Received: 5 March 2019 | Revised: 24 May 2020 | Accepted: 10 June 2020 | Published on: 14 July 2020

DOI: 10.1002/smj.3213

RESEARCH ARTICLE

STRATEGIC MANAGEMENT JOURNAL WILEY

The crisis-response match: An empirical investigation

Sascha Raithel[1] | **Stefan J. Hock**[2]

[1]Freie Universität Berlin, Berlin, Germany

[2]University of Connecticut, Storrs, Connecticut, USA

Correspondence
Sascha Raithel, Freie Universität Berlin, Arnimallee 11, 14195 Berlin, Germany.
Email: sascha.raithel@fu-berlin.de

Abstract

Research Summary: Prior crisis-response literature outlines zones of conformity (i.e., response meets stakeholder expectations), underconformity (i.e., response falls short of expectations), and overconformity (i.e., response exceeds expectations). We utilize a mixed-method approach to empirically test the impact of different response strategies on customers (Study 1: experiment) and investors (Study 2: event study). We not only find empirical support that a conforming strategy outperforms both nonconforming strategies concerning stakeholders' affective evaluations of reputation, but extend this proposition to stakeholders' cognitive evaluations of reputation and the financial implications for the firm. The most counterintuitive finding is that overconforming strategies result in lower firm reputation and stock returns relative to conforming strategies. Thus, exceeding stakeholder expectations during a crisis can have unintended negative consequences.

Managerial Summary: How should firms react to product recalls? Previous research suggests that exceeding expectations of external stakeholders should have a neutral or even positive impact on firm reputation and financial performance, while falling short of expectations should have a negative impact. In this article, we test the impact of different product recall strategies.

Sascha Raithel and Stefan J. Hock contributed equally to this work.

The most counterintuitive finding is that exceeding stakeholder expectations during a product crisis can have unintended negative consequences on both customers and investors.

KEYWORDS

conformity, crisis, firm response, reputation, product recall, stock return

1 | INTRODUCTION

A crisis is an "unexpected, publicly known, and harmful event that has high levels of initial uncertainty, interferes with the normal operations of an organization, and generates widespread, intuitive, and negative perceptions among evaluators" (Bundy & Pfarrer, 2015: p. 345). A firm's crisis-response strategy that meets or exceeds expectations of external stakeholders should have a neutral or even positive impact on their evaluations of the firm, while falling short of expectations should have a negative impact (Fombrun & Shanley, 1990; Gray & Balmer, 1998; Philippe & Durand, 2011). In their conceptual piece, Bundy and Pfarrer (2015; hereafter: B&P), based on Coombs' situational crisis communication theory (e.g., Coombs, 2007), categorize potential crisis-response strategies into three general groups: underconforming (i.e., response falls short of expectations), conforming (i.e., response meets expectations), and overconforming (i.e., response exceeds expectations). B&P argue that social approval loss is lowest for conforming strategies. While it is intuitive that conforming (underconforming) has a neutral or positive (negative) impact on external stakeholders' evaluations of the firm, B&P develop the counterintuitive idea that overconforming could create negative consequences for firms.

To date, there is no empirical evidence for B&P's main proposition and how these strategies shape evaluators' perceptions of the firm. We close this gap and offer the following two extensions. First, we utilize a two-component conceptualization of reputation (Raithel & Schwaiger, 2015) that captures both the intuitive and affective aspects (as discussed by B&P) as well as the deliberate and analytical aspects of reputation. Second, in addition to firm reputation, we also utilize stock returns to understand the financial impact of B&P's main proposition.

2 | THE CRISIS-RESPONSE MATCH THEORY

According to B&P's framework, stakeholders evaluate firm crises based on two primary factors: situational attributions (low vs. high) and a firm's response strategy (more defensive vs. more accommodative). *Situational attributions* refer to "the perceived degree of an organization's responsibility based on the characteristics of the crisis" (p. 346), such as "its perceived intentionality, controllability, and severity" (p. 351), while a firm's *response strategy* refers to "the set of coordinated communication and actions used to influence evaluators' crisis perceptions" (p. 346). Based on these two factors, B&P derive their proposed crisis-response match, and argue that in order to reduce social approval loss, "a crisis with higher situational attributions of responsibility should be matched with a response strategy that accepts more responsibility, and

a crisis with lower situational attributions of responsibility should be matched with a response strategy that accepts less responsibility" (p. 352).

Conforming strategies match higher (lower) situational attributions of responsibility with a response strategy that accepts more (less) responsibility. As B&P argue, both conforming strategies are more likely to create cognitive consonance, and external stakeholders are not only more likely to agree with a firm's response, but also less likely to alter their initial judgments about the firm (Nickerson, 1998; Traut-Mattausch, Shulz-Hardt, Greitemeyer, & Frey, 2004; Tversky & Kahneman, 1974). Further, conforming strategies are more likely to satisfy external stakeholders' normative expectations. For crises with higher situational attributions and perceived responsibility, a firm's acceptance of more responsibility satisfies social expectations of justice, sincerity, and fairness (Coombs, 2007; Dean, 2004; Pfarrer, DeCelles, Smith, & Taylor, 2008; Tyler, 1997) and can support perceptions of competence and likeability (Lee, Peterson, & Tiedens, 2004). For crises with lower situational attributions and perceived responsibility, evaluators typically have lower normative expectations of a firm. Thus, a strategy that accepts less responsibility is appropriate for such crises.

Nonconforming strategies either underconform by accepting less responsibility than evaluators expected, or overconform by accepting more. Both nonconforming strategies, as B&P argue, challenge stakeholders' initial perceptions of a crisis and trigger cognitive dissonance (Festinger, 1962; Pratt, 2000), which in turn increases information search (Ashforth, Harrison, & Corley, 2008; Elliot & Devine, 1994) that can lead to the conclusion that a firm's response was careless, inaccurate, wrong, or even deceitful (Benoit, 1995; Dean, 2004; Elsbach, 2003; Pfarrer et al., 2008). Further, "evaluators' attribution biases (i.e., confirmation bias and anchoring bias) can lead them to resist an underconforming or overconforming mismatched strategy" (B&P, 2015: p. 356). While it is not surprising that an *underconforming* strategy can cause negative judgments, it seems counterintuitive that an *overconforming* strategy can do the same. Although exceeding expectations is usually a positive signal, external stakeholders might become suspicious and wonder why a firm is accepting more crisis responsibility than it needs to. Is the potential damage/hazard much bigger than expected? The anchoring bias (Tversky & Kahneman, 1974) ensures that lower initial attributions are not easily changed by the overconforming response. Instead, it raises other concerns, such as is the firm insincere, overreaching, or hiding anything? As B&P argue, looking for such counterfactuals to reduce/eliminate the dissonance might result in more negative judgments by evaluators (Ashforth & Gibbs, 1990).

3 | EXTENDING THE CRISIS-RESPONSE MATCH THEORY

First, B&P only theorize about the effects of the crisis-response match on evaluators' perceptions of social approval. They define social approval as an intuitive and affective construct, which captures the "perception of general affinity toward an organization" (p. 347). B&P argue that social approval is different from perceptions of organizational reputation, which they define as a deliberate and analytical construct that captures the "assessment of an organization's ability to deliver value" (p. 347). B&P's conceptualizations of social approval and reputation do, however, reflect the two-component conceptualization of corporate reputation (e.g., Raithel & Schwaiger, 2015) that captures external evaluators' perceptions of a firm's abilities ("competence") and their affinity toward the firm ("likeability"). B&P acknowledge this two-component view on reputation (p. 347) and the inherent overlap between their

conceptualization of social approval and reputation (p. 348: footnote 2). Therefore, we extend B&P's main proposition by applying it to the two-component conceptualization of reputation.

Second, B&P acknowledge the missing focus on the financial implications of the crisis-response match. On the one hand, financial performance is directly related to firm reputation (e.g., Raithel & Schwaiger, 2015; Robert & Dowling, 1997). Hence, conforming responses could outperform nonconforming responses financially. On the other hand, different response strategies involve different financial costs for the firm (Liu, Liu, & Luo, 2016), and investors might evaluate the financial net effect of the crisis-response match differently. Therefore, in addition to examining the impact of B&P's main proposition on firm reputation, we also analyze stock returns to gain insights into the financial consequences of the different response strategies.

We utilize a mixed-method approach to empirically test B&P's crisis-response match and the two proposed extensions. Study 1 (experiment) manipulates situational attributions and firm response in a controlled environment and examines customer response by using a two-component view on reputation as the dependent measure. In Study 2 (event study) we draw on observational data and analyze investor response by using abnormal returns as the dependent measure. This two-study design allows us to assess the generalizability of B&P's prediction to different stakeholder groups and performance outcomes that cannot be addressed through one data source alone (Davis, Golicic, & Boerstler, 2011; Hamilton, 2016).

4 | RESEARCH CONTEXT: PRODUCT RECALLS

We use a product recall context to test B&P's conceptual framework empirically. First, product recalls are one of the most frequent firm crises and can even threaten a firm's existence (e.g., Japanese airbag manufacturer Takata went bankrupt after a major recall; CNN, 2017). Second, and most importantly, for many industries product recalls are regulated by the Consumer Product Safety Commission (CPSC), and therefore offer standardized data, which enable the comparison of empirical results from the field with those of experiments. In short, CPSC's basic product recall process is as follows. CPSC releases a standardized recall announcement together with the affected firm if CPSC, the firm, a consumer, or any other supply chain member identifies a significant product hazard.[1] The primary two objectives of any recall are to (a) locate and remove defective products as quickly as possible and (b) "communicate accurate and understandable information in a timely manner to the public about the product defect, the hazard, and the corrective action" (CPSC, 2012: p. 18).

4.1 | Study 1: The crisis-response match and customer response

4.1.1 | Participants, method, and design

Five hundred and sixty-nine U.S. based adults (M_{age} = 37.34, 61% female) completed this survey for a small payment through the TurkPrime application (Litman, Robinson, & Abberbock, 2017). This study employed a two (situational attributions: low = external cause vs. high = internal cause) × 2 (firm response: accept vs. deny responsibility) between subjects design with random assignment, plus a control group. Participants read the recall of a

[1]For a detailed explanation of CPSC's recall process, see Chen, Ganesan, and Liu (2009).

smartphone producer, which we tailored to their specific smartphone brand to increase involvement. After reading about the recall, we first manipulated situational attributions. In the low [high] situational attributions condition, participants read a brief newspaper excerpt in which the firm blamed their supplier [internal testing], and we assessed participants' situational attributions (I believe that [X] had the power to control the product failure/[X] intentionally sold the malfunctioning product/the product malfunction is a severe event; $\alpha = .481$).[2] After that, we manipulated firm response. In the accept [deny] responsibility condition, participants read another brief newspaper excerpt in which the firm accepted full [denied any] responsibility. In the control condition, participants only read the initial recall announcement, see Web Appendix A for more study details.

Next, participants rated their brand's post-recall reputation, which served as dependent variable. We used a measure from Raithel and Schwaiger (2015), which consists of six 7-point Likert scales and captures both people's perceptions of a firm's abilities ("competence") and their feelings about the firm ("likeability"). Sample items include "[X] is a top competitor in its market" (competence) and "I regard [X] as a likeable company" (likeability). The extracted factor shows a good fit (AVE = 0.657, $\alpha = .880$). Lastly, participants indicated how old their phone was, how much it cost, and provided demographics. We thanked and debriefed participants to prevent carry-over effects from the experiment to actual brand perceptions.

In order to ensure high data quality, we included three manipulation checks and participants were immediately brought to the end of the survey if they answered any of them incorrectly. First, after the recall announcement, we asked what the product hazard was (overheating only, burn hazard, burn and fire hazard, choking, laceration). Second, after the situational attributions manipulation, we asked who the firm blamed for the hazardous batteries (internal testing, supplier, article did not specify). Third, after the firm response manipulation, we asked if the firm accepted responsibility for the recall (yes, no, article did not specify).

4.1.2 | Results

Equivalence of groups

ANOVAs revealed that the five groups did not differ with regard to the age ($p = .791$) or price of the phone ($p = .648$), or participants' age ($p = .983$) or gender ($p = .577$), suggesting that the five groups ($N_{internal \times accept} = 100$, $N_{internal \times deny} = 91$, $N_{external \times accept} = 119$, $N_{external \times deny} = 111$, $N_{control} = 148$)[3] were by and large comparable.

Situational attributions

A successful manipulation satisfies two criteria: First, both internal conditions should yield higher situational attributions than both external conditions. Second, both internal conditions should not differ from each other and both external conditions should not differ from each other either. This is exactly what we find. Most importantly, both internal conditions (internal × accept: 0.134; internal × deny: 0.137) yield higher situational attributions than both external conditions (external × accept: −0.128; external × deny: −0.128). The effect difference between the two internal and the two external conditions is significant ($F[1, 419] = 7.238$,

[2]The lower alpha can be explained by the fact that the first item is affected by the situational attribution manipulation, but the other two items are not. Our effects get even stronger when we only include the first item.

[3]Group sizes differ because of the three manipulation checks described earlier.

$p = .006$, $d = 0.269$), while there is no difference between the two internal ($F[1, 189] = 0.001$, $p = .982$, $d = 0.003$) and the two external conditions ($F[1, 228] = 0.000$, $p = .998$, $d = 0.001$). Therefore, our manipulation of situational attributions was successful. In order to provide a complete picture of all conditions, we also compared the control condition with the two internal and external conditions. The effect difference between the two internal conditions and the control condition (−0.068) is marginally significant ($F[1, 337] = 3.714$, $p = .055$, $d = 0.209$), while there is no difference between the two external conditions and the control condition ($F[1, 376] = 0.314$, $p = .576$, $d = 0.060$). Next, we discuss how firm responses (accept vs. deny) impact perceptions of post-recall reputation once these situational attributions are formed.

Reputation (aggregate)

A successful test of B&P's main proposition entails two main findings. First, both conforming conditions should yield higher situational attributions than both nonconforming conditions. Second, both conforming conditions should not differ from each other and both nonconforming conditions should not differ from each other either. This is exactly what we find. Figure 1a displays the group means of the *z*-standardized post-recall reputation factor score for the five different groups. Most importantly, both conforming conditions (internal × accept (C1): 0.328; external × deny (C2): 0.157) yield a higher post-recall reputation than the two nonconforming conditions (internal × deny (UC): −0.290; external × accept (OC): −0.154). The effect difference between the two conforming and the two nonconforming strategies is significant ($F[1, 419] = 24.859$, $p = .000$, $d = 0.486$), while there is no difference between the two conforming ($F[1, 209] = 2.200$, $p = .140$, $d = 0.204$) and the two nonconforming conditions ($F[1, 208] = 0.918$, $p = .339$, $d = 0.132$). Therefore, we find support for B&P's main proposition. In order to provide a complete picture of all conditions, we also compared the control condition with the two conforming and nonconforming conditions. The effect difference between the two conforming conditions and the control condition (0.049; ($F[1, 357] = 3.955$, $p = .047$, $d = 0.211$), and for the two nonconforming conditions and the control condition is significant ($F[1, 356] = 6.140$, $p = .014$, $d = 0.266$).

Competence

These reputation effects replicate for the competence factor. Figure 1b (black bars) displays the group means of the *z*-standardized factor score. Most importantly, both conforming conditions (internal × accept (C1): 0.237; external × deny (C2): 0.207) yield a higher post-recall competence than nonconforming conditions (internal × deny (UC): −0.204; external × accept (OC): −0.125). The effect difference between the two conforming and the two nonconforming strategies is significant ($F[1, 419] = 17.373$, $p = .000$, $d = 0.407$), while there is no significant difference between the two conforming ($F[1, 209] = 0.063$, $p = .802$, $d = 0.036$) and the two nonconforming conditions ($F[1, 208] = 0.303$, $p = .583$, $d = 0.176$). The effect difference between the two conforming conditions and the control condition is significant (−0.001; ($F[1, 357] = 5.185$, $p = .023$, $d = 0.238$), while there is no difference between the two nonconforming conditions and the control condition ($F[1, 356] = 2.108$, $p = .147$, $d = 0.158$).

Likeability

These reputation effects also replicate for the likeability factor (see gray bars in Figure 1b). Most importantly, both conforming conditions (internal × accept (C1): 0.349; external × deny (C2): 0.093) yield a higher post-recall likeability than the two nonconforming conditions (internal × deny (UC): −0.312; external × accept (OC): −0.153). The effect difference between the

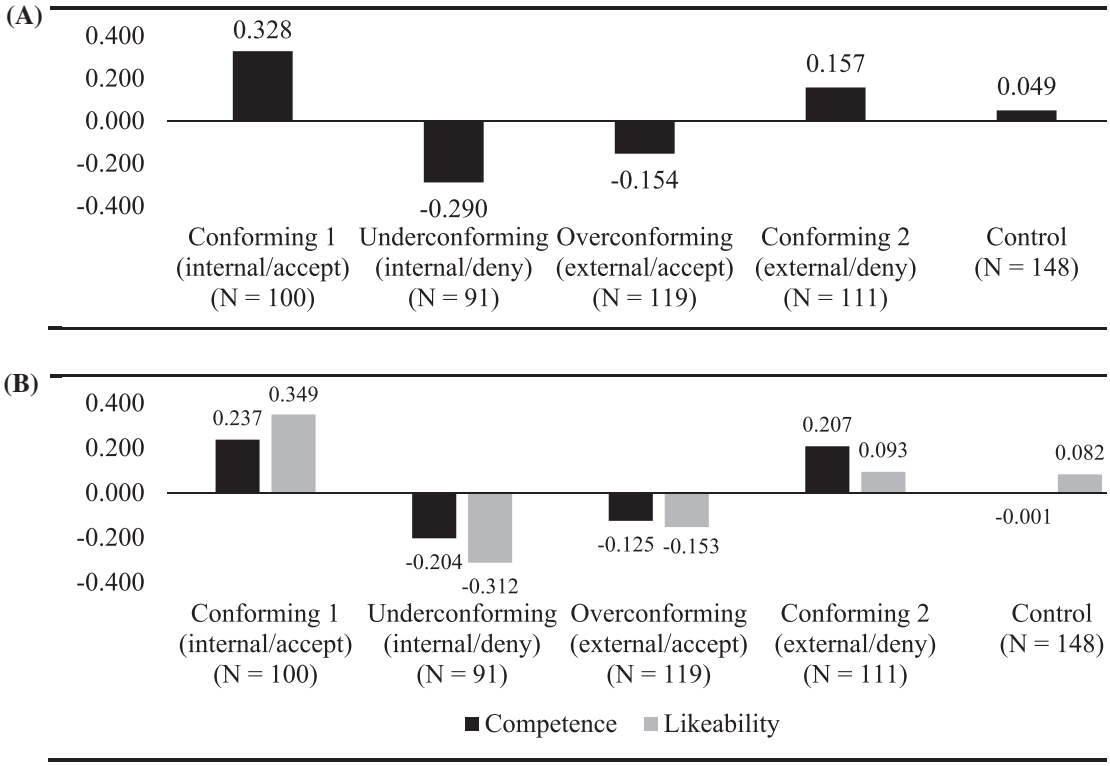

FIGURE 1 (a) Impact of crisis-response match on aggregate reputation measure (Study 1). *Note*: Reputation is measured as *z*-standardized factor score of the six reputation items. Scores can only be interpreted relative to each other but not in absolute terms. (b) Impact of crisis-response match on competence and likeability (Study 1). *Note*: Competence [Likeability] is measured as *z*-standardized factor score of the three competence [likeability] items. Scores can only be interpreted relative to each other but not in absolute terms

two conforming and the two nonconforming conditions is significant (F[1, 419] = 22.734, p = .000, d = 0.461), and there is a significant difference between the two conforming conditions (F[1, 209] = 4.662, p = .032, d = 0.302), while there is no difference between the two nonconforming conditions (F[1, 208] = 1.286, p = .258, d = 0.158). The effect difference between the two conforming conditions and the control condition is not significant (0.082; (F[1, 357] = 1.931, p = .165, d = 0.145), while there is a difference between the two nonconforming conditions and the control condition (F[1, 356] = 8.520, p = .004, d = 0.301).

4.1.3 | Discussion

We find empirical support for B&P's main proposition: post-recall reputation is higher for conforming strategies than for under- and overconforming strategies. This finding holds for the affective as well as cognitive aspects of reputation. The goal of the next study is to extend B&P's proposition to a financial context by identifying the financial implications of the crisis-response match. In Study 2, we analyze the impact of CPSC product recalls on stock returns.

4.2 | Study 2: The crisis-response match and investor response

4.2.1 | Mapping the field data onto B&P's framework

The standardized CPSC data does not contain explicit information about situational attributions (i.e., internal vs. external cause) and whether firms accept or deny any responsibility. In Study 2 (event study), we therefore use the two implicit factors, timing (reactive vs. proactive) and remedy (partial vs. full), as proxies. Although these proxies do not fully reflect stakeholders' perceptions about situational attributions and firms' acceptance of responsibility, timing and remedy do map onto B&P's two core dimensions of their theoretical framework, and implicitly provide evaluators with important cues about the nature of and responsibility for the crisis.

First, Chen et al. (2009) classify recall strategies as either proactive or reactive. Proactive (reactive) recalls occur early (much later) in the process and are often triggered through internal inspections (external complaints) and before any (after at least one) consumer safety incident has been reported to the firm or CPSC. A proactive firm voluntarily discloses information about a product malfunction, approaches the authorities directly, and wants to remove/repair defective products as quickly as possible (p. 216). Thus, by avoiding consumer safety incidents, a proactive firm could reduce situational attributions about crisis severity and intentionality compared to a reactive firm, which has been either incompetent at discovering, or has intentionally concealed the product malfunction until consumer safety incidents became public. Consequently, and all else being equal, situational attributions and accompanying perceived responsibility should be lower (higher) for proactive (reactive) recalls.

Second, Liu et al. (2016) classify firm responses as either partial or full remedy, and firms can freely choose between the two. A firm's decision to offer partial or full remedy as compensation contributes to stakeholders' perception of the firm's stand on its crisis responsibility. "As crisis response strategies become more accommodative, (...) stakeholders perceive the organization as taking greater responsibility for the crisis (Coombs & Holladay, 2004, 2005)" (Coombs, 2007; p. 170). Partial remedies (e.g., discount for future purchase or do-it-yourself repair kits) signal a more defensive response and indicate that a firm accepts less responsibility, because they either do not fully compensate customers or customers are expected to fix the malfunction themselves. Full remedies (e.g., full refund or free repair/replacement) signal a more accommodative response, because customers either receive all of their money back or the firm provides them with a new/repaired product. Thus, full remedies take on more responsibility because the firm incurs a higher compensation cost (Liu et al., 2016: p. 81).

In our context, a *conforming* strategy could offer either (a) full remedy (accepts more responsibility; Coombs, 2007; Liu et al., 2016) for a reactive recall (higher situational attributions and perceived responsibility; Chen et al., 2009) or (b) partial remedy (accepts less responsibility; Coombs, 2007; Liu et al., 2016) for a proactive recall (lower situational attributions and perceived responsibility; Chen et al., 2009). *Nonconforming* strategies either underconform (partial/reactive) by accepting less responsibility than evaluators expected, or overconform (full/proactive) by accepting more responsibility (see Web Appendix B).

4.2.2 | Data sources and sample

We analyze CPSC product recalls from January 1996 to December 2014. Each recall announcement includes the exact recall date, product details, hazard, remedy, incidents and

injuries, number of units recalled, time frame during which the product has been sold, and price. CPSC does not allow any news releases before the official recall announcement. This enabled us to pinpoint the event date. Both the announcement and the firm's recall strategy are not anticipated by the public, which is the ideal setting for an event study. We obtained daily stock returns from the Center for Research in Security Prices (CRSP) for product recalls of publicly traded firms (NYSE or NASDAQ). Similar to Chen et al. (2009), we excluded all retailer recalls (e.g., Walmart) and focused only on manufacturer recalls, because CPSC regulations differ between these two groups. CPSC's product recall announcement serves as the event and the announcement date serves as the event day (Day 0). Our final sample consists of 443 CPSC product recalls of 112 different publicly traded firms from January 1996 to December 2014.[4]

4.2.3 | Measures

Abnormal stock returns

We apply the Fama and French Four-Factor Model (Carhart, 1997; Fama & French, 1993) and regress each firm's stock returns on the market, size, value, and momentum factor over the estimation period $t = -255$ to $t = -11$, relative to the event day $t = 0$:

$$R_{it} - R_{RFt} = \alpha_i + \beta_i \cdot RMRF_t + s_i \cdot SMB_t + h_i \cdot HML_t + u_i \cdot UMD_t + \varepsilon_{it}, \quad (1)$$

where R_{it} is the stock return i on day t. R_{RFt} is the risk-free rate of return on day t. $RMRF_t$ is the risk-free adjusted market return on day t. SMB_t is the difference between small and large stock returns on day t, HML_t between high and low book-to-market stock returns on day t, and UMD_t between stock returns with an upward and downward momentum factor. ε_{it} refers to the error term of stock i on day t, and α_i is the intercept. The four slope estimates β_i, s_i, h_i, and u_i measure the sensitivity of stock i's risk-free adjusted return to the four risk factors. We focus on the cumulative abnormal returns, *CAR*, for the event window $t = -1$, capturing information breaches and insider trading, to $t = +1$, capturing delay if investors use only follow-up information, such as news tickers picking up the CPSC announcement. *CAR* (−1,+1) is defined as the sum of the differences between actual and expected stock returns on days $t = -1, 0, +1$.

CPSC provides objective information about the *Firm Response*. In line with Chen et al. (2009), we code a response strategy as reactive (proactive) if incidents/injuries have (not) occurred. Following Liu et al. (2016), we classify firm responses as either partial or full remedy. Combining both codes we categorize firm responses as either (a) *conforming* (C1: reactive/full or C2: proactive/partial), (b) *underconforming* (UC: reactive/partial), or (c) *overconforming* (OC: proactive/full). For instance, in 2014, Emerson Electric Co. recalled 3.7 million travel charger kits with a loose wire that posed an electrocution hazard. They issued the recall after receiving 300 reports of injuries (= reactive), and offered free replacement (= full remedy), making it a conforming (C1) firm response (Table 1 lists examples for C1, C2, UC, and OC).

[4]A Factiva search for each recall yielded 12 concurring events, such as quarterly earnings announcements, which might have overshadowed the impact of the recall. Thus, we performed all analyses with the reduced set of observations. Our findings are robust to these events. Results for the reduced sample can be obtained upon request.

TABLE 1 Examples of conforming and nonconforming firm responses (Study 2)

Group	Firm	Product	Hazard	Injuries	Remedy
C1 (reactive/full)	Emerson Electric Co.	Travel charger kit	Electrocution	300	Free replacement
C2 (proactive/partial)	Steelcase Inc.	Desk chair	Fall	0	Free adhesive cover repair kit
UC (reactive/partial)	Graco	Stroller	Amputation and laceration	7	Free protective cover repair kit
OC (proactive/full)	Reebok	Sneaker	Choking	0	Full refund

4.2.4 | Correction for observed self-selection and omitted variable bias

Firms do not randomly choose their recall strategy. For example, firms might be more likely to offer full instead of partial remedy if the product hazard is high versus low. Further, as B&P argue, managers of higher and lower approval firms may accept less crisis responsibility than average approval firms (proposition 4, p. 362). Both variables might also shape investors' response to the product recall.[5] Accordingly, the estimate for the relationship between the crisis-response match and investor response is biased if this self-selection remains unaccounted for.

We set up a model that corrects the potential outcome (here: stock returns) in such a way that the observed covariates, which are correlated with the treatment (here: response strategy), do not affect the potential outcome anymore. Based on prior literature (e.g., B&P, Chen et al., 2009; Liu et al., 2016), we include important covariates into the model including hazard, reputation, financial risk (value of recalled products relative to the total revenue), product sell time, recall experience, product category, media coverage of the recall, firm size, profitability, and liability. These covariates can have an impact on response strategy and investor response, and thus have to be controlled for. Web Appendix C describes these covariates. To estimate the multivalued average treatment effects (here: underconforming, conforming, and overconforming), Cattaneo (2010) proposes the semiparametric efficient-influence estimator, which has the double-robust property (Wooldridge, 2010).[6] In stage one, this procedure estimates two models which give (a) for each observation the predicted probabilities for the three treatment statuses to control for self-selection effect on firm response and (b) the conditional outcome means to control for the covariates' effects on investor response. These first stage models include the control covariates mentioned above (for more details, see Web Appendix C). This first stage estimation procedure ensures that the three treatment groups are comparable and that the relationship between treatment statuses and outcome is independent of the covariates. In stage two, the treatment effect model then describes the bias corrected relationship between firm response and stock returns. Since we lack a "true" control group, i.e., firms which did *not* select in either treatment condition, the estimation procedure will produce a Local Average Treatment Effect (LATE).

[5]For example, B&P discuss the endowment effect, whereby "higher and lower levels of social approval [reputation] act as either a buffer or a burden to modify evaluators' crisis perceptions and attributions" (p. 362).
[6]Cattaneo, Drukker, and Holland (2013) discuss technical details and an implementation of this estimator in Stata.

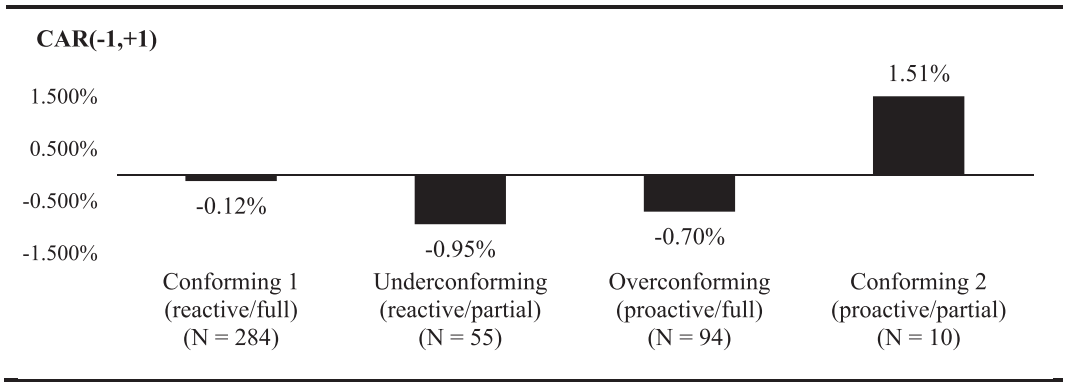

FIGURE 2 The crisis-response match and cumulative abnormal stock returns (Study 2)

4.2.5 | Results

Self-selection/omitted variable bias unadjusted results

Figure 2 shows that conforming strategies (C1: −0.12%, CI 95% (−5.41%, 5.80%), $N = 284$; C2: 1.51%, CI 95% (−1.04%, 4.52%), $N = 10$) receive more positive *CARs* than under- (UC: −0.95%, CI 95% [−8.41%, 2.78%], $N = 55$) and overconforming (OC: −0.70%, CI 95% [−5.71%, 3.49%], $N = 94$) strategies. A heteroscedasticity-robust Welch-test indicates that group differences are very likely (W-test [3, 43.65] = 4.681, $p = .006$). The relatively wide confidence intervals of the four groups might raise concerns about the influence of outliers. We therefore winsorize the variable *CAR*(−1,+1) at the 95% confidence interval.[7] Further, since proactive/partial strategies (C2) only account for about 2% in our sample (10 events), we combine this group with reactive/full (C1) for all subsequent analyses to represent conforming strategies.[8] The three resulting treatment groups are also very likely to have different *CARs* (W-test [2, 131.30] = 3.132, $p = .047$).

Self-selection/omitted variable bias adjusted results

Table 2 summarizes the multivalued treatment model results and provides two important insights. First, this model estimates the potential outcome for each treatment group. The potential outcome mean is negative for under- (−0.931%, $p = .001$) and overconforming (−0.747%, $p = .003$) groups, whereas it is not different from zero for the conforming group (−0.037%, $p = .814$). Second, the model provides information about treatment group differences. Conforming responses yield more positive *CARs* compared to under- (0.895%, $p = .004$) and overconforming (0.710%, $p = .016$) responses. Under- and overconforming responses do not have different *CARs* (−0.184%, $p = .617$).

Robustness check

The two-stage analysis might be insufficient to remedy the self-selection bias because firm responses do not result from a random assignment to all four conditions of the crisis-response

[7]Results for outlier unadjusted data are qualitatively similar and replicate the outlier adjusted findings reported below. Results are available upon request.

[8]A small sample adjusted nonparametric test shows that *CARs* are unlikely to differ between conforming groups ($\chi^2(1) = 0.932$, $p > .10$), whereby this merging does not introduce bias.

TABLE 2 Local average treatment effects (Study 2)

Group **Firm response**	**1** ***Underconforming***	**2** ***Conforming***	**3** ***Overconforming***
N	55	294	94
Potential outcome: CAR(−1,+1) in %[a]	−0.931 (0.274)	−0.037 (0.157)	−0.747 (0.248)
p-Value	.001	.814	.003

	LATE[a,b] (%)	***p*-Value**	**95% CI lo**	**95% CI hi**
Conforming vs. *Underconforming*	0.895 (0.312)	.004	0.284	1.505
Conforming vs. *Overconforming*	0.710 (0.294)	.016	0.001	1.286
Underconforming vs. *Overconforming*	−0.184 (0.369)	.617	−0.907	0.539

Note: Heteroscedasticity robust standard errors in parentheses. Covariates for modeling predicted probabilities (to control for covariates' effects on firm response): *Hazard*, *Reputation*, *Hazard*Reputation*, *Reputation*Reputation*, *Hazard*Reputation*Reputation*, *Financial Risk*. Covariates for modeling conditional outcome means (to control for the observed covariates' effects on investor response): *Hazard*, *Reputation*, *Hazard*Reputation*, *Reputation*Reputation*, *Hazard*Reputation*Reputation*, *Financial Risk*, *Product Sell Time*, *Recall Experience*, *Product categories Toy* (*binary*), *Child* (*binary*), *Sports* (*binary*), *Specialty* (*binary*), *Corporate Brand* (*binary*), *Firm Size*, *Return on Sales*, *Liability*, *Media Coverage* (−1,+1).

Abbreviation: CAR, cumulative abnormal stock return.

[a]Outlier adjusted results shown (outliers winsorized at the 95% confidence interval).

[b]LATE estimation based on efficient-influence-function (EIF) estimator with double-robust property (Cattaneo et al., 2013).

match matrix (Web Appendix B). Therefore, we conducted an additional experiment via Amazon's Mechanical Turk and randomly assigned participants (M_{age} = 36.39, 52% female) to one of the four conditions: C1 (reactive/full), C2 (proactive/partial), UC (reactive/partial), and OC (proactive/full). In short, the additional experiment replicates the findings of the event study and, also in line with Study 1, suggests that both under- and overconforming strategies yield lower perceptions of reputation by external evaluators than conforming strategies (see Web Appendix D for all study details).

4.2.6 | Discussion

We find empirical support that under- *and* overconforming firms would have had higher *CARs* if they had adopted a conforming strategy. Hence, this finding does not only replicate Study 1 and thus provides evidence for B&P's main proposition, but it also extends B&P's prediction to the financial effects of the crisis-response match. Further, Study 2 also shows that this effect is robust not only with regard to explicit but also to implicit informational cues about situational attribution and firms' willingness to accept crisis responsibility.

5 | GENERAL DISCUSSION

B&P (2015) provide a theoretical framework for optimal crisis-response strategies. Based on situational attributions and a firm's response strategy, they outline zones of conformity,

underconformity, and overconformity, and argue that social approval/reputation loss is lowest for conforming strategies. The present study is the first to empirically test this main proposition.

We find empirical support for this proposition, that is, situational attributions and response strategy must "match:" Conforming strategies outperform nonconforming strategies. We also find support for B&P's most surprising and counterintuitive prediction: overconforming strategies yield both lower firm reputation and stock returns relative to conforming strategies. Although exceeding stakeholders' expectations is essential for building a superior reputation, exceeding evaluators' (consumers' or investors') expectations during a crisis can have unintended negative consequences.

We extend B&P's main proposition in two ways. First, we not only find that conforming strategies outperform nonconforming strategies with regard to stakeholders' affective and intuitive evaluations (B&P call this social approval), but also with regard to cognitive and analytical evaluations. By jeopardizing a firm's perceived competence and likeability, nonconforming response strategies, including overconforming, pose a threat to a firm's reputation.

Second, while B&P do not differentiate between consumers and investors (i.e., they use the term "evaluators"), we make this distinction and analyze the financial consequences of the crisis-response match. The stock market punishes nonconforming strategies with a 0.5–1% reduction in market capitalization, which represents about $75 to $150 million for an S&P 500 stock with median market capitalization. This observation also holds for overconforming strategies. Investors are not only concerned about the financial costs of highly accommodative behavior in the face of lower situational attributions, but also about the unclear motivation behind this potentially overreaching and insincere response. Thus, it is essential for firms to clearly outline their rationale for using overconforming strategies so that external evaluators will not initiate counterfactual thinking and question the sincerity of the overconforming response. For underconforming strategies, which might involve lower financial costs in the short-term, investors seem to take a more comprehensive and longer-term view. They balance potential benefits for reputation and associated costs of this response strategy. Thus, firms should refrain from defensive responses and focus on crisis resolution (Pfarrer et al., 2008), which ultimately benefits the firm, its stakeholders, and society (B&P, 2015).

5.1 | Boundary conditions of the crisis-response match and future research

In this research note, we focus on B&P's main Proposition 1. However, B&P also offer insights into the boundary conditions of the crisis-response match. To test B&P's Propositions 2 and 3, which state that firms with higher (Proposition 2) and lower (Proposition 3) social approval can afford to offer more defensive response strategies, we conducted an additional experiment with students of a large U.S. public university (M_{age} = 26.58, 66% female), see Web Appendix E for all study details. In short, we find support for Proposition 2. Firms with high (vs. medium) levels of pre-recall reputation can afford to offer more defensive response strategies. However, we do not find support for Proposition 3, which suggests that firms with low (vs. medium) levels of pre-recall reputation can afford to offer more defensive response strategies. In fact, we find quite the opposite. Firms with low levels of pre-recall reputation can use product recalls to improve their reputation. Although this finding contradicts B&P's prediction and sounds quite paradoxical, prior research in the field of service failure management reports similar effects (service recovery paradox; e.g., De Matos, Henrique, & Alberto Vargas Rossi, 2007): customers

can have an even better opinion of the firm after a service failure if the firm handles the failure unexpectedly well.

Since this paradox is at odds with B&P's rationale for Proposition 3, future studies could examine B&P's Proposition 3 more theoretically and empirically. Moreover, B&P's Propositions 2 and 3 focus on a single boundary condition (pre-recall social approval/reputation). It is, however, very likely that other factors such as crisis type (e.g., performance- versus value-based misconduct) have a moderating impact as well.

The secondary data for Study 2 (event study) allows us to also test B&P's final Proposition 4, which states that managers of a higher or lower (vs. average) approval organization will be more likely to accept less crisis responsibility. In short, we find that B&P's prediction holds only for lower hazard product malfunctions, see Web Appendix F for all study details. Firms are less likely to choose underconforming responses if the product hazard is high. However, if the product hazard is low, managers of low and high reputation firms are (a) more likely to accept less crisis responsibility and select more underconforming (vs. conforming) responses, and (b) more likely to accept less crisis responsibility and not select overconforming (vs. conforming) responses. This result points to a partial gap between actual and ideal decisions. To understand the reasons for this gap, experiments and in-depth surveys could examine the psychological processes of managerial decision making in times of crises.

ORCID

Sascha Raithel https://orcid.org/0000-0002-2776-6802
Stefan J. Hock https://orcid.org/0000-0003-3530-7005

REFERENCES

Ashforth, B. E., & Gibbs, B. W. (1990). The double-edge of organizational legitimation. *Organization Science*, *1*(2), 177–194.

Ashforth, B. E., Harrison, S. H., & Corley, K. G. (2008). Identification in organizations: An examination of four fundamental questions. *Journal of Management*, *34*(3), 325–374.

Benoit, W. L. (1995). *Accounts, excuses, and apologies: A theory of image restoration strategies*. Albany, NY: State University of New York Press.

Bundy, J., & Pfarrer, M. D. (2015). A burden of responsibility: The role of social approval at the onset of a crisis. *Academy of Management Review*, *40*(3), 345–369.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82.

Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics*, *155*(2), 138–154.

Cattaneo, M. D., Drukker, D. M., & Holland, A. D. (2013). Estimation of multivalued treatment effects under conditional independence. *The Stata Journal*, *13*(3), 407–450.

Chen, Y., Ganesan, S., & Liu, Y. (2009). Does a firm's product-recall strategy affect its financial value? An examination of strategic alternatives during product-harm crises. *Journal of Marketing*, *73*(6), 214–226.

CNN. (2017). *Takata, brought down by airbag crisis, files for bankruptcy*. https://money.cnn.com/2017/06/25/news/companies/takata-bankruptcy/index.html.

Consumer Product Safety Commission. (2012). *Recall handbook*. https://www.cpsc.gov/s3fs-public/8002.pdf.

Coombs, W. T. (2007). Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate Reputation Review*, *10*(3), 163–176.

Coombs, W. T., & Holladay, S. J. (2004). Reasoned action in crisis communication: An attribution theory-based approach to crisis management. In D. P. Millar & R. L. Heath (Eds.), *Responding to crisis: A rhetorical approach to crisis communication* (pp. 95–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Coombs, W. T., & Holladay, S. J. (2005). Exploratory study of stakeholder emotions: Affect and crisis. In N. M. Ashkanasy, W. J. Zerbe, & C. E. J. Hartel (Eds.), *Research on emotion in organizations: Volume 1: The effect of affect in organizational settings* (pp. 271–288). New York: Elsevier.

Davis, D. F., Golicic, S., & Boerstler, C. N. (2011). Benefits and challenges of conducting multiple methods research in marketing. *Journal of the Academy of Marketing Science*, *39*(3), 467–479.

De Matos, C. A., Henrique, J. L., & Alberto Vargas Rossi, C. (2007). Service recovery paradox: A meta-analysis. *Journal of Service Research*, *10*(1), 60–77.

Dean, D. H. (2004). Consumer reaction to negative publicity: Effects of corporate reputation, response, and responsibility for a crisis event. *International Journal of Business Communication*, *41*(2), 192–211.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, *67*(3), 382–394.

Elsbach, K. D. (2003). Relating physical environment to self-categorizations: Identity threat and affirmation in a non-territorial office space. *Administrative Science Quarterly*, *48*(4), 622–654.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *The Academy of Management Journal*, *33*(2), 233–258.

Gray, E. R., & Balmer, J. M. T. (1998). Managing corporate image and corporate reputation. *Long Range Planning*, *31*(5), 695–702.

Hamilton, R. (2016). Consumer-based strategy: Using multiple methods to generate consumer insights that inform strategy. *Journal of the Academy of Marketing Science*, *44*(3), 281–285.

Lee, F., Peterson, C., & Tiedens, L. Z. (2004). Mea culpa: Predicting stock prices from organizational attributions. *Personality and Social Psychology Bulletin*, *30*(12), 1636–1649.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442.

Liu, A., Liu, Y., & Luo, T. (2016). What drives a firm's choice of product recall remedy? The impact of remedy cost, product hazard, and the CEO. *Journal of Marketing*, *80*(3), 79–95.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Pfarrer, M. D., DeCelles, K. A., Smith, K. G., & Taylor, M. S. (2008). After the fall: Reintegrating the corrupt organization. *The Academy of Management Review*, *33*(3), 730–749.

Philippe, D., & Durand, R. (2011). The impact of norm-conforming behaviors on firm reputation. *Strategic Management Journal*, *32*(9), 969–993.

Pratt, M. G. (2000). The good, the bad, and the ambivalent: Managing identification among Amway distributors. *Administrative Science Quarterly*, *45*(3), 456–493.

Raithel, S., & Schwaiger, M. (2015). The effects of corporate reputation perceptions of the general public on shareholder value. *Strategic Management Journal*, *36*(6), 945–956.

Robert, P. W., & Dowling, G. R. (1997). The value of a firm's corporate reputation: How reputation helps and sustain superior profitability. *Corporate Reputation Review*, *1*(1/2), 72–76.

Traut-Mattausch, E., Shulz-Hardt, S., Greitemeyer, T., & Frey, D. (2004). Expectancy confirmation in spite of disconfirming evidence: The case of price increases due to the introduction of the euro. *European Journal of Social Psychology*, *34*(6), 739–760.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Tyler, L. (1997). Liability means never being able to say you're sorry: Corporate guilt, legal constraints, and defensiveness in corporate communication. *Management Communication Quarterly*, *11*(1), 51–73.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: The MIT Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Raithel S, Hock SJ. The crisis-response match: An empirical investigation. *Strat Mgmt J*. 2021;42:170–184. https://doi.org/10.1002/smj.3213