The Effect of Compensation Size on Recovery Satisfaction after Group Service Failures

The Role of Group Versus Individual Service Recovery

Albrecht, Arne K.; Schäfers, Tobias; Walsh, Gianfranco; Beatty, Sharon E.

Document Version
Accepted author manuscript

Published in:
Journal of Service Research

DOI:
10.1177/1094670518802059

Publication date:
2019

License
Unspecified

Citation for published version (APA):
Albrecht, A. K., Schäfers, T., Walsh, G., & Beatty, S. E. (2019). The Effect of Compensation Size on Recovery Satisfaction after Group Service Failures: The Role of Group Versus Individual Service Recovery. *Journal of Service Research*, *22*(1), 60-74. https://doi.org/10.1177/1094670518802059

The Effect of Compensation Size on Recovery Satisfaction after Group Service Failures: The Role of Group Versus Individual Service Recovery

Arne K. Albrecht, Tobias Schäfers, Gianfranco Walsh, and Sharon E. Beatty

Journal article (Accepted manuscript*)

Please cite this article as:
Albrecht, A. K., Schäfers, T., Walsh, G., & Beatty, S. E. (2019). The Effect of Compensation Size on Recovery Satisfaction after Group Service Failures: The Role of Group Versus Individual Service Recovery. *Journal of Service Research*, *22*(1), 60-74. DOI: 10.1177/1094670518802059

DOI: https://doi.org/10.1177/1094670518802059

The effect of compensation size on recovery satisfaction after group service failures: The role of group versus individual service recovery

Abstract

Two experimental studies reveal that customers' reactions to different levels of recovery compensation differ between a recovery that occurs at the group level (such that every customer knows that every other affected customer receives the same compensation) and one that occurs at the individual level (such that the individual does not know if and how much compensation other affected customers receive). In both cases, recovery compensation exhibits diminishing returns on compensation size in terms of recovery satisfaction. However, at the group level, the rate at which the returns on compensation diminish is greater and satisfaction reaches a plateau at lower compensation levels than at the individual level. The salient social comparison made during a group service recovery, as evidenced by the mediating role of distributive justice, explains these effects. Finally, we note that at mid-range compensation levels, GSR and ISR did not lead to different levels of recovery satisfaction, suggesting a zone of tolerance or indifference at these levels. Further, our findings yield important managerial implications for the efficient allocation of service recovery resources after a group service failure.

Keywords: compensation, individual service recovery, group service failure, group service recovery, social comparison

Customers can affect one another in the normal course of service delivery (Argo, Dahl, and Manchanda 2005). However, relatively little is known about group level impacts on service failures and recoveries. Group service failures (GSF), which affect multiple customers simultaneously, are common, such as in delayed air travel or impaired hospitality services. Research has only started to examine the resulting effect of other customers, yielding initial evidence that people react differently to a GSF than to an individual service failure (ISF), which only affects one customer. Specifically, compared with an ISF, GSFs tend to result in more negative service outcomes (Du, Fan, and Feng 2014), because customers attribute more blame to the service provider (Albrecht, Walsh, and Beatty 2017). However, this does not necessarily mean that service firms have to apply different recovery efforts or allocate more resources to ISF- vs. GSF-related recovery efforts. What it does mean, though, is that service managers need to find ways to recover effectively from GSFs (Du, Fan, and Feng 2014) and need insight into managing ISF- vs. GSF-related customer recovery efforts in both a customer-oriented and cost-effective way.

Research suggests that service firms can restore equity by compensating aggrieved customers (e.g., De Ruyter, and Wetzels 2000) and that financial compensation is an efficient recovery method after service failures, with greater compensation resulting in more favorable customer outcomes (Hess, Ganesan, and Klein 2003; Smith, Bolton, and Wagner 1999). However, prior research has not yet differentiated GSFs from ISFs with regard to compensation size effects (e.g., Gelbrich, Gäthke, and Grégoire 2015; Smith, Bolton, and Wagner 1999). Few, if any, studies investigate how the size of financial compensation could affect customer outcomes. The need to understand in more detail how aggrieved customers respond to post-failure compensation offers is clear (Hess, Ganesan, and Klein 2003) given that marketing

managers face increasing pressure to justify spending valuable firm resources on service recovery efforts (Cambra-Fierro, Melero, and Sese 2015).

Accordingly, we seek to extend knowledge on service recovery by considering two potential choices firms face in providing financial compensation after a group service failure. With a group service recovery (GSR), the individual affected recognizes that all affected customers receive the same recovery, but for an individual service recovery (ISR), focal customers only know what their compensation is, and the compensation offered to other affected parties is not known (Zhou et al. 2013). By distinguishing between these two forms of service recovery (group vs. individual), we then examine the contingent effect of compensation size on recovery satisfaction. Additionally, we show that the interactive effect between compensation size and recovery form is mediated by distributive justice. Furthermore, we draw on and seek to replicate Gelbrich, Gäthke, and Grégoire's (2015) ISR research, which reveals diminishing returns of compensation size on satisfaction; with our replication we show that GSRs also exhibit diminishing returns, but that the curve progression differs from ISRs. This difference for GSR vs. ISR is important because service firms might inadvertently allocate money to service recovery efforts without yielding a meaningful increase in recovery satisfaction. Our findings in turn show that after a GSR, service providers can achieve higher levels of satisfaction with lower compensation and that greater compensation levels are inefficient, as they do not substantially increase satisfaction. This finding contrasts with Zhou et al.'s (2013) postulation that it is always advantageous to frame an economic recovery effort as a GSR to imply fair compensation for everyone.

We begin by discussing the effects of compensation size on recovery satisfaction, expound on why we expect the two to be nonlinearly related, and briefly discuss the literature on recovery

efforts in relation to GSFs. Then, we present our hypotheses and describe the research methodology as well as our empirical findings.

Conceptual Background

Compensation size and recovery satisfaction

Financial compensation is a useful, frequently applied means to achieve recovery after a service failure (Gelbrich and Roschk 2011; Grewal, Roggeveen, and Tsiros 2008; Roschk and Gelbrich 2017). More compensation after a service failure generally results in higher customer recovery satisfaction (Hess, Ganesan, and Klein 2003; Smith, Bolton, and Wagner 1999), due to differences in perceptions of the level of distributive justice (Gelbrich and Roschk 2011). In equity theory terms, higher compensation positively affects customer perceptions of their outcomes (compensation received) and the service provider's inputs (compensation offered) (Hess, Ganesan, and Klein 2003). Thus, for service failure and recovery, the level of distributive justice experienced will reflect the degree to which the customer perceives the recovery outcome to be fair (Gelbrich and Roschk 2011; Smith, Bolton, and Wagner 1999). Customers compare recovery outcomes with the degree of inconvenience caused by the service failure to estimate its fairness (Mattila and Patterson 2004). Therefore, greater compensation increases the individual's perceptions of distributive justice, given the same level of inconvenience (Gelbrich, Gäthke, and Grégoire 2015; Smith, Bolton, and Wagner 1999), and distributive justice predicts the customer's satisfaction with the recovery (Gelbrich and Roschk 2011; Roschk and Gelbrich 2017; Tax, Brown, and Chandrashekaran 1998), holding all else constant (e.g., severity).[1]

[1] Other justice dimensions have been considered in service recovery studies, namely procedural and interactional justice (e.g., Gelbrich and Roschk 2011). However, because the specific question of group versus individual service recovery effects relates to social comparison processes and thus distributional perceptions, we focus on distributive justice as a process variable.

However, some past research indicates that the effect of compensation size on customer satisfaction cannot be described adequately by a linear relationship and questions the notion that more compensation will always result in higher satisfaction (Gelbrich, Gäthke, and Gregoire 2015; Hess, Ganesan, and Klein 2003).

Nonlinear Effect of Compensation Size

Gelbrich, Gäthke, and Grégoire (2015) investigate compensation size and satisfaction after a service failure for an 'accepted service', that is, one that falls below the customer's expectations, such that the customer still consumes the product but the experience is diminished relative to its value-in-use. They find a concave relationship between compensation size and satisfaction after a service failure, reflecting the law of diminishing marginal utility with saturation. Thus, increases in compensation should have the greatest impact on customer satisfaction at low compensation levels. At higher compensation, further increases instead invoke continuously smaller gains in customer satisfaction. Gelbrich and colleagues (2015) affirm that recovery satisfaction levels reach a plateau after a certain compensation size. What has not been studied before is if a concave curve will describe the outcomes of both a GSR and an ISR, and whether and how these curves differ in concaveness.

Recovery after a Group Service Failure

After a GSF, a service recovery may be presented as a GSR, with a clear statement that every affected customer will receive the same recovery, or it may be presented as an ISR, in which case it is framed as a personal recovery for the individual involved. Some evidence exists indicating that a social recovery effort (e.g., an apology) should be framed as an ISR to produce maximum effectiveness, whereas an economic recovery effort should be framed as a GSR (Zhou et al. 2013). Zhou et al. (2013) argue that a GSR provides the knowledge that other customers

receive the same compensation, in contrast to ISR cases, which provide no such information. Thus, GSR (vs. ISR) reduces the customer's uncertainty about whether the service provider is offering fair compensation. This notion is consistent with social comparison theory, which posits that people tend to evaluate their compensation through social comparisons with relevant referents (Festinger 1954; Gerber, Wheeler, and Suls 2018). Thus, Zhou et al.'s (2013) findings would predict that an economic compensation will always result in more favorable customer outcomes with a GSR setup versus an ISR. However, we challenge this postulation by proposing that the recovery form does not exert a direct effect on recovery satisfaction but rather it interacts with other recovery characteristics, such as the amount of the compensation, to shape recovery satisfaction.

Specifically, according to social comparison theory, the GSR curve should be characterized by a steeper incline and a plateau at lower values on the x-axis (i.e., compensation size) than the ISR curve, because it sparks a focus on the compensation of fellow customers. This strong lateral social comparison may affect customers' judgments about the favorability of this comparison, consistent with the tenets of equity theory (Oliver and Swan 1989), and may prevent them from experiencing greater satisfaction when compensation is further increased. The curve in the ISR condition instead should incline at a steadier rate and reach the plateau at a higher value on the x-axis (i.e., at a higher compensation size), than the GSR curve, because no social comparison is available, which makes the amount of compensation more relevant to customer satisfaction.

Compensation Size and Recovery Form

Prior research on the influence of compensation size after a service failure focused on the effect of compensation size on customer satisfaction, but did not differentiate between GSFs and ISFs (e.g., Gelbrich, Gathke, and Grégoire 2015; Smith, Bolton, and Wagner 1999). It therefore

remains unclear what role compensation size plays after a GSF in terms of relevant customer outcomes. In particular, it is unclear if the effect of compensation size after a GSF depends on the form of the recovery (i.e., GSR vs. ISR). Prior assumptions, implying a universally positive effect of GSR versus ISR (Zhou et al. 2013), suggest main effects of compensation size and recovery form on recovery satisfaction; customers who receive a recovery should be more satisfied with it when compensation size is greater and when they receive the recovery as a GSR rather than an ISR. However, we instead predict an interaction effect between the recovery form (i.e., GSR vs. ISR) and compensation. In particular, according to fairness heuristic theory, the best type of information for assessing distributive justice comes from social comparisons (Collie, Bradley, and Sparks 2002; van den Bos et al. 1997), such that people interpret relevant information about themselves and their current situation by comparing themselves and their own situation to those of others in similar situations (Collie, Bradley, and Sparks 2002; Festinger 1954).

With a GSR, the social comparison information about the compensation offered to other customers is available and salient. Thus, a customer can use this information to assess the fairness of the outcome (McColl-Kennedy, Daus, and Sparks 2003; van den Bos et al. 1997). Here, we focus on a situation in which after a GSF every customer affected by the failure receives the same compensation during the GSR, which represents a lateral social comparison (Strickhouser and Zell 2015). That is, all affected customers receive the same compensation for the same failure and they know it. The salient comparison dimension is social comparison, and the relative outcome for the affected customer is the same for any compensation size. In this case, the focal customer's justice perceptions and satisfaction are heavily shaped by the compensation size he or she believes all affected parties are receiving (McColl-Kennedy, Daus,

and Sparks 2003; van den Bos et al. 1997). Thus, perceived distributive justice and recovery satisfaction should not vary considerably across different compensation sizes. In other words, in a GSR, even small compensation sizes should be perceived as fair (owing to the lateral comparison) and increases in compensation likely invoke smaller increases in customer satisfaction compared to an ISR. Thus, following the findings from Gelbrich, Gäthke, and Grégoire (2015), we predict diminishing returns on compensation size, but the rate at which the returns diminish should be greater for GSR than for ISR.

In contrast, an ISR offers no information about the compensation of other customers. In this case, customers will tend to use the absolute amount of compensation to weigh its value against the value of their loss (due to the service failure), thus estimating the fairness of the compensation (Bonifield and Cole 2008; Smith, Bolton, and Wagner 1999). An ISR does not provide any cues for making social comparisons, so customers only have the absolute compensation size to use in assessing the fairness of the recovery. More compensation indicates a better outcome and should result in increased customer satisfaction. In this ISR scenario, in the absence of social comparison information, recovery satisfaction should increase as compensation size increases, although with diminishing returns (Gelbrich, Gäthke, and Grégoire 2015). More importantly, the return on additional compensation for an ISR should diminish at a lower rate than for a GSR. We thus predict an interaction of compensation size and recovery form, such that the relationship between compensation size and recovery satisfaction should be more concave in a GSR setting than in an ISR setting.

H1: The recovery form (group versus individual service recovery) moderates the effect of compensation size on recovery satisfaction, such that for a GSR, the rate at which the

returns diminish is greater and satisfaction reaches a plateau at lower compensation levels than for an ISR.

Furthermore, we predict distributive justice will explain the hypothesized differences in the effects of compensation size on recovery satisfaction for GSR and ISR. Prior research found justice perceptions to mediate the effect of compensation size on recovery satisfaction (e.g., Gelbrich, Gäthke, and Grégoire 2015; Roschk and Gelbrich 2017). Taking into account that customers judge the fairness of varying compensation sizes differently, depending on the recovery form, we expect distributive justice to mediate the predicted moderation effect of recovery form.

H2: Distributive justice mediates the moderated effect of compensation size, by recovery form, on recovery satisfaction.

Experimental Studies

We conducted two experiments to test our hypotheses. The studies differ in several respects, including sample, service price, type of service failure, and compensation size, to increase the generalizability of our findings. In the first study, we investigate the assumed nonlinear relationship between compensation size and recovery satisfaction, as well as the moderating influence of recovery form (H1). In Study 2, we validate these findings, rule out alternative explanations by including several control variables, and test the mediating role of distributive justice (H2). Both studies feature German respondents, and we used translations/back-translations by bilingual, native English speakers to modify all the scales employed in both studies. With a scenario-based approach, we follow prior service failure research (e.g., Albrecht, Walsh, and Beatty 2017; Gelbrich 2010) that notes their advantages, in that they do not require customers to endure artificially created service failures. Unlike Gelbrich,

Gäthke, and Grégoire (2015), who used six compensation levels between 0% and 100% (0/20/40/60/80/100), we use 7 percentage levels (0/17/33/50/67/83/100) to allow for a more nuanced assessment of the relationship between compensation size and recovery satisfaction.

Study 1

Study Design and Method

The aim of Study 1 is to empirically test H1. We thus evaluate the effects of different compensation sizes on recovery satisfaction in group and individual recovery settings.

Procedure. Study 1 used a 7 (compensation = €0, €10, €20, €30, €40, €50, €60) × 2 (GSR vs. ISR) between-subjects experimental design. We applied a hotel context, similar to the one used by Gelbrich (2010), in which the participant is asked to imagine a stay in a hotel for the price of €60. First, participants read the service-failure scenario which pertains to an unpleasant hotel stay due to waking up early because of noise. Then, they read the response letter which begins with an apology to the customer and contains the recovery form and compensation size manipulations. The detailed scenario descriptions appear in Appendix A. After the manipulations, participants responded to items capturing the dependent variable.

Participants. Graduate research assistants recruited students from a German university by approaching them on campus and asking them to participate voluntarily in a study on service failure/recovery; they were offered chocolate as participation incentives. The assistants recruited a sample of 511 participants (mean age = 22.87, S.D. = 3.83; 234 (45.8%) male, 277 (54.2%) female), who were then scheduled to attend a lab session. During these sessions they were randomly assigned to one of the 14 conditions. Fifteen participants recalled the wrong hotel price and thus were dropped. Of the remaining 496 participants, 268 were women (54%) with an average sample age of 22.9 years (*SD* = 3.85). Student samples are advantageous in reducing

extraneous variance (because of relative sample homogeneity) and have been found suitable for scenario-based studies into service failures (e.g., Hess et al. 2003).

Manipulation Check. As a manipulation check, we asked participants, “Who receives the compensation?” with the possible answers “only me,” “me and other people,” “only other people,” “nobody,” and “I don’t know.” In this sample, 102 participants (21%) did not answer the manipulation check correctly and were removed from further analyses, leaving 394 respondents; the demographic characteristics are essentially the same as those of the full sample.

Measures. We measured recovery satisfaction with three items (see Appendix B). Parallel to prior research on nonlinear effects of service recovery (Gelbrich, Gäthke, and Grégoire 2015), we anchored the recovery satisfaction items on an 11-point scale (1 = “strongly disagree” to 11 = “strongly agree”) to create adequate variance in the curve progression. An assessment of the scale’s psychometric properties, as shown in Appendix B, indicates satisfactory levels of reliability and convergent validity.

Results

Prior research suggests using a square root function for models with diminishing returns (Deighton, Henderson, and Neslin 1994; Homburg, Müller, and Klarmann 2011; Nam, Manchanda, and Chintagunta 2010). We thus estimated a structural equation model—in order to account for the latent nature of recovery satisfaction—that captures the diminishing returns by including the effects of compensation size (CS) and the square root of compensation size ($CS^{0.5}$). Together, both variables determine the concaveness of the nonlinear effect and thus the hypothesized plateau. Data analysis was conducted using Mplus version 8 (Muthén and Muthén 1998-2017).

In a first step, we assessed whether the hypothesized diminishing returns structure explained the data better than alternative models. Specifically, we compared the square root model (Table 1, column A), to a linear, a quadratic, and a cubic model (not shown in the table). Because this assessment results in comparisons of non-nested models, commonly used fit indices (e.g., chi-square difference test) do not provide meaningful results; instead, we assessed the models based on Bayesian Information Criterion, or BIC (Raftery 1995) and explained variance. The square root model, which explains 54.3% of the variance in recovery satisfaction exhibits a substantially better fit than the linear (ΔBIC = 63.53, R^2 = .442), quadratic (ΔBIC = 17.44, R^2 = .519), and cubic (ΔBIC = 10.73, R^2 = .537) models. More importantly, the difference in BIC exceeds 10 for all models, which provides very strong evidence for superior fit of the square root model (Raftery 1995).

--Insert Table 1 about here--

After establishing the diminishing returns of compensation size on recovery satisfaction, we tested the hypothesized moderation of this effect by recovery form. As explained, the predictors CS and $CS^{0.5}$ jointly determine the concave curve progression. Additionally, the two variables are necessarily highly correlated, as one is calculated based on the other. Therefore, testing the moderating effect with two interaction terms (i.e., CS × recovery form, $CS^{0.5}$ × recovery form) is likely to produce biased estimates. Instead, we analyzed a multi-group SEM to jointly test for differences in the two predictors' parameters between ISR and GSR. First, we estimated a model that was fully invariant across the two groups (Table 1, column B). Second, we estimated a model in which the two structural paths from CS and $CS^{0.5}$ to recovery satisfaction were unconstrained (Table 1, column C). This partly unconstrained model exhibits more favorable values across all fit indices; a chi-square difference test indicates that it fits the

data better than the multi-group invariant model in column B of Table 1 ($\Delta\chi^2$ (3) = 10.442, p < .05). Moreover, a Wald test reveals that the two path coefficients for CS and $CS^{0.5}$ on recovery satisfaction differs across the two groups ($W\chi^2$ (2) = 7.660, p < .05). The results thus support the assumption that the nonlinear effect of compensation size on recovery satisfaction differs in concaveness between GSR and ISR.

To obtain a better understanding of how the concaveness differs between the two recovery forms, we used the path coefficients for CS and $CS^{0.5}$ to plot the two estimated curves. Furthermore, we included the latent variable means for each compensation level, as estimated in a multi-group confirmatory factor analysis with a mean structure (Figure 1). It is clear that for the GSR context, the increase in satisfaction yielded from additional compensation is smaller than the increase achieved in an ISR context. Thus, the rate at which the returns on compensation size diminish is greater and recovery satisfaction reaches a plateau at lower compensation levels for GSR than for ISR, supporting H1.

--Insert Figure 1 about here--

Discussion

Study 1 confirms the nonlinear relationship between compensation size and recovery satisfaction as well as the moderating role of recovery form (GSR vs. ISR). That is, both curves feature a concave shape, but the curve for GSR is characterized by a stronger incline at low compensation values. The plateau occurs at a lower compensation size than in the ISR curve, which keeps growing with increases in compensation and reaches its saturation point at a higher compensation size, consistent with our expectations.

Study 2

Study Design and Method

The second study was again a 7 (compensation: €0, €75, €150, €225, €300, €375, €450) × 2 (GSR vs. ISR) between-subjects experiment. To assess the robustness and generalizability of the findings from Study 1, we used the same context (i.e., hotel), but recruited a broader sample, made changes to the employed scenario, and captured several control variables to serve as covariates in the analysis and to rule out alternative explanations, all of which we describe in detail next.

Procedure. We employed an online survey to allow for a more vivid description of the experimental scenario by including pictures and to ensure a standardized procedure for all participants. The study began by asking participants to imagine spending three nights at a hotel in another city for a total price of €450. We deliberately selected a higher priced context than in Study 1, which also resulted in higher absolute compensation amounts, to ensure that the findings are robust across various price ranges. The service failure pertained to the limited availability of the hotel's fitness and pool facilities. Instead of a service failure relating to the core service, Study 2 focused on a failure that affects peripheral aspects. Next, participants judged the severity of the service failure, to be included as a covariate, and were randomly assigned to one of the fourteen experimental conditions. The manipulations of the compensation amount and the recovery form, which were then presented, were similar to the ones used in Study 1, but were complemented with pictures and visualizations (e.g., hotel entrance, pool facility). Appendix A provides the detailed scenario descriptions. After the manipulations, participants responded to items capturing the dependent variable, the mediator, and the additional covariates.

To ensure that respondents carefully read the stimulus material, we used three manipulation and attention check questions at the end of the survey. Specifically, we asked

participants to recall the price of the hotel ("What was the initial price for the three night hotel stay?" €150 / €250 / €350 / €450 / €550), the compensation amount ("How high was the compensation that the hotel manager offered?" He did not offer any compensation [€0] / €75 / €150 / €225 / €300 / €375 / €450), and whether it was a GSR or an ISR scenario ("The hotel manager's response…" …was referring only to me personally. / …was referring to all affected guests.).

Participants. In contrast to the student sample used in Study 1, we paid a German market research firm to recruit participants of a professionally hosted national online panel. Panel members receive monetary incentives from the panel provider. The panel provider adheres to ICC/ESOMAR standards (www.esomar.org) by using a double opt-in registration process and by actively managing and continuously refreshing the panel via online and offline recruitment to address aspects such as recruitment bias and conditioning bias (Göritz 2004). Panel members agree to participate in, on average, one self-administered online survey per month in return for small monetary incentives. In total, 703 individuals completed the survey in a median time of 6.68 minutes (mean = 8.80 minutes, *SD* = 9.69 minutes). We checked for invalid responses following the suggestions of DiLalla and Dollinger (2006): 24 participants sped through the questionnaire (< 3 minutes), 93 failed at least one of the three check questions, leaving 586 respondents. Furthermore, five respondents were outliers in their responses regarding recovery satisfaction or distributive justice (e.g., indicating the highest satisfaction score in the no compensation condition). Data cleaning thus led to an analyzable sample of 581 (82.65%) respondents (50.9% female, mean age = 43.93 years, *SD* = 14.09). Cell sizes for the 14 experimental conditions ranged from 35 to 46 individuals.

Measures. All measures appear in Appendix B. The same three-item scale as in Study 1 was used to measure recovery satisfaction (Gelbrich, Gäthke, and Grégoire 2015). Distributive justice was assessed using the four-item scale by Smith, Bolton, and Wagner (1999). To measure failure severity, we employed a three-item scale from Maxham and Netemeyer (2002). All scales exhibit satisfactory reliability and convergent validity scores, as illustrated in Appendix B. Moreover, discriminant validity was supported, as each scale's average variance extracted (AVE) exceeds the squared correlations with the other two constructs (Fornell and Larcker 1981).

In addition to the three latent constructs, we assessed age, gender, and the following control variables using single questions. The relevance of the service failure to respondents was captured by asking whether they would use a hotel's fitness and pool facilities in real life (7-point scale, anchored by "absolutely not" and "absolutely"). This variable was included to account for possible differences in reactions depending on whether the service failure affects a service aspect that respondents view as an integral part of their stay. Complaint experience ("How often have you complained to a company about a problem with a service?") and recovery experience ("How often have you received compensation from a company in reaction to a complaint?"), measured on a 7-point scale (1 - "never", 7 - "very often"), were included to account for possible internal reference points that respondents may use when assessing the compensation (Gerber, Wheeler, and Suls 2018) and because previous experience with complaint and service recovery can create expectations about future recovery situations and thus reduce the impact of other reference points such as the compensation delivered to other customers (Seta and Seta 1992).

Common Method Variance. Because we collected data from the same respondents at the same point in time, the possibility of common method bias was considered as recommended by Podsakoff et al. (2003). As procedural remedies, respondents were encouraged to answer honestly and informed that there were no right or wrong answers. To reduce potential common method variance we captured recovery satisfaction on an 11-point scale and distributive justice on a 7-point scale, adhering to Gelbrich, Gäthke and Grégoire's (2015) work. To assess whether common method variance impacted the obtained results, we employed the marker variable technique (Lindell and Whitney 2001). As a marker variable that is theoretically unrelated to the focal variables, we captured respondents' agreement to the statement "On social media (e.g., Facebook), I react to other messages politely and respectfully, even if they are impolite or insulting." In a confirmatory factor analysis, we regressed each of the items measuring one of the three latent constructs (i.e., recovery satisfaction, distributive justice, and failure severity) on the marker variable and compared the results to a model in which the loadings of the marker variable were set to zero. The two models did not differ in model fit ($\Delta\chi^2$ (10) = 15.245, p = .123) and the correlations between the latent constructs were not affected by including the marker variable. Moreover, when we included the marker variable in the estimations required for hypothesis testing, the results remained consistent, with only minor differences in some of the path coefficients. Overall, we thus judged that common method variance did not bias our results.

Results

In order to test for the hypothesized moderated and mediated nonlinear effects of recovery compensation, we again analyzed a multi-group SEM using Mplus (Muthén and Muthén 1998-2017). The independent variables were CS, the $CS^{0.5}$, the recovery form, and the covariates (i.e., age, gender, failure severity, failure relevance, complaint experience, and recovery experience).

For recovery satisfaction, the model estimation revealed a good fit to the data, as shown in Table 2, column A. Together, the negative path coefficient for CS and the positive path coefficient for $CS^{0.5}$ reveal diminishing returns, in line with the findings from Study 1.[2] To assess the hypothesized moderation of this effect by recovery form, we again analyzed a multi-group SEM to compare a constrained model, in which all structural paths were set invariant (Table 2, column B), to a model in which the structural parameters were freely estimated across the two groups (Table 2, column C). The partly unconstrained model exhibited a better fit across various indices. A chi-square difference test revealed that the partly unconstrained model fit the data better than the multi-group invariant model ($\Delta\chi^2$ (3) = 20.645, $p < .001$). Additionally, a Wald test indicates that the two path coefficients for CS and $CS^{0.5}$ on recovery satisfaction differs across the two groups ($W\chi^2$ (2) = 19.129, $p < .001$). Overall, the results provide evidence that the concaveness of the curves predicting recovery satisfaction based on compensation differs between the GSR and ISR contexts.

—Insert Table 2 and Figure 2 about here—

Based on the partly unconstrained multi-group analysis results, we plotted the two nonlinear functions, depicted in Figure 2. This plot illustrates that for GSR, the effect of compensation size on recovery satisfaction exhibits a more concave pattern than for ISR. For GSR, the level of satisfaction reaches a plateau at substantially lower compensation levels than for ISR. These findings are consistent with Study 1's findings and support H1.

Before testing the hypothesized mediating effect of distributive justice, we conducted individual analyses for this variable, which appear in Table 3. Similar to recovery satisfaction,

[2] We again assessed competing relationships between compensation size and recovery satisfaction, all of which exhibited a poorer model fit and explained less variance than the square root model (linear: $R^2 = .543$, $\Delta BIC = 353.31$; squared: $R^2 = .729$, $\Delta BIC = 76.67$; cubic: $R^2 = .760$, $\Delta BIC = 19.96$).

compensation size exhibits diminishing returns to distributive justice (see Table 3, column A).[3] The multi-group SEM results, shown in columns B and C of Table 3, show that not constraining the nonlinear effects of compensation size to be equal increases model fit over the multi-group invariant model ($\Delta\chi^2$ (2) = 11.633, $p < .01$). Furthermore, a Wald test supports the assumption that the path coefficients for CS and $CS^{0.5}$ are different between GSR and ISR ($W\chi^2$ (2) = 11.748, $p < .01$). These results thus indicate that recovery form moderates the diminishing returns of compensation size on distributive justice.

—Insert Table 3 and Figure 3 about here—

The comparison between the two recovery forms, illustrated in Figure 3, reveals that for GSR, compensation size produces greater diminishing returns than for ISR. In line with the results for recovery satisfaction, a plateau is reached at a lower compensation size for GSR compared to ISR.

Finally, we analyzed the proposed mediation via distributive justice with a multi-group path analysis using bootstrapping with 10,000 samples. To test for differences between the two recovery forms, we compared the fit of a constrained model, in which all structural paths were set invariant (χ^2 (144) = 316.182, χ^2/df = 2.196, CFI = .973, TLI = .971, RMSEA = .064 [90% CI: .055 .074], SRMR = .040, BIC = 13,124.08), to the fit of the hypothesized model in which the structural parameters for CS and $CS^{0.5}$ were freely estimated across the two groups (see Figure 4). The latter model fits the data better than the former invariant model ($\Delta\chi^2$ (6) = 24.948, $p < .001$). According to a Wald test, the path coefficients of CS and $CS^{0.5}$ on distributive justice and recovery satisfaction differ across the two groups ($W\chi^2$ (4) = 23.254, $p < .001$). As shown in Figure 4, and in agreement with H2, distributive justice acts as a mediator for the nonlinear

[3] Comparisons with alternative models provided support for the hypothesized square root model (linear: R^2 = .488, ΔBIC = 287.07; squared: R^2 = .688, ΔBIC = 36.69; cubic: R^2 = .710, ΔBIC = 5.19).

effects of compensation size on recovery satisfaction in both GSR and ISR contexts. It is notable, however, that the direct effects of compensation (i.e., CS and $CS^{0.5}$) remain, thus indicating a partial mediation of distributive justice.

—Insert Figure 4 about here—

We also tested for differences in the indirect and the total effects of CS and $CS^{0.5}$on recovery satisfaction, as suggested by Ryu and Cheong (2017). The results indicate that the indirect nonlinear effect of compensation size on recovery satisfaction via distributive justice differs between GSR and ISR ($W\chi^2$ (2) = 10.725, $p < .01$). Similarly, the total nonlinear effect, comprised by the direct effect and the indirect effect via distributive justice, is different ($W\chi^2$ (2) = 19.258, $p < .001$). Thus, the diminishing returns on compensation size to recovery satisfaction via distributive justice are greater in a GSR context than in an ISR context. This is evidenced by the greater negative indirect and total effects of CS in combination with the greater positive indirect and total effects of $CS^{0.5}$. That is, when a GSR is used, the returns on recovery compensation diminish at a greater rate and satisfaction reaches a plateau at lower compensation levels than when ISR is used, with distributive justice providing the explanation of this effect.

Discussion

Overall, the results of Study 2 validate the findings of Study 1. We find support for the hypothesized diminishing returns on compensation regarding recovery satisfaction as well as the moderation of this effect by recovery form. In addition to replicating the findings, Study 2 provides evidence for the robustness of the results, due to the broader sample, a different type of failure, and the included covariates. Moreover, we tested for the hypothesized role of distributive justice in explaining the effects we uncovered for recovery satisfaction (H2). Our findings indicate that after experiencing a group service failure, consumers who subsequently receive a

GSR perceive lower compensation levels as fairer than individuals who receive an ISR. However, the increases in distributive justice due to higher compensation size are relatively small in GSR, thus creating a plateau, whereas in ISR, increases in compensation size lead to more gains in distributive justice.

General Discussion

After experiencing a service failure, customers often receive remuneration (Mattila 2001). How generously service firms should compensate customers in different failure situations and with different recovery forms is not well understood (Gelbrich, Gäthke, and Grégoire 2015). This research seeks to specify how service recovery following a group service failure can achieve higher customer satisfaction by accounting for compensation size (i.e., costs) and recovery form. Across two studies, we find evidence for a moderating effect, indicating that compensation levels and recovery forms should be considered together rather than in isolation. The results confirm prior research that more compensation results in higher customer satisfaction (Smith, Bolton, and Wagner 1999), but that the relationship between compensation size and customer satisfaction is not necessarily linear (Gelbrich, Gäthke, and Grégoire 2015; Hess, Ganesan, and Klein 2003). For a range of compensation from 0% to 100% of the price of the failed service, the curve between compensation size and satisfaction was concave in our two studies. Thus, our results align with previously found diminishing returns on compensation regarding recovery satisfaction (Gelbrich, Gäthke, and Grégoire 2015).

Importantly, we also show that recovery form moderates the effect of compensation on recovery satisfaction. After a group service failure, the shape of the compensation–satisfaction curve depends on the delivery of the compensation, as either a group service recovery or an individual service recovery. For a GSR, a strong increase in satisfaction results from raising

compensation at lower ranges. However, saturation occurs, and no further improvement in satisfaction results from higher amounts of compensation. For an ISR, the pattern differs and compensation increases recovery satisfaction even at higher levels of compensation, although diminishing returns are observed in this context as well.

The moderation results can be explained by the availability of social comparison information in the group service recovery condition. In social comparison theory terms, customers prefer to evaluate their situation or outcome using objective and nonsocial standards; however, if no such objective information is available, then customers will compare their situation or outcome against that of other (affected) customers (Gerber, Wheeler, and Suls 2018). Knowing that other customers receive the same amount of compensation creates a lateral social comparison. This comparison indicates the fair treatment of the customer and quickly leads to a certain level of satisfaction, but satisfaction does not increase with further enhancements to compensation. More compensation fails to prompt greater satisfaction because the social comparison with the other customers is the most salient aspect during the group service recovery situation. For individual service recovery situations, no social comparison is available, so customers can rely only on the compensation they received to assess its fairness. Further, our final assessment indicates that distributive justice provides an explanation as to what drives the ISR and GSR difference effects.

Additionally, although not hypothesized, our results contradict prior assumptions that group service recovery always leads to higher satisfaction in economic recovery settings (Zhou et al. 2013, 2014). The positive effect of a group service recovery on recovery satisfaction receives support only at low compensation levels (i.e., 17% and 33%). At high compensation

levels (i.e., 83%), the opposite pattern emerges.[4] Moreover, while the effect of recovery form is specific to the compensation size, we note that at mid-compensation levels in both studies (i.e., at 50% and 67%), there are no differences in recovery satisfaction between GSR and ISR strategies. Thus, these mid-range levels appear to form a zone of tolerance or indifference (cf. Johnston 1995; Zeithaml, Berry, and Parasuraman 1993) in which recovery satisfaction is not dependent on GSR/ISR effects.

Managerial Implications

Service firms may try to avoid compensating aggrieved customers or keep compensation costs low (Aula 2010; Estelami 2000), but if they do not understand how compensation size and recovery framing affect customer recovery satisfaction, they may make suboptimal allocation decisions. Our findings have implications for how service managers should frame recovery efforts after group service failures (as a group service recovery or individual service recovery) and the amount of compensation they should offer. Our findings suggest that after group service failures, firms should frame service recoveries as either a group service recovery or an individual service recovery in coordination with the amount of compensation that the firm can offer (i.e., depending on the resources allocated to service recovery efforts) and depending on the number of affected customers. If only moderate compensation is offered, a group service recovery results in maximal customer satisfaction; if higher compensation levels are desirable, an individual service recovery would result in maximal customer satisfaction. Whether framing the service recovery or adjusting the compensation size is more preferable depends on which variable the service provider can influence most effectively. If both can be managed equally well, service

[4] The described findings emerged in a post-hoc analysis of the individual compensation levels. In Study 1, recovery satisfaction was greater for GSR than for ISR at a compensation level of 33% ($p < .05$). In Study 2, recovery satisfaction was greater for GSR than for ISR at compensation levels of 17% ($p < .01$); at a compensation level of 83%, satisfaction was greater for ISR than for GSR ($p < .05$).

firms can decide between a cost-efficient (low–medium compensation as a group service recovery) or a satisfaction-maximizing (high compensation as an individual service recovery) strategy.

More importantly perhaps, our findings demonstrate that, compared with an ISR, a GSR enables service providers to attain higher levels of recovery satisfaction at lower compensation levels. Therefore, ceteris paribus, GSR is a better option than ISR. However, we caution against applying a blanket approach here because the advantage of lower costs (i.e., less money paid for compensation) should be considered in combination with the group size that needs to be compensated. In other words, compensating a few customers with €100 each might still be cheaper than compensating many at €25 each.

Further, given a zone of tolerance or indifference relative to mid-range compensation levels, these findings suggest that firms study and assess their customers' zones of tolerance as to recovery compensation levels. With this information, firms may find that customers are indifferent as to how compensation is delivered (i.e., GSR versus ISR) at mid-range compensation levels; thus, lower compensation levels might suffice and could help to maximize profits while maintaining equivalent recovery satisfaction levels.

In line with increasing research on methods to determine returns on marketing activities (e.g., Hamilton et al. 2017), service firms may be interested in assessing the return on the compensation they offer using different recovery forms. Our findings should encourage service firms to track their customers' recovery satisfaction across these specific variables. For example, they could email customers a link to a short survey that asks them to report on their service recovery experience after either a group service recovery or individual service recovery, at varying compensation sizes. Such a fine-grained analysis could help firms determine the optimal

compensation size. In a parallel effort, service firms could log service failures and customers' responses to compensation in customer databases. Then they can better determine customer profitability and whether to adjust the compensation size offered to a given customer. Extant research suggests the usefulness of differentiating customer groups in relation to recovery efforts; Gelbrich, Gäthke, and Grégoire (2016) indicate that when a service firm offers higher compensation, strong (vs. weak) relationship customers may exhibit greater recovery satisfaction.

Limitations and Further Research

Several limitations to our study should be noted. First, we gathered our data with two written scenarios involving fictitious service providers. With this approach, we could control what information study participants received. A field design using a real service provider would represent a less abstract situation for participants and improve the external validity of the experiments.

Second, the service failures presented in these scenarios were one-off incidents and relatively moderate failures. Noting that failure severity in Study 2 was generally significant as a covariate, it is safe to say that the compensation–satisfaction curve progressions for group service recovery and individual service are likely to differ based on failure severity as well as for situations involving repeated service failures with the same provider. Continued studies should consider different failure levels and repeated service failures.

Future studies should also examine the effects of other important mediators, such as the effect of the customer's perceived controllability of the failure (Choi and Mattila 2008), the customer's involvement with and loyalty to the firm, and possible contextual effects, such as recovery time (Hogreve, Bilstein, and Mandl 2017), on the compensation-satisfaction link.

Additionally, it would be worthwhile to consider other mediators, such as individual differences (e.g., personality variables such as self-efficacy or sensitivity to environmental stimuli and other demographic variables, such as income or education), as well as potential situational effects (e.g., perceived scarcity; Seta and Seta 1992).

Third, our results reflect a group service failure context, such that customers are aware that they are not the only ones affected. Our results likely do not transfer to individual service failure contexts in which customers perceive themselves as the only ones affected. For example, a group service recovery after an individual service failure may result in universally negative results because the customer may suddenly realize that the failure is a group service failure, not an individual service failure, only after being offered a group service recovery. This reframing could be detrimental in terms of post recovery outcomes, because group service failures are usually perceived as worse than individual service failures. Further research should investigate the effects of these recovery forms in the contexts of both group service failures and individual service failures.

Fourth, we only considered distributive justice as a mediator for the effects of compensation on recovery satisfaction. This focus was based on the notion that group versus individual service recovery mainly leads to differences in social comparison effects, which should be reflected in distributive justice perceptions. However, depending on the implementation of both types of service recovery (e.g., group service recovery in the presence/absence of other customers), interactional and procedural justice perceptions may also be affected and should be considered in future studies.

Finally, our research is premised on the assumption that aggrieved customers who receive compensation will evaluate their compensation through social comparisons with relevant

referents. While we varied the size of the compensation received by all aggrieved customers in the experiments, we did not manipulate the size of the compensation that the other customers received. Further study of the relationship between compensation size and recovery satisfaction in conditions of lower and higher compensation for other customers can help us understand better how customers' recovery satisfaction is shaped.

Appendix A. Experimental Stimuli

	Group service recovery	**Individual service recovery**
Study 1		
Introduction:	You stay in a hotel during a short vacation for a price of €60. Your room is close to the hotel's breakfast area. You wake up at 4:30 a.m. due to loud noises and can't go back to sleep. Later during breakfast, you learn that other guests were affected by the noise as well. The noise was due to the hotel's breakfast preparations. You go to the hotel reception to complain. You leave your address and get reassured that you will be contacted.	
Compensation level:		
€0 (0%)	"Dear Mrs./Mr. XXX, We are sorry to hear that you were not fully satisfied with your stay on January 10, 2016. We want to apologize for the problems during your stay. Unfortunately, we cannot offer *any of the affected customers any financial compensation*. Kind regards Marcus Miller, Customer Service"	"Dear Mrs./Mr. XXX, We are sorry to hear that you were not fully satisfied with your stay on January 10, 2016. We want to apologize for the problems during your stay. Unfortunately, we cannot offer *you any financial compensation*. Kind regards Marcus Miller, Customer Service"
€10 (17%) €20 (34%) €30 (50%) €40 (67%) €50 (83%) €60 (100%)	"Dear Mrs./Mr. XXX, We are sorry to hear that you were not fully satisfied with your stay on January 10, 2016. We want to apologize for the problems during your stay. We want to offer *all affected customers a compensation* of [*€10/€20/€30/€40/€50/€60*]. Please find a check enclosed with this letter. Kind regards Marcus Miller, Customer Service"	"Dear Mrs./Mr. XXX, We are sorry to hear that you were not fully satisfied with your stay on January 10, 2016. We want to apologize for the problems during your stay. We want to offer *you a personal compensation* of [*€10/€20/€30/€40/€50/€60*]. Please find a check enclosed with this letter. Kind regards Marcus Miller, Customer Service"
Study 2		
Introduction:	Imagine that you are spending an extended weekend in another city. For that reason, you booked three nights in a hotel for a total price of €450. The price includes access to the hotel's fitness center and pool. When you booked the hotel on the Internet, you saw that other customers evaluated the fitness center and the pool very positively. You are very much looking forward to using these facilities extensively during your stay. On the day after your first night you want to use the fitness center and the pool. However, you find that in the gym, four of the five cross trainer machines are out of order. One cross trainer and the only available treadmill are currently being used by other guests. Two more guests are already waiting for the machines to become available again. In the weights section, several machines are also not usable. Moreover, you overhear three guests talking about the pool currently being out of order for maintenance. On your third and fourth day of your stay, most of the fitness machines and the pool are still unusable. On the day of your departure, you go to the reception to pay the bill. When the hotel employee asks if you were satisfied with your stay, you complain about the limited availability of the fitness center and the pool. You point out that, after all, your total price of €450 included the fee for using these facilities. The employee apologizes for the problems and calls the hotel manager. The manager tells you:	
Compensation		

level:		
€0 (0%)	"I am very sorry that you were not fully satisfied with your stay. Unfortunately, our company cannot offer any of the affected customers a financial compensation. I hope that despite these problems, we can welcome you as our guest in the future."	"I am very sorry that you were not fully satisfied with your stay. Unfortunately, we can't offer you a personal financial compensation. I hope that despite these problems, we can welcome you as our guest in the future."
€75 (17%) €150 (33%) €225 (50%) €300 (67%) €375 (83%) €450 (100%)	"I am very sorry that you were not fully satisfied with your stay. Our company wants to offer *all affected customers a compensation* of *[€75/€150/€225/€300/€375/€450]*. We will deduct this amount from your hotel bill of €450, so that you [only have to pay €375/€300/€225/€150/€75 / do not have to pay anything] for your stay. I hope that despite these problems, we can welcome you as our guest in the future."	"I am very sorry that you were not fully satisfied with your stay. Our company wants to offer *you personally a compensation* of *[€75/€150/€225/€300/€375/€450]*. We will deduct this amount from your hotel bill of €450, so that you [only have to pay €375/€300/€225/€150/€75 / do not have to pay anything] for your stay. I hope that despite these problems, we can welcome you as our guest in the future."

Appendix B. Items and Reliability Measures (Study 1 / Study 2)

	Cronbach's alpha	Construct reliability	Average variance extracted	Factor loading[d]	Indicator reliability
Recovery satisfaction[a] (Gelbrich, Gäthke and Grégoire 2015)	.946 / .978	.946 / .979	.854 / .938		
1. In my opinion, the hotel provided a satisfactory resolution to the problem on this particular occasion.				.924 / .970	.854 / .940
2. I am satisfied with the hotel's handling of the problem.				.949 / .965	.901 / .931
3. I am satisfied with the hotel's response to the problem.				.898 / .971	.806 / .944
Distributive justice[b] (Smith, Bolton and Wagner 1999)	- / .931	- / .927	- / .761		
1. The outcome I received was fair.				- / .969	- / .940
2. I did not get what I deserved. (r)				- / .775	- / .601
3. In resolving the problem, the hotel gave me what I needed.				- / .893	- / .798
4. The outcome I received was not right. (r)				- / .840	- / .705
Failure severity[c] (Maxham and Netemeyer 2002)	- / .932	- / .933	- / .842		
In my opinion, the described situation represents…					
1. …a minor problem. / …a major problem.				- / .885	- / .784
2. …a small inconvenience. / …a big inconvenience.				- / .958	- / .918
3. …a minor aggravation / …a major aggravation.				- / .877	- / .769
CFA model fit Study 2: χ^2 (31) = 75.506; χ^2/df = 2.436; RMSEA = .050 (90% C.I.: .036 .064); SRMR = .012; CFI = .993; NNFI = .991					

Note: a: an eleven-point Likert scale anchored by "totally disagree" (1) and "totally agree" (11). b: a seven-point Likert scale anchored by "totally disagree" (1) and "totally agree" (7). c: a seven-point semantic differential scale. d: standardized factor loadings.

References

Albrecht, Arne K., Gianfranco Walsh, and Sharon E. Beatty (2017), "Perceptions of Group Versus Individual Service Failures and Their Effects on Customer Outcomes: The Role of Attributions and Customer Entitlement," *Journal of Service Research*, 20 (2), 188-203.

Argo, Jennifer J., Darren W. Dahl, and Rajesh V. Manchanda (2005), "The Influence of a Mere Social Presence in a Retail Context," *Journal of Consumer Research*, 32 (2), 207–212.

Aula, Pekka (2010), "Social Media, Reputation Risk and Ambient Publicity Management," *Strategy & Leadership*, 38 (6), 43–49.

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 (4), 565–577.

Cambra-Fierro, Jesus, Iguacel Melero, and F. Javier Sese (2015), "Managing Complaints to Improve Customer Profitability," *Journal of Retailing*, 91 (1), 109–124.

Choi, Sunmee, and Anna S. Mattila (2008), "Perceived Controllability and Service Expectations: Influences on Customer Reactions Following Service Failure," *Journal of Business Research*, 61 (1), 24-30.

Collie, Therese, Graham Bradley, and Beverley A. Sparks (2002), "Fair Process Revisited: Differential Effects of Interactional and Procedural Justice in the Presence of Social Comparison Information," *Journal of Experimental Social Psychology*, 38 (6), 545–555.

Deighton, John, Caroline M. Henderson, and Scott A. Neslin (1994), "The Effects of Advertising on Brand Switching and Repeat Purchasing," *Journal of Marketing Research*, 31 (1), 28–43.

De Ruyter, Ko, and Martin Wetzels (2000), "Customer Equity Considerations in Service Recovery: A Cross-Industry Perspective," *International Journal of Service Industry Management*, 11 (1), 91-108.

DiLalla, David L. and Stephen J. Dollinger (2006), "Cleaning up Data and Running Preliminary Analyses," in *The Psychology Research Handbook*, Frederick T.L. Leong and James T. Austin, eds. Thousand Oaks, CA: SAGE, 241-254.

Du, Jiangang, Xiucheng Fan, and Tianjun Feng (2014), "Group Emotional Contagion and Complaint Intentions in Group Service Failure: The Role of Group Size and Group Familiarity," *Journal of Service Research*, 17 (3), 326–338.

Estelami, Hooman (2000), "Competitive and Procedural Determinants of Delight and Disappointment in Consumer Complaint Outcomes," *Journal of Service Research*, 2 (3), 285-300.

Festinger, Leon (1954), "A Theory of Social Comparison Processes," *Human Relations*, 7 (2), 117–140.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (1), 39-50.

Gelbrich, Katja (2010), "Anger, Frustration, and Helplessness after Service Failure: Coping Strategies and Effective Informational Support," *Journal of the Academy of Marketing Science*, 38 (5), 567–585.

———, Jana Gäthke, and Yany Grégoire (2015), "How Much Compensation Should a Firm Offer for a Flawed Service? An Examination of the Nonlinear Effects of Compensation on Satisfaction," *Journal of Service Research*, 18 (1), 107–123.

———, ———, and ——— (2016), “How a Firm’s Best Versus Normal Customers React to Compensation After a Service Failure,” *Journal of Business Research*, 69 (10), 4331–4339.

——— and Holger Roschk (2011), “A Meta-Analysis of Organizational Complaint Handling and Customer Responses,” *Journal of Service Research*, 14 (1), 24–43.

Gerber, J. P., Ladd Wheeler, and Jerry Suls (2018), “A Social Comparison Theory Meta-Analysis 60+ Years On,” *Psychological Bulletin*, 144 (2), 177-197.

Göritz, Anja S. (2004), “Recruitment for Online Access Panels,” *International Journal of Market Research*, 46 (4), 411-425.

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.

Hamilton, Rebecca W., Roland T. Rust, Michel Wedel, and Chekitan S. Dev (2017), “Return on Service Amenities,” *Journal of Marketing Research*, 54 (1), 96-110.

Hess, Ronald L., JR., 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 (2), 127–145.

Hogreve, Jens, Nicola Bilstein and Leonhard Mandl (2017), “Unveiling the Recovery Time Zone of Tolerance: When Time Matters in Service Recovery,” *Journal of the Academy of Marketing Science*, 45 (6), 866-883.

Homburg, Christian, Michael Müller, and Martin Klarmann (2011), “When Should the Customer Really Be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters,” *Journal of Marketing*, 75 (2), 55–74.

Johnston, Robert (1995), "The Zone of Tolerance: Exploring the Relationship between Service Transactions and Satisfaction with the Overall Service," *International Journal of Service Industry Management*, 6 (2), 46-61.

Lindell, Michael K. and David J. Whitney (2001), "Accounting for Common Method Variance in Cross-Sectional Research Designs," *Journal of Applied Psychology*, 86 (1), 114-121.

Mattila, Anna S. (2001), "The Effectiveness of Service Recovery in a Multi-Industry Setting," *Journal of Services Marketing*, 15 (7), 583-596.

——— and Paul G. Patterson (2004), "The Impact of Culture on Consumers' Perceptions of Service Recovery Efforts," *Journal of Retailing*, 80 (3), 196–206.

Maxham, James G. and Richard G. Netemeyer (2002), "A Longitudinal Study of Complaining Customers' Evaluations of Multiple Service Failures and Recovery Efforts," *Journal of Marketing*, 66 (4), 57-71.

McColl-Kennedy, Janet R., Catherine S. Daus, and Beverley A. Sparks (2003), "The Role of Gender in Reactions to Service Failure and Recovery," *Journal of Service Research*, 6 (1), 66–82.

Muthén, Linda K. and Bengt O. Muthén (1998-2017), *Mplus User's Guide*, 8th ed., Los Angeles, CA: Muthén and Muthén.

Nam, Sungjoon, Puneet Manchanda, and Pradeep K. Chintagunta (2010), "The Effect of Signal Quality and Contiguous Word of Mouth on Customer Acquisition for a Video-on-Demand Service," *Marketing Science*, 29 (4), 690–700.

Oliver, Richard L. and John E. Swan (1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach," *Journal of Marketing*, 53 (2), 21–35.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee and Nathan P. Podsakoff (2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (5), 879-903.

Raftery, Adrian E. (1995), "Bayesian Model Selection in Social Research [with discussion]," in *Sociological Methodology 1995*, Peter V. Marsden, ed. Cambridge, MA: Blackwell, 111-195.

Roschk, Holger and Katja Gelbrich (2017), "Compensation Revisited," *Journal of Service Research*, 20 (4), 393–408.

Ryu, Ehri and Jeewon Cheong (2017), "Comparing Indirect Effects in Different Groups in Single-Group and Multi-Group Structural Equation Models," *Frontiers in Psychology*, 8, 747, available at https://doi.org/10.3389/fpsyg.2017.00747.

Seta, John J. and Catherine E. Seta (1992), "Personal Equity-Comparison Theory: An Analysis of Value and the Generation of Compensatory and Noncompensatory Expectancies," *Basic and Applied Social Psychology*, 13 (1), 47–66.

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 (3), 356–372.

Strickhouser, Jason E. and Ethan Zell (2015), "Self-Evaluative Effects of Dimensional and Social Comparison," *Journal of Experimental Social Psychology*, 59, 60–66.

Tax, Stephen S., Stephen W. Brown, and Murali Chandrashekaran (1998), "Customer Evaluations of Service Complaint Experiences: Implications for Relationship Marketing," *Journal of Marketing*, 62 (2), 60–76.

van den Bos, Kees, E. A. Lind, Riel Vermunt, and Henk A. M. Wilke (1997), "How do I Judge my Outcome When I Do not Know the Outcome of Others? The Psychology of the Fair Process Effect," *Journal of Personality and Social Psychology*, 72 (5), 1034–1046.

Zhou, Yuanyuan, Minxue Huang, Alex S. Tsang, and Nan Zhou (2013), "Recovery Strategy for Group Service Failures: The Interaction Effects Between Recovery Modes and Recovery Dimensions," *European Journal of Marketing*, 47 (8), 1133–1156.

———, Alex S. L. Tsang, Minxue Huang, and Nan Zhou (2014), "Group Service Recovery Strategies Effectiveness: The Moderating Effects of Group Size and Relational Distance," *Journal of Business Research*, 67 (11), 2480–2485.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21 (1), 1-12.

Table 1. Study 1: SEM Results for Recovery Satisfaction

	A: Overall estimation	B: Multi-group invariant	C: Multi-group partly unconstrained	
Antecedents			GSR	ISR
Compensation size	−.051**	–.051***	–.093***	.000
Compensation size$^{0.5}$	1.245***	1.245**	1.527***	.908***
Recovery form	.361†	N/A	N/A	N/A
R^2	.543	.542	.534	.569
Adjusted R^2	.540	.740	.529	.565
Model fit				
χ^2 (df)	11.234 (6)	61.306 (19)	50.864 (16)	
χ^2 / df	1.872	3.227	3.179	
CFI	.996	.969	.974	
TLI	.992	.971	.971	
RMSEA (90% CI)	.047 (.000 .089)	.107 (.077 .137)	.105 (.074 .139)	
SRMR	.009	.090	.060	
BIC	4,788.80	4,750.77	4,748.73	

Note. Recovery form dummy-coded as GSR = 1 and ISR = 0; Unstandardized path coefficients.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Table 2. Study 2: SEM Results for Recovery Satisfaction

	A: Overall estimation	B: Multi-group invariant	C: Multi-group partly unconstrained	
Antecedents			GSR	ISR
Compensation size	−.018***	–.018***	–.023***	–.012***
Compensation size$^{0.5}$	.801***	.801***	.881***	.719***
Recovery form	.193	N/A	N/A	N/A
Covariates				
Age	.000	.000	.000	
Gender	–.223	–.215	–.252†	
Failure severity	−.234***	–.233***	–.221***	
Failure relevance	.068	.066	.055	
Complaint experience	−.187**	–.194**	–.190**	
Recovery experience	.138**	.137**	.126*	
R^2	.766	.766	.759	.788
Adjusted R^2	.762	.762	.752	.782
Model fit				
χ^2 (df)	45.771 (18)	108.524 (49)	87.879 (46)	
χ^2 / df	2.542	2.215	1.910	
CFI	.992	.983	.988	
TLI	.987	.981	.986	
RMSEA (90% CI)	.052 (.033 .070)	.065 (.048 .081)	.056 (.038 .074)	
SRMR	.004	.030	.013	
BIC	6,048.35	5,989.82	5,978.74	

Note. Recovery form dummy-coded as GSR = 1 and ISR = 0; Unstandardized path coefficients.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

Table 3. Study 2: SEM Results for Distributive Justice

	A: Overall estimation	B: Multi-group invariant	C: Multi-group partly unconstrained	
Antecedents			GSR	ISR
Compensation size	–.011***	–.011***	–.012***	–.009***
Compensation size$^{0.5}$	.460***	.460***	.474***	.446***
Recovery form	.066	N/A	N/A	N/A
Covariates				
Age	.000	.000	.000	
Gender	–.008	–.008	–.024	
Failure severity	–.109**	–.109**	–.106**	
Failure relevance	–.012	–.012	–.018	
Complaint experience	–.065	–.065	–.065†	
Recovery experience	.076*	.076*	.069*	
R^2	.710	.710	.685	.742
Adjusted R^2	.705	.705	.676	.735
Model fit				
χ^2 (df)	89.708 (29)	151.609 (71)	139.976 (69)	
χ^2 / df	3.093	2.135	2.029	
CFI	.977	.970	.974	
TLI	.967	.968	.971	
RMSEA (90% CI)	.060 (.046 .074)	.063 (.049 .076)	.060 (.045 .074)	
SRMR	.014	.036	.026	
BIC	7,318.17	7,251.50	7,246.25	

Note. Recovery form dummy-coded as GSR = 1 and ISR = 0; Unstandardized path coefficients.
*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

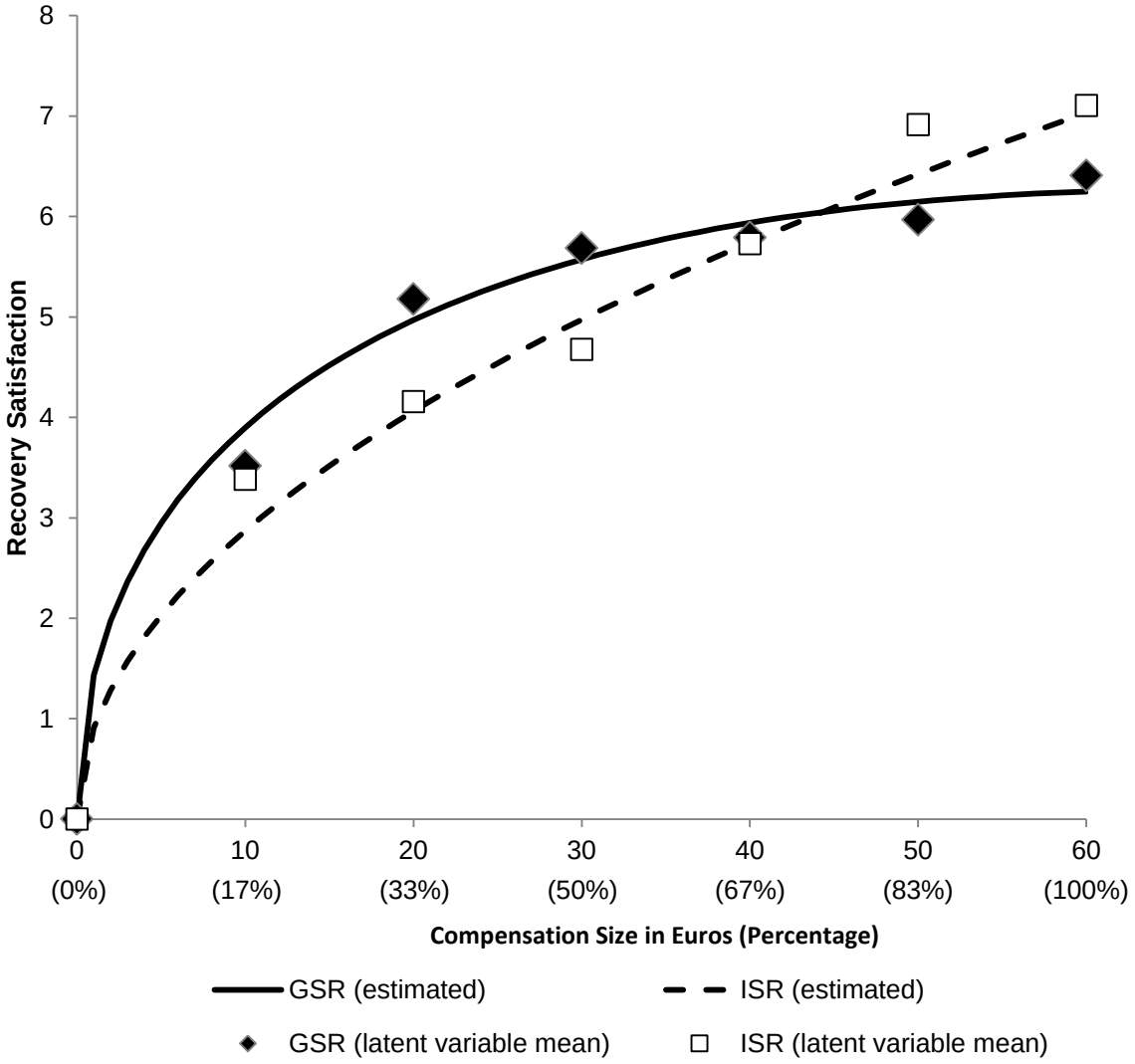


Figure 1. Study 1: Estimated Curve Progressions and Latent Variable Means for Recovery Satisfaction by Recovery Form

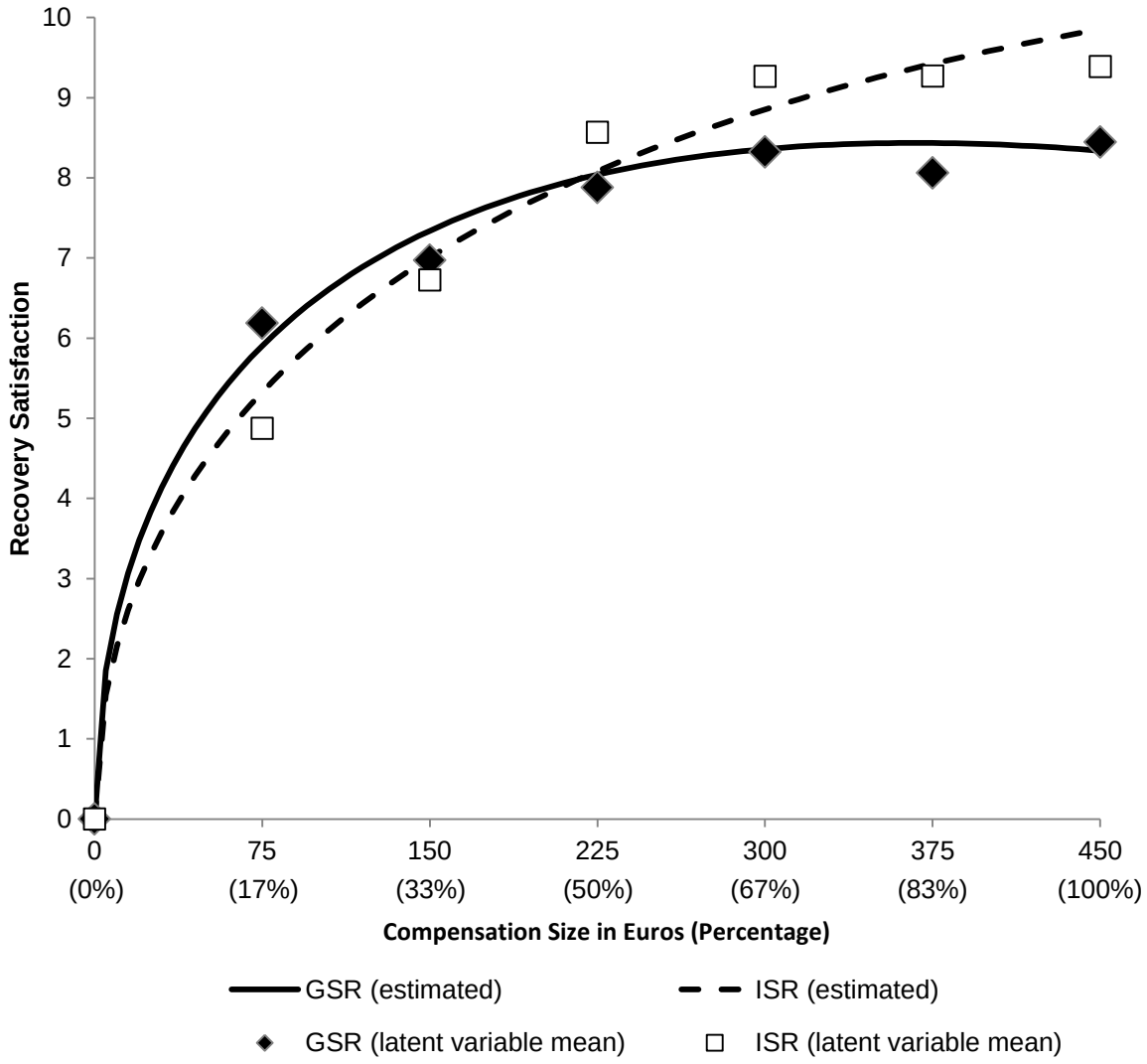


Figure 2. Study 2: Estimated Curve Progressions and Observed Latent Variable Means for Recovery Satisfaction by Recovery Form

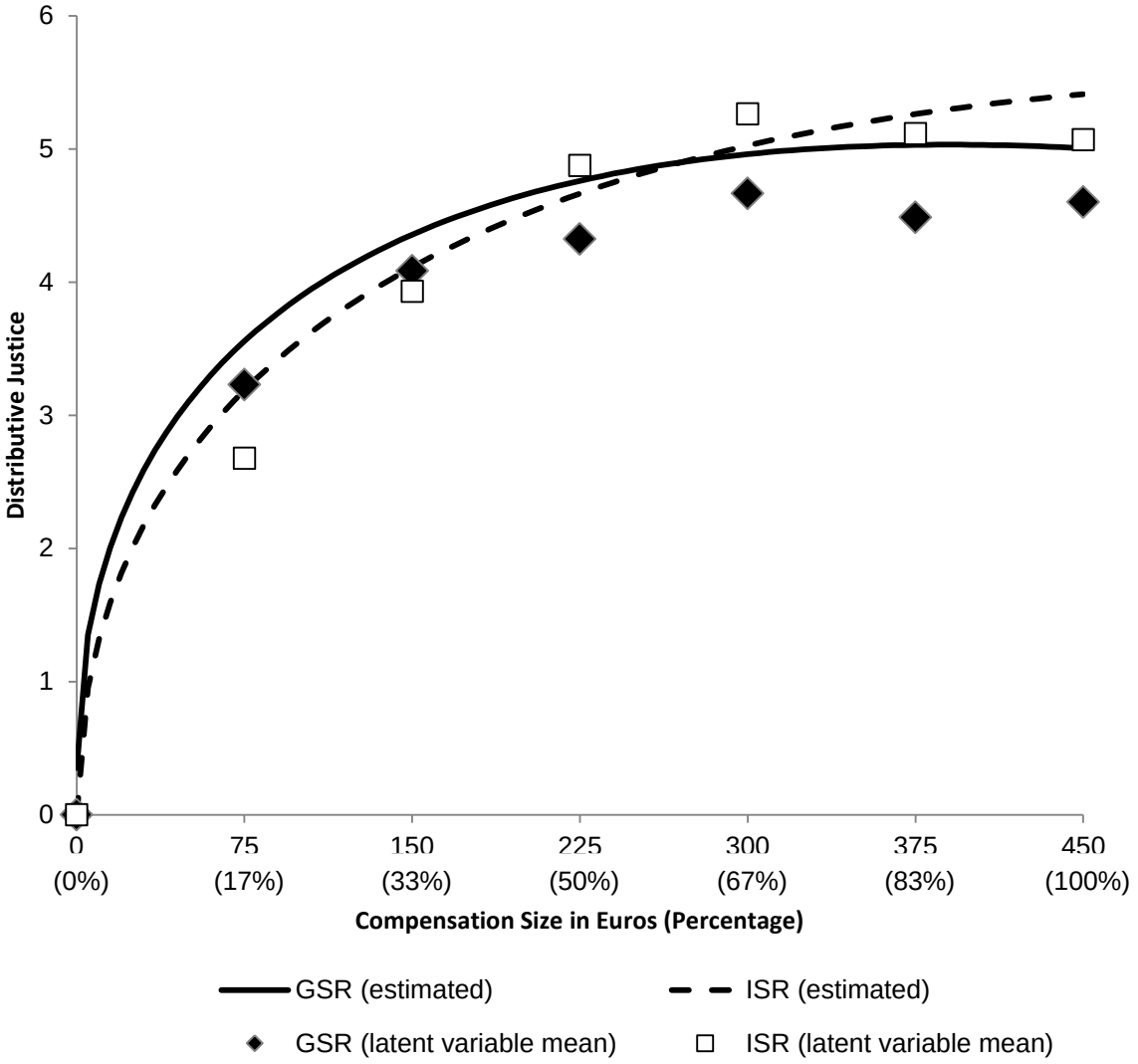
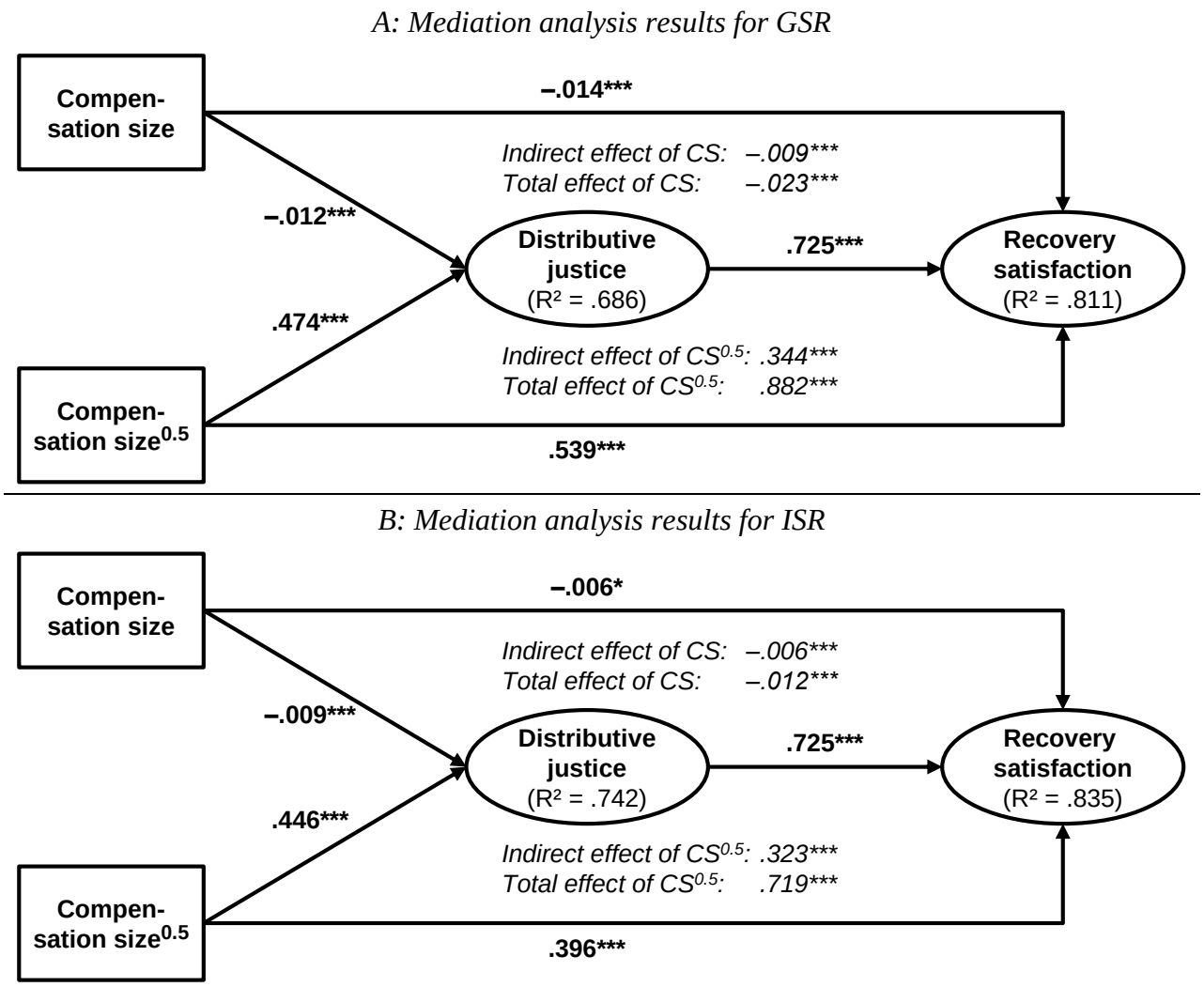


Figure 3. Study 2: Estimated Curve Progressions and Observed Latent Variable Means for Distributive Justice by Recovery Form



Model fit: χ^2 (138) = 291.234, χ^2/df = 2.110, CFI = .976, TLI = .973, RMSEA = .062 [90% CI: .052 .072], SRMR = .026, BIC = 13,118.28

Figure 4. Study 2: Moderated Mediation Analysis Results

Note: Covariates included in the estimation were age, gender, failure severity, failure relevance, complaint experience, and recovery experience. Estimation based on bias-corrected bootstrapping, using 10,000 bootstrap samples. *** $p < .001$