*, 8 (2), 110-130.

Brinberg, David and Ronald Wood (1983), "A Resource Exchange Theory Analysis of Consumer Behavior," *Journal of Consumer Research*, 10 (3), 330-338.

Chan, Elaine and Anirban Mukhopadhyay (2010), "When Choosing Makes a Good Thing Better: Temporal Variations in the Valuation of Hedonic Consumption," *Journal of Marketing Research*, 47 (June), 497-507.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cropanzano, Russel, Zinta S. Byrne, D. R. Bobocel, and Deborah E. Rupp (2001), "Moral Virtues, Fairness Heuristics, Social Entities, and Other Denizens of Organizational Justice," *Journal of Vocational Behavior*, 58 (2001), 164-209.

Davidow, Moshe (2003), "Organizational Responses to Customer Complaints: What Works and what Doesn't," *Journal of Service Research*, 5 (February), 225-250.

Donnenwerth, Gregory V. and Uriel G. Foa (1974), "Effect of Resource Class on Retaliation to Injustice in Interpersonal Exchange," *Journal of Personality and Social Psychology*, 29 (6), 785-793.

Eisend, Martin (2009), "A Meta-Analysis of Humor in Advertising," *Journal of the Academy of Marketing Science*, 37 (Summer), 191-203.

Estelami, Hooman (2000), "Competitive and Procedural Determinants of Delight and Disappointment in Consumer Complaint Outcomes," *Journal of Service Research*, 2 (February), 285-300.

Fehr, Ryan and Michele J. Gelfand (2010), "When Apologies Work: How matching Apology Components to Victims' Self-construals Facilitates Forgiveness," *Organizational Behavior and Human Decision Processes*, 113 (1), 37-50.

Foa, Edna B. and Uriel G. Foa (1976), "Ressource Theory of Social Exchange," in *Contemporary Topics in Social Psychology*, John W. Thibaut, Janet T. Spence, and Robert C. Carson, eds. Morristown, NJ: General Learning Press, 99-134.

Foa, Uriel G. and Edna B. Foa (1974), *Societal Structures of the Mind*. Springfield, IL: Charles C, Thomas.

Folkes, Valerie S., Susan Koletsky, and John L. Graham (1987), "A Field Study of Causal Inferences and Consumer Reaction: The View from the Airport," *Journal of Consumer Research*, 13 (4), 534-539.

Garrett, Dennis E. (1999), "The Effectiveness of Compensation Given to Complaining Consumers: Is More Better?" *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 12, 26-34.

Gelbrich, Katja (2010), "Anger, Frustration, and Helplessness After Service Failure: Coping Strategies and Effective Informational Support," *Journal of the Academy of Marketing Science*, 38 (2010), 567-585.

Gelbrich, Katja and Holger Roschk (2011a), "A Meta-Analysis of Organizational Complaint Handling and Customer Responses," *Journal of Service Research*, 14 (February), 24-43.

Gelbrich, Katja and Holger Roschk (2011b), "Do Complainants Appreciate Overcompensation? A Meta-Analysis on the Effect of Simple Compensation Vs. Overcompensation on Post-Complaint Satisfaction," *Marketing Letters*, 22 (1), 31-47.

Geyskens, Inge, Jan-Benedict E. M. Steenkamp, and Nirmalya Kumar (1999), "A Meta-Analysis of Satisfaction in Marketing Channel Relationships," *Journal of Marketing Research*, 36 (May), 223-238.

Gilly, Mary C. and Betsy D. Gelb (1982), "Post-Purchase Consumer Processes and the Complaining Consumer," *Journal of Consumer Research*, 9 (December), 323-328.

Glass, Gene V., Berry McGaw, and Mary L. Smith (1981), *Meta-Analysis in Social Research*. Newbury Park, CA: Sage.

Gouldner, Alvin W. (1960), "The Norm of Reciprocity: A Preliminary Statement," *American Sociological Review*, 25 (2), 161-178.

Grewal, Dhruv, Anne L. Roggeveen, and Michael Tsiros (2008), "The Effect of Compensation on Repurchase Intentions in Service Recovery," *Journal of Retailing*, 84 (4), 424-434.

Hess, Ronald L., Shankar Ganesan, and Noreen M. Klein (2003), "Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction," *Journal of the Academy of Marketing Science*, 31 (Spring), 127-145.

Hunter, John E. and Frank L. Schmidt (2004), *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Thousand Oaks, CA: Sage.

Hunter, John E. and Gregg G. Jackson (1982), *Meta-Analysis: Cumulating Research Findings Across Studies*. Beverly Hills, CA: Sage.

Kelley, Scott W., K. D. Hoffman, and Mark A. Davis (1993), "A Typology of Retail Failures and Recoveries," *Journal of Retailing*, 69 (4), 429-452.

Kim, Namin and Francis M. Ulgado (2012), "The Effect of On-the-spot versus Delayed Compensation: The Moderating Role of Failure Severity," *Journal of Services Marketing*, 26 (3), 158-167.

Knox, George and Rutger van Oest (2012), "Customer Complaints and Lifetime Value," working paper (accessed August 30, 2013), [available at: http://ssrn.com/abstract=1427265.]

Lee, Eun-Jung and JungKun Park (2010), "Service Failures in Online Double Deviation Scenarios: Justice Theory Approach," *Managing Service Quality*, 20 (1), 46-69.

Liao, Hui (2007), "Do it Right this Time: The Role of Employee Service Recovery Performance in Customer-Perceived Justice and Customer Loyalty After Service Failures," *Journal of Applied Psychology*, 92 (2), 475-489.

Mattila, Anna S. and Paul G. Patterson (2004), "Service Recovery and Fairness Perceptions in Collectivist and Individualist Contexts," *Journal of Service Research*, 6 (May), 336-346.

Maxham III, James G., and Richard G. Netemeyer (2002a), "Modeling Customer Perceptions of Complaint Handling Over Time: The Effects of Perceived Justice on Satisfaction and Intent," *Journal of Retailing*, 78 (4), 239-252.

Maxham III, James G., and Richard G. Netemeyer (2002b), "A Longitudinal Study of Complaining Custorners' Evaluations of Multiple Service Failures and Recovery Efforts," *Journal of Marketing*, 66 (October), 57-71.

Noone, Breffni M. and Chung H. Lee (2011), "Hotel Overbooking: The Effect of Overcompensation on Customers' Reactions to Denied Service," *Journal of Hospitality & Tourism Research*, 35 (3), 334-357.

Orsingher, Chiara, Sara Valentini, and Matteo de Angelis (2010), "A Meta-Analysis of Satisfaction with Complaint Handling in Services," *Journal of the Academy of Marketing Science*, 38 (Spring), 169-186.

Peterson, Robert A. and Steven P. Brown (2005), "On the use of Beta Coefficients in Meta-Analysis," *Journal of Applied Psychology*, 90 (1), 175-181.

Pyone, Jin S. and Alice M. Isen (2011), "Positive Affect, Intertemporal Choice, and Levels of Thinking: Increasing Consumers' Willingness to Wait," *Journal of Marketing Research*, 48 (June), 532-543.

Roehm, Michelle L. and Harper A. Roehm (2011), "The Influence of Redemption Time Frame on Responses to Incentives," *Journal of the Academy of Marketing Science*, 39 (3), 363-375.

Roschk, Holger and Susanne Kaiser (2013), "The Nature of an Apology: An Experimental Study on how to Apologize after a Service Failure," *Marketing Letters*, 24 (3), 293-309.

Rosenthal, Robert (1979), "The 'File-Drawer Problem' and Tolerance for Null Results," *Psychological Bulletin*, 86 (3), 638-641.

Rust, Roland T. and Bruce Cooil (1994), "Reliability Measures for Qualitative Data: Theory and Implications," *Journal of Marketing Research*, 31 (February), 1-14.

Samuelson, Paul A. (1937), "A Note on Measurement of Utility," *The Review of Economic Studies*, 4 (February), 155-161.

Schmidt, Frank L. and John E. Hunter (1999), "Comparison of Three Meta-Analysis Methods Revisited: An Analysis of Johnson, Mullen, and Salas (1995)," *Journal of Applied Psychology*, 84 (1), 144-148.

Shrout, Patrick E. and Joseph L. Fleiss (1979), "Intraclass Correlations: Uses in Assessing Rater Reliability," *Psychological Bulletin*, 86 (2), 420-428.

Smith, Amy K., Ruth N. Bolton, and Janet Wagner (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36 (August), 356-372.

Sparks, Beverly A. and Janet R. McColl-Kennedy (2001), "Justice Strategy Options for Increased Customer Satisfaction in a Services Recovery Setting," *Journal of Business Research*, 54 (3), 209-218.

Teichman, Meir and Uriel G. Foa (1975), "Effect of Resources Similarity on Satisfaction with Exchange," *Social Behavior and Personality: An International Journal*, 3 (2), 213-224.

Walster, Elaine, Ellen Berscheid, and G. W. Walster (1973), "New Directions in Equity Research," *Journal of Personality and Social Psychology*, 25 (2), 151-176.

Weaver, Kimberlee, Stephen M. Garcia, and Norbert Schwarz (2012), "The Presenter's Paradox," *Journal of Consumer Research*, 39 (3), 445-460.

Webster, Cynthia and D. S. Sundaram (1998), "Service Consumption Criticality in Failure Recovery," *Journal of Business Research*, 41 (2), 153-159.

Whatley, Mark A., J. M. Webster, Richard H. Smith, and Adele Rhodes (1995), "The Effect of a Favor on Public and Private Compliance: How Internalized is the Norm of Reciprocity?" *Basic and Applied Social Psychology*, 21 (3), 251-259.

Wirtz, Jochen and Anna S. Mattila (2004), "Consumer Responses to Compensation, Speed of Recovery and Apology After a Service Failure," *International Journal of Service Industry Management*, 15 (2), 150-166.

Zauberman, Gal (2003), "The Intertemporal Dynamics of Consumer Lock-in," *Journal of Consumer Research*, 30 (3), 405-419.

Zauberman, Gal, B. K. Kim, Selin A. Malkoc, and James R. Bettman (2009), "Discounting Time and Time Discounting: Subjective

Time Perception and Intertemporal Preferences,'' *Journal of Marketing Research*, 46 (August), 543-556.

Zeithaml, Valarie A., A. Parasuraman, and Leonard L. Berry (1985), ''Problems and Strategies in Services Marketing,'' *Journal of Marketing*, 49 (Spring), 33-46.

Zhu, Zhen, K. Sivakumar, and A. Parasuraman (2004), ''A Mathematical Model of Service Failure and Recovery Strategies,'' *Decision Sciences*, 35 (Summer), 493-525.

Author Biographies

Holger Roschk (PhD in Business Administration, Ilmenau University of Technology) is a postdoc candidate at Catholic University Eichstaett–Ingolstadt. His research focuses on service recovery and meta-analysis. He has published in the *Journal of Service Research*, *Marketing Letters*, and *Journal of Retailing and Consumer Services*.

Katja Gelbrich (PhD in Business Administration, University of Stuttgart) is a full professor of International Management at Catholic University Eichstaett–Ingolstadt. Major fields of research are service recovery, international management, and cross-cultural marketing. She has published articles in the *Journal of the Academy of Marketing Science*, *Journal of Service Research*, *Journal of Retailing, Marketing Letters*, *Schmalenbach Business Review*, *Journal of Retailing and Consumer Services*, and *Journal of Services Marketing*.