Can bad news be good? On the positive and negative effects of including moderately negative information in CSR disclosures

Johannes Jahn*, Rolf Brühl

ESCP Europe Business School Berlin, Heubnerweg 8-10, 14059 Berlin, Germany

ABSTRACT

Most companies are eager to present themselves as "green," "sustainable" or "socially responsible" by making overly positive and optimistic corporate social responsibility (CSR) disclosures. Failures, setbacks and mistakes, however, are usually withheld. This study critically examines the advantageous and disadvantageous effects of voluntarily disclosing moderately negative information. We argue that two countervailing effects exist that affect corporate trustworthiness. First, attribution theory predicts that voluntarily reporting negative information is likely to be attributed to the honesty of the company. However, second, readers may perceive the voluntary disclosure of negative information as implausible, which adversely affect trustworthiness. The study additionally investigates the effects of moderately negative information and trustworthiness on CSR perception. An experiment was conducted to test these relationships, and the results provide support for the positive and negative effects on trustworthiness. Moreover, no negative effect on CSR perception was detected.

1. Introduction

A plethora of studies show that the vast majority of companies that release corporate social responsibility (CSR) disclosures favor publishing good news (e.g., Hackston & Milne, 1996; Holder-Webb, Cohen, Nath, & Wood, 2009; Scalet & Kelly, 2010) and try to distract the readers from corporate irresponsibility (Perks, Farache, Shukla, & Berry, 2013). The literature provides several insights that highlight the importance of a favorable CSR perception, such as its positive impact on corporate reputation (Hur, Kim, & Woo, 2014) and on the job performance of employees (Korschun, Bhattacharya, & Swain, 2014). Negative CSR associations, on the other hand, cause disadvantages; e.g., they adversely affect product evaluations (Brown & Dacin, 1997) and word-of-mouth (Skarmeas & Leonidou, 2013). Likewise, green skepticism (i.e., consumers' doubts regarding the environmental benefits of a firm) negatively affects purchase intentions (Leonidou & Skarmeas, 2017). Thus, the importance of "avoiding bad" to maintain a positive CSR perception is often emphasized (Lin-Hi & Müller, 2013). However, the communication of overly positive information is not without risk. When companies are believed to claim to be something that they are not, also referred to as corporate hypocrisy, consumers' attitudes toward the company are negatively affected (Wagner, Lutz, & Weitz, 2009). This seems to lead to a dilemma for companies: Should they push forward positive information about the CSR activities or reveal negative information to avoid stakeholders' disapproval?

Research in attribution theory argues that good news is often consistent with incentives of the issuing firm because it presents the company's products, achievements or capabilities in a positive light. Bad news, on the other hand, is contrary to a firm's incentives because it highlights failures and mistakes. Consequently, the latter may be perceived to be more credible because management has less of an incentive to publish it when it is untrue (Crowley & Hoyer, 1994; Mercer, 2004). This effect particularly holds true for voluntary disclosures, because forced disclosure of negative information is unlikely to be attributed to the credibility of the company (Jones & Davis, 1965; Smith & Hunt, 1978). The literature on voluntary earnings forecasts (Hutton, Miller, & Skinner, 2003; Williams, 1996) revealed the higher credibility of negative compared to positive information. Moreover, marketing research provides strong support for the positive effect of disclosing negative information on corporate credibility (Crowley & Hoyer, 1994; Eisend, 2006). The current study will refer to the latter research stream, in which attribution theory is a dominant approach (Golden, 1977; Mizerski, Golden, & Kernan, 1979; Settle & Golden, 1974).

Disclosures that contain positive as well as negative information are often called two-sided messages.[1] While marketing scholars normally refer to positive and negative product features, we seek to investigate

---

* Corresponding author.
  E-mail addresses: jjahn@escpeurope.eu (J. Jahn), rbruehl@escpeurope.eu (R. Brühl).
  [1] Other conceptualizations of two-sided messages are also conceivable. Rim and Song (2016) referred to a two-sided message as when a company admits to having financial motives for engaging in a pro-social activity.

the effects of publishing two-sided CSR disclosures, i.e., disclosures that mention the positive and the negative aspects of a company's social performance. Whereas evaluations of product features are mainly driven by instrumental motives of consumers, an evaluation of social performance may be based on moral grounds. However, in the case of moral judgments, negative information is particularly meaningful regarding the attribution of negative traits (Skowronski & Carlston, 1989). Thus, in a CSR setting, the hurdle for disclosing negative information might be higher.

This study seeks to address two research gaps. First, while attribution theory proposes a linear increase in credibility for higher proportions of negative information (Crowley & Hoyer, 1994), Eisend (2006) found in his meta-analysis that evidence for this relationship is weak at best. We aim to investigate whether this linear effect exists and whether there is another, countervailing effect at work. Mercer (2004) argued that missing plausibility, defined as deviance from previous expectations, negatively affects credibility. We posit that such an effect might be particularly pronounced for very high proportions of negative information and thus weakens the credibility-enhancing effect of including negative information as proposed by attribution theory.

Second, past research on the effect of negative information in CSR disclosures is scarce and provides contradicting findings. While Chan and Milne (1999) indicated that disclosing negative information detrimentally affects investment decisions, Reimsbach and Hahn (2015) did not find adverse reactions. One reason for the contradictory findings could be that neither of these studies differentiates between different proportions of negative information. Moreover, although previous marketing research and attribution theory emphasize the importance of credibility (e.g., Eisend, 2006; Settle & Golden, 1974), far too little attention has been paid to this construct when referring to two-sided CSR disclosures. Finally, we propose two effects on CSR perception triggered by two-sided disclosures: a positive effect that is driven by corporate credibility and a negative effect caused by the negative information itself.

To test how various proportions of negative CSR information affect credibility and CSR perception, an online experiment was conducted. While previous research focused on credibility in general, we concentrate on one dimension of credibility: trustworthiness. Trustworthiness refers to the willingness of the company to honestly disclose information and is therefore particularly relevant in the given context. In a between-subject design, participants had to assess CSR disclosures that contained 0%, 20%, 40% or 60% negative information. In line with previous research, we focus on moderately negative information, which is supposed to be more likely to cause positive effects compared to severe or important negative information (Crowley & Hoyer, 1994; Eisend, 2006). We found that two-sided CSR disclosures indeed have a direct and positive effect on trustworthiness, but also that there is a distinct, contradicting effect for high proportions of moderately negative information that is triggered by the plausibility of the message. Moreover, we did not find an adverse effect on CSR perception for any proportion of negative information.

2. Theoretical framework

2.1. Positive and negative effects of two-sided messages on trustworthiness as a dimension of corporate credibility

While companies face strong incentives to publish overly positive information, they should also consider that "[w]hether or not ethical positioning and image repair campaigns are successful depends, to a large extent, on the company's credibility" (Schlegelmilch & Pollach, 2005).

Depending on the field of research, credibility can have different meanings. Here, we refer to credibility as the "communicator's positive characteristics that affect the receiver's acceptance of a message" (Ohanian, 1990). With respect to corporate credibility, the two major characteristics are trustworthiness and expertise (Newell & Goldsmith, 2001). Expertise refers to a company's experience, knowledge and capabilities regarding a particular issue; thus, expertise reflects management's ability to provide truthful information. Trustworthiness, on the other hand, can be regarded as the management's willingness to honestly disclose information. Particularly, this refers to the extent to which the recipient of a message believes that the communicated content reflects the assertions the management perceives as valid (Hovland, Janis, & Kelley, 1953).

CSR disclosures are, to a large extent, unregulated; i.e., there are only a few laws that determine the scope and content of such disclosures. Thus, whether or not to publish only positive or also negative information is a managerial decision. From this follows that management's willingness—not its ability—to unveil negative information is crucial. In consequence, the current study refers to the trustworthiness dimension of corporate credibility.

Scholars who investigate the credibility of positive and negative information regularly refer to attribution theory (Eisend, 2010; Golden, 1977; Settle & Golden, 1974). People attribute observed behavior either to external causes, such as environmental or situational factors, or to internal causes, such as an actor's personal characteristics or disposition (Heider, 1958). However, the latter are only inferred in the case of uncommon behavior (Jones & Davis, 1965), such as the disclosure of negative information (Smith & Hunt, 1978).

For our study, Kelley's *discounting principle* is particularly relevant. He states that "the role of a given cause in producing a given effect is discounted if other plausible causes are also present" (Kelley 1971, p. 8). Observing the disclosure of positive information allows two different conclusions. First, the company is telling the truth (internal cause). Second, the information is disclosed to present the company in a favorable light (external cause). Because both explanations are possible, each is discounted. However, if negative information is disclosed, the latter reason is less likely, and one is more willing to attribute the disclosure to the honesty of the corporation—its internal disposition to tell the truth.

The effect described above only emerges if the disclosure is made voluntarily (Jones & Davis, 1965). If a company is forced by law to publish negative information, it can hardly be attributed to the company's willingness to tell the truth, but rather to the external, legal environment. Hence, trustworthiness does not increase in such a case.

The trustworthiness-enhancing effect is driven by company incentives: The higher the proportion of negative information, the stronger the contradiction to company incentives. This, in turn, results in higher perceived trustworthiness of the company. Consequently, attribution theory predicts trustworthiness "to be a monotonically increasing function of the amount of negative information communicated" (Crowley & Hoyer, 1994, p. 570).

To ensure comparability with previous research, we focus on moderately negative information, which is supposed to be more likely to cause beneficial effects (Crowley & Hoyer, 1994; Eisend, 2006).

Thus, we hypothesize the following:

Hypothesis 1 (H1). Two-sided CSR disclosures cause an increase in trustworthiness compared to one-sided CSR disclosures. This increase in trustworthiness becomes stronger the higher the proportion of (moderate) negative information included in a CSR disclosure.

In contrast to the predictions made by attribution theory, empirical studies investigating this effect do not find support for this relationship. Anderson and Golden (1984) found one negative attribute to be perceived as more credible than two. Golden and Alpert (1987) varied the proportion of negative information in their experiment between two-sevenths and two-thirds but did not find evidence that these differences

affect credibility. Still, Eisend (2006) found a positive, but not significant, relationship between the amount of negative information and source credibility in his meta-analysis.

A possible explanation for the discrepancy between attribution theory and empirical findings is provided by Mercer (2004). She argued that plausibility, defined as deviation from the existing expectations of the readers, affects credibility perception. Because companies normally only publish positive information, with higher levels of negative information, behavior increasingly deviates from expectations and disclosures are perceived as more implausible.

Hypothesis 2a (H2a). Two-sided CSR disclosures cause a decrease in plausibility compared to one-sided CSR disclosures. This decrease in plausibility becomes stronger the higher the proportion of (moderate) negative information included in a CSR disclosure.

Plausibility, in turn, is expected to be positively related to credibility (Mercer, 2004). The relationship between both is probably familiar to most people; if we hear stories about events and behaviors that deviate from what we would normally expect, we often become skeptical about whether this story indeed reflects the truth. If someone comes late to an appointment and apologizes by saying that he or she was stuck in a traffic jam, we doubt this story if we know that during that time, there is normally less traffic. Likewise, we become skeptical when someone who normally rarely plays sports tells us that he or she ran a marathon last weekend. In both cases, the stories are implausible and thus lack trustworthiness: "In a sense, already known information has a privileged truth status or priority: Subjects believe the information they possess to be true, and they doubt information that contradicts it" (Bacon, 1979). The same principle applies to corporate communication. Barton and Mercer (2005) argue that implausible explanations for bad performance are less likely to be believed by analysts and could show that implausible disclosures result in lower earnings forecasts and stock price evaluations. Thus, we hypothesize:

Hypothesis 2b (H2b). Plausibility is positively related to trustworthiness.

Taken together, this implies that the disclosure of moderately negative information causes a second, negative and indirect effect on trustworthiness, next to the direct positive effect described above (see H1). The voluntary disclosure of negative information is likely to be perceived as implausible and will thus adversely affect trustworthiness. As companies normally have no incentives to publish such information, people might doubt that the disclosure was indeed voluntary; maybe the company claimed to have published it voluntarily but was actually forced to do so. If readers do not believe in the voluntariness of the disclosure, no attribution toward trustworthiness occurs. Moreover, it might be possible that readers might think that in such a case, the company tries to distract the readers' attention from other negative

events so that readers doubt the motives of the company for voluntary disclosure. Attributed motives might then cause an adverse reaction (Ellen, Webb, & Mohr, 2006). Thus, we hypothesize:

Hypothesis 2c (H2c). There is a negative effect of two-sided CSR disclosures on trustworthiness that is mediated by plausibility.

Fig. 1 illustrates the conceptual model comprising H1 and H2a, H2b, H2c as well as the subsequently developed hypotheses H3 and H4.

2.2. Positive and negative effects of two-sided messages on CSR perception

CSR perception can be viewed as stakeholder judgment regarding the fulfillment of the various responsibilities of corporations, such as social and environmental issues that go beyond profit maximization (Alvarado-Herrera, Bigne, Aldas-Manzano, & Curras-Perez, 2017). We argue that two independent and contradicting effects exist that affect CSR perception: It is positively related to trustworthiness but negatively related to the proportion of moderately negative information.

First, Schwartz and Carroll (2008) considered accountability, which is supported by the "normative principles of […] transparency, promise keeping, […] honesty, [and] trustworthiness" (p. 172), to be an important dimension of CSR. Furthermore, avoiding deception is likewise regarded as a central responsibility of corporations (Donaldson, 1982; Friedman, 1970). Along with this normative relatedness, the positive relationship between trustworthiness/credibility and CSR perception is also empirically supported (Alcañiz, Cáceres, & Pérez, 2010; Schmeltz, 2012; Swaen & Vanhamme, 2004). Additionally, CSR disclosures made by trustworthy companies can be interpreted as more serious commitments to the underlying goals than those of less trustworthy companies.

Thus, we posit the following:

Hypothesis 3 (H3). Trustworthiness is positively related to CSR perception.

Second, moderately negative information in CSR disclosures can refer to missed objectives, e.g., regarding the reduction of CO_2 emissions or unsafe working conditions. One can imagine that honestly admitting a few such incidences could increase trustworthiness and, in turn, CSR perception. However, the sole disclosure of such information is hardly likely to do so, because companies that constantly fall short of social and ecological expectations are acting irresponsibly rather than responsibly. Moreover, because most companies do not publish negative information, companies that do so are more likely to be perceived as unique. However, this also applies to the negative features they communicate (Eisend, 2010). Hence, CSR perception is negatively affected by the disclosure of negative information.

Hypothesis 4 (H4):. A higher proportion of (moderate) negative information is negatively related to CSR perception.
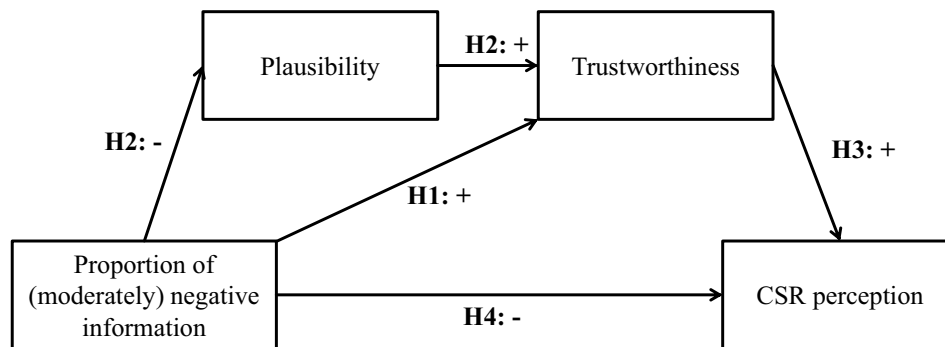

Fig. 1. Conceptual model.

The optimal proportion of negative information (i.e., whether the joint effect of trustworthiness [H3] and proportion of negative information [H4] is positive or negative) is at the heart of two-sided message research (Crowley & Hoyer, 1994). Meta-analytic results for two-sided advertisements show positive effects on various variables (e.g., purchase intention and attitude toward the brand) for moderate proportions of negative information (Eisend, 2006). With respect to two-sided CSR disclosures, Chan and Milne (1999) found negative effects, while Reimsbach and Hahn (2015) found no adverse effects from the disclosure of negative information. This difference might be explained by the fact that the latter study used disclosures that contained positive and negative information, while the former only used negative information (Reimsbach & Hahn, 2015). However, along with the proportion of negative information, the effectiveness of two-sided messages depends on several message characteristics and context factors (e.g., Eisend, 2006; Golden, 1979). This makes precise hypothesizing about the strength and direction of a joint effect exceedingly speculative. Still, it is reasonable to assume that there will be no decrease in CSR perception for moderate proportions of negative information.

3. Method

3.1. Experimental design

A vignette experiment with four experimental groups (between-subject design) was conducted. The independent variable, message sidedness, was manipulated with respect to the proportion of moderately negative information included in a CSR disclosure. We varied the proportion of negative information from 0% to 60% with intermediate stages at 20% and 40%. Fig. 2 illustrates the experimental design.

3.2. Study material

Participants were asked to read information about a fictitious company ("Lainko AG") and to assess the company afterward. A fictitious company was chosen to prevent varying prior associations of the participants that would influence company assessment. Material regarding the company was presented to the participants in two steps.

First, participants were asked to read a short company description that contained information regarding industry, products, year of foundation, location of production facilities and sales regions, product quality and financial success. The company description was identical for all four treatments. In accordance with prior studies on two-sided messages (e.g., Chebat & Picard, 1988; Golden & Alpert, 1987) and credibility (e.g., Alcañiz et al., 2010; Lafferty, 2007), we chose the cosmetics industry for Lainko AG. The cosmetics industry is well suited to the experiment because its products are used by almost everybody and are generally noncontroversial (compared to tobacco or alcohol), and product choice is less involving (Golden & Alpert, 1987; Parguel, Benoît-Moreau, & Larceneux, 2011). The other attributes (e.g., the year of foundation, product quality) were described in a way that presented the company as moderately credible (Lafferty, 2007). This was done to allow for an increase in trustworthiness due to two-sided CSR disclosures.

Second, after reading the company description, participants were asked to read a *letter to the stakeholders* (see Appendix A). A letter to the stakeholders contains a summary of the company's CSR goals and achievements and can be found at the beginning of most CSR reports as well as on company websites. The presentation of the most important aspects of the report in letter form makes the given information comprehensible and assessable while ensuring external validity.

The fictitious letter to the stakeholders contained information on the company's activities regarding its employees and social engagement, which was consistent among all treatments, as well as information concerning its ecological measures. Similar to previous research (e.g., Parguel et al., 2011), we manipulated only the company's ecological engagement, because environmental issues can be considered to represent the most prominent dimension of CSR. The instructions clearly indicated that the disclosure was made voluntarily.

The proportion of moderately negative information was manipulated through statements that indicated whether or not the company would achieve its various ecological objectives. This is in line with a proposal from Hahn and Lülfs (2014) that a negative aspect of sustainability reporting refers to corporate conduct with a negative impact on the realization of sustainability. To increase the external validity of our study, the goals were identified by analyzing the CSR reports of leading companies within the cosmetics industry (e.g., Beiersdorf, Procter & Gamble). We ensured that negative information was placed neither first nor last because Eisend (2006) showed that such positioning would additionally influence participants' judgment.

3.3. Sample

Participants were recruited with the support of an online panel provider. Thus, the participants read the material and answered the
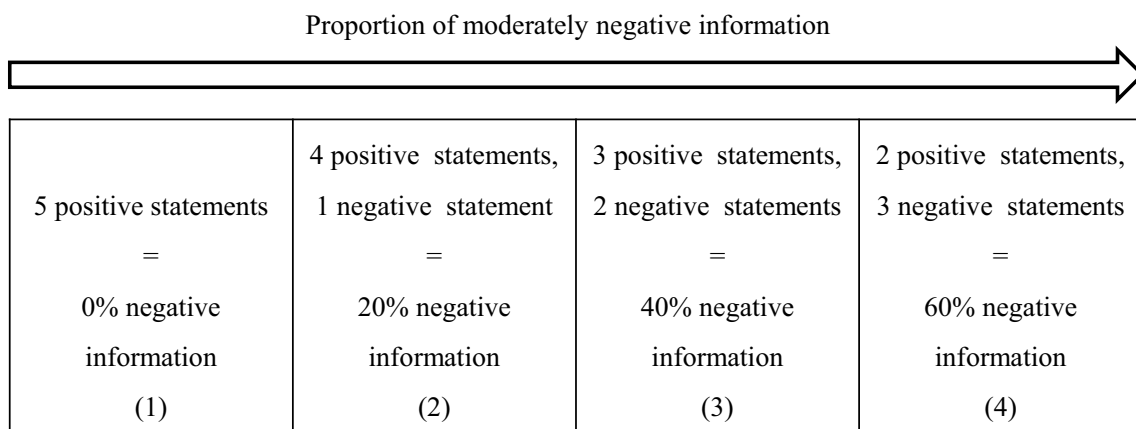
Proportion of moderately negative information

5 positive statements = 0% negative information (1)	4 positive statements, 1 negative statement = 20% negative information (2)	3 positive statements, 2 negative statements = 40% negative information (3)	2 positive statements, 3 negative statements = 60% negative information (4)

Fig. 2. Experimental design, four experimental groups.

subsequent questionnaire (see below, "Variables and measures") at their own computers. The online setting was fitting because information regarding the social and environmental activities of corporations, such as corporate websites, press releases or CSR reports, is regularly available as an online source.

The initial sample contained 596 participants who completed the questionnaire. We applied three exclusion criteria to ensure that only participants who completed the questionnaire diligently were included. First, we excluded participants who did not read the company description carefully. We assessed this by including a set of five questions (e.g., "Does the company have a production site in Poland?" or "Does the company sell shaving foam?") placed after the company description. Those participants who were not able to answer at least three of the five questions correctly were excluded. This reduced our sample to n = 448. Second, we excluded participants who clicked through the letter to the stakeholders. Applying a cutoff criterion of at least 10 s for reading the letter further reduced our sample to n = 382. Third, participants who took < 4 min to answer the complete questionnaire were excluded, which led to a sample size of 379 participants.[2] Due to missing data, the final sample contained 318 participants.

The sample was recruited within Germany. The participants were, on average, 44.0 years old; 56.3% were female, and 39.9% held the general qualifications for university entrance or a university degree. Comparing these figures with the German population (Federal Statistical Office, 2015, figures for Germany are: average age 43.2 years, 51.2% female, 29.4% hold university entrance qualification) shows that the sample is suitably representative of the German population. The application of such a sample for the given setting is appropriate because CSR disclosures, in contrast to financial disclosures, are addressed to all stakeholders and not only to investors and creditors (e.g., Van der Laan Smith, Adhikari, & Tondkar, 2005).

3.4. Variables and measures

The independent variable of this study was message sidedness. To conduct a manipulation check, participants had to indicate the proportion of negative information in the letter to the stakeholders on an 11-point Likert scale (1 = "only negative information"; 11 = "only positive information"). The mean answers were as follows: 9.25 (0% negative information), 8.17 (20% negative information), 7.58 (40% negative information) and 6.62 (60% negative information). An ANOVA with pairwise comparison revealed that these differences were significant at the 1% level, except for the difference between the 20% and 40% groups (p = .17). These results indicate successful manipulation of the independent variable.

To measure trustworthiness, we referred to a scale developed by Newell and Goldsmith (2001). CSR perception was measured with a scale developed by Alvarado-Herrera et al. (2017). Because we manipulated environmental issues, we only used the items that refer to environmental perception from the CSR perception scale. Both constructs were measured on 7-point Likert scales. The items are depicted in the Appendix (see Table B1 in Appendix B).

In order to assess plausibility, we followed Mercer (2004) and regarded it as deviance from previous expectations. Participants had to indicate on a scale (1 = "very unlikely"; 7 = "very likely") whether they believed that companies would voluntary publish the letter to the stakeholders they had just read. A low score implies that such a disclosure was perceived as implausible because they deemed such behavior unlikely and would not expect companies to act in this manner. A

high score, in contrast, implies plausibility.

Furthermore, two control variables were included in the questionnaire. A trust disposition scale (Wiedmann, Wüstefeld, & Klibert, 2011) was included because it indicates a person's willingness to trust others (McKnight, Cummings, & Chervany, 1998). Moreover, we included a scale measuring the participants' sensitivity to social performance (Liu, Wang, & Wu, 2010) because this is expected to affect participants' reaction to negative information (Liu et al., 2010). Both constructs were measured on 7-point Likert scales. The items are depicted in the Appendix (see Table B1 in Appendix B).

3.5. Reliability and validity

Several analyses were run to assure the reliability and validity of our measures. First, we ran an exploratory factor analysis (EFA) with all measures (maximum likelihood, oblimin rotation) in order to assess factor loadings and cross-loadings of items. After eliminating items, we reran the EFA (see Table B2 in Appendix B for cross-loadings). We then ran a confirmatory factor analysis (CFA) with this 5-factor model with good fit indices (χ^2 = 160.18, df = 110, p = .00, Comparative Fit Index [CFI] = 0.986, Tucker-Lewis Index [TLI] = 0.983, Root Mean Square Error of Approximation [RMSEA] = 0.038). We assessed individual item reliability according to the factor loadings of the measurement model (see Table B1 in Appendix B). All factor loadings are above the recommended value of 0.5 in order to be practically significant (Hair, Black, Babin, & Anderson, 2014). The reliability of the scales was assessed using Cronbach's alpha. Alpha should be above 0.7 (Hair et al., 2014), which is the case for our constructs (see Table B1 in Appendix B). We examined convergent validity with the average variance extracted (AVE, see Table B1 in Appendix B), which should be above 0.5 for the constructs; this is the case for all of our constructs (Fornell & Larcker, 1981). Apart from assessing the cross-loadings (see Table B2 in Appendix B), discriminant validity was assessed by comparing the square root of AVE of each construct with the correlations to the other constructs. All square roots of AVE exceeded their respective correlations (Fornell & Larcker, 1981), which shows discriminant validity (see Table B1 in Appendix B). In summary, the results show the validity and reliability of our constructs.

3.6. Common method variance

We controlled for common method variance by introducing a single unmeasured latent method factor (ULMF) (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). All manifest variables are related to the ULMF, those factor loadings are constrained to be equal and the covariance of the ULMF with the other factors is constrained to be 0. Model 1 is a trait-only model in which each measure is loaded on its respective construct (χ^2 = 160.18, df = 110, p = .00, CFI = 0.986, TLI = 0.983, RMSEA = 0.038), and Model 2 is a trait-and-method model in which a ULMF is added to Model 1 (χ^2 = 159.55, df = 108, p = .00, CFI = 0.986, TLI = 0.982, RMSEA = 0.039). Comparing the two models reveals that the difference in model fit (ΔCFI) between Model 1 and Model 2 is < 0.010, which is not a significant difference (Cheung & Rensvold, 2002). This indicates that the variance in the data can be explained by the traits rather than by a common-method variance.

3.7. Data analyses

A mediation analysis according to the procedure suggested by Hayes (2017) (10,000 bootstrap samples, 95% confidence interval) was conducted, whereby mediation does not require a direct effect of the independent on the dependent variable; an indirect effect is sufficient (Shrout & Bolger, 2002; Wood, Goodman, Beckmann, & Cook, 2008; Zhao, Lynch Jr, & Chen, 2010). The proportion of moderately negative information was treated as a multi-categorical independent variable, and dummy coding was applied as follows: 0% negative information

---

[2] While the application of these criteria is justified, the actual cutoff criteria were to some extent arbitrarily chosen. In order to show the robustness of our results, we repeated our analysis several times with varying cutoff criteria (lowest n was 287, highest n was 372). These changes did not affect the results of our analysis.

Table 1

Means, Standard Deviations, and correlations for all dependent, independent, and control variables in analyses.

	M (SD)	1	2	3	4	5	6	7	8	9
1. Age	43.98 (12.62)	–								
2. Gender	–	0.12*	–							
3. Education	–	−0.16**	−0.01	–						
4. Proportion of negative information	–	−0.00	−0.05	−0.02	–					
5. Plausibility	4.70 (1.52)	0.07	0.07	−0.01	−0.34**	–				
6. Trustworthiness	4.99 (1.09)	0.05	0.06	−0.04	0.03	0.39**	–			
7. CSR perception	5.16 (1.09)	0.13*	−0.03	−0.18**	−0.06	0.38**	0.62**	–		
8. Trust disposition	4.11 (1.19)	0.02	0.06	−0.05	−0.04	0.32**	0.37**	0.31**	–	
9. Sensitivity to social performance	5.09 (1.18)	0.09	−0.15**	−0.03	0.04	0.17**	0.28**	0.32**	0.33**	–

Note. Pearson correlation for all pairings except those that involve Education (Spearman).

* $p < 0.05$.

** $p < 0.01$.

Table 2

Direct effects for a mediation analysis with proportion of negative information as multi-categorical independent variable, plausibility and trustworthiness as mediator and CSR perception as dependent variable.

Direct effects	M1 (plausibility)			M2 (trustworthiness)			Y (CSR perception)		
Antecedent	Coeff.	SE	*p*	Coeff.	SE	*p*	Coeff.	SE	*p*
D1–20% negative information	−0.33	0.21	.13	0.27	0.15	.07	−0.12	0.14	.37
D2–40% negative information	−0.93	0.22	< .01	0.31	0.16	.04	−0.03	0.14	.85
D3–60% negative information	−1.28	0.21	< .01	0.44	0.16	.01	−0.12	0.14	.39
M1 – Plausibility	–	–	–	0.24	0.04	< .01	0.10	0.04	.01
M2 – Trustworthiness	–	–	–	–	–	–	0.52	0.05	< .01
C1 – Sensitivity to social performance	0.11	0.06	.08	0.12	0.05	.01	0.12	0.02	< .01
C2 – Trust disposition	0.35	0.07	< .01	0.19	0.05	< .01	0.03	0.05	.52
	$R^2 = 0.22$			$R^2 = 0.26$			$R^2 = 0.42$		
	$F_{(5, 312)} = 17.34, p \leq .01$			$F_{(6, 311)} = 17.88, p \leq .01$			$F_{(7, 310)} = 32.49, p \leq .01$		

was coded as 0, 20% as D1, 40% as D2 and 60% as D3.

In order to test whether trustworthiness (H1) and plausibility (H2a) increases with higher message sidedness, additional ANCOVAs were conducted to compare the various groups that received negative information.

4. Results

4.1. Descriptive statistics

Correlations for all independent, dependent and control variables as well as demographic figures are depicted in Table 1. The table shows that the included control variables (sensitivity to social performance, trust disposition) are positively correlated with all three dependent variables. Furthermore, education (operationalized in groups: low, middle and high education level) is negatively correlated with CSR perception, while age is positively correlated with CSR perception. However, no randomization imbalance was found for education (χ^2 [2, n = 318] = 2.63, p = .85) or age (F [3, 314] = 0.05, p = .99), so neither is included as a control variable.

4.2. Hypotheses tests

Participants who read a two-sided disclosure showed two different reactions with respect to trustworthiness. First, there was a direct positive effect on trustworthiness (see Table 2 for all direct effects; direct effects that affect trustworthiness are also depicted in Fig. 3). This effect prevailed for disclosures that contained 20% (significant at the 0.10 level), 40% and 60% (both significant at the 0.05 level) of moderately negative information, which provides initial support for H1. However, as proposed by H2a, those who received disclosures with 40% and 60%

negative information also perceived the disclosure as less plausible. Low plausibility, in turn, negatively affected trustworthiness. Thus, H2b is supported. The mediation analysis[3] also shows an indirect negative effect on trustworthiness through plausibility. This indirect effect is statistically different from zero, as revealed by a 95% bootstrap confidence interval that is entirely above zero, for the 40% and 60% group (see Table 3 for all indirect and total effects). Thus, H2c is supported.

An ANCOVA with trustworthiness as a dependent variable (see Table 4) confirmed that all groups that received two-sided messages perceived the company to be more trustworthy compared to the company that only disclosed positive information. However, we did not find significant differences between the three groups that received moderately negative information, though the differences between the groups are in the proposed direction. Thus, the ANCOVA does not indicate a linear increase in trustworthiness. Still, the coefficients of the direct effects for proportion of negative information on trustworthiness in the mediation model are stepwise increasing. Thus, the data at least provides some indication for a linear effect.[4] Nevertheless, H1 is only partly supported.

In order to determine whether plausibility decreases with an increasing proportion of moderately negative information, a further ANCOVA was conducted. The results (see Table 5) show that higher

---

[3] The mediation analysis is presented here without suppressing the effect from plausibility on CSR perception, so we have not built a hypothesis on this relationship. Repeating the mediation analysis while suppressing this effect did not affect the results.

[4] Repeating the mediation analysis and treating proportion of negative information as a continuous variable (percentage values) showed a significant effect of proportion of negative information on trustworthiness (coeff. = 0.69, $p < 0.01$).
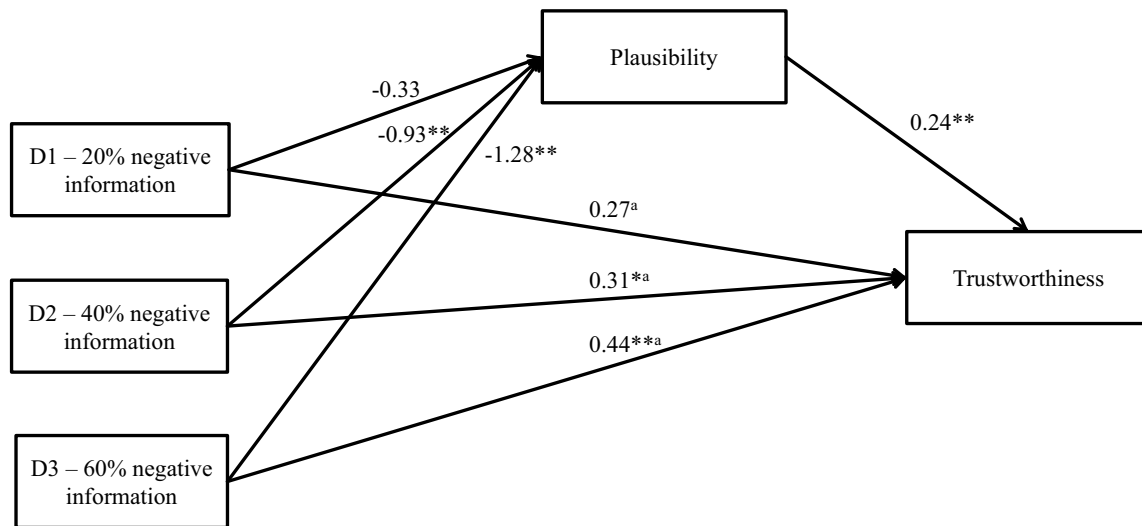
Fig. 3. Unstandardized regression coefficients for the relationship between proportion of negative information and trustworthiness as mediated by plausibility. Note. $*p < 0.05$, $**p < 0.01$, a – unstandardized regression coefficient after controlling for plausibility.

Table 3
Indirect and total effects for a mediation analysis with proportion of negative information as multi-categorical independent variable, plausibility and trustworthiness as mediator and CSR perception as dependent variable.

Indirect effects	Unstandardized indirect effect	SE	LLCI	ULCI
D1 ⇒ Plausibility ⇒ Trustworthiness[a]	−0.08	0.06	−0.20	0.02
D2 ⇒ Plausibility ⇒ Trustworthiness[a]	−0.22	0.07	−0.37	−0.10
D3 ⇒ Plausibility ⇒ Trustworthiness[a]	−0.31	0.08	−0.47	−0.17
D1 ⇒ Plausibility ⇒ CSR perception	−0.03	0.03	−0.10	0.01
D2 ⇒ Plausibility ⇒ CSR perception	−0.09	0.05	−0.19	−0.01
D3 ⇒ Plausibility ⇒ CSR perception	−0.12	0.06	−0.25	−0.02
D1 ⇒ Trustworthiness ⇒ CSR perception	0.14	0.08	0.00	0.32
D2 ⇒ Trustworthiness ⇒ CSR perception	0.16	0.09	−0.01	0.36
D3 ⇒ Trustworthiness ⇒ CSR perception	0.22	0.10	0.05	0.43
D1 ⇒ Plausibility ⇒ Trustworthiness ⇒ CSR perception	−0.04	0.03	−0.10	0.01
D2 ⇒ Plausibility ⇒ Trustworthiness ⇒ CSR perception	−0.12	0.04	−0.20	−0.05
D3 ⇒ Plausibility ⇒ Trustworthiness ⇒ CSR perception	−0.16	0.04	−0.25	−0.09

Total effects	Relative total effects	SE	LLCI	ULCI
D1 ⇒ Trustworthiness[a]	0.19	0.16	−0.12	0.51
D2 ⇒ Trustworthiness[a]	0.09	0.16	−0.23	0.40
D3 ⇒ Trustworthiness[a]	0.13	0.16	−0.18	0.43
D1 ⇒ CSR perception	−0.05	0.16	−0.37	0.27
D2 ⇒ CSR perception	−0.07	0.16	−0.39	0.25
D3 ⇒ CSR perception	−0.18	0.16	−0.49	0.13

Note. LLCI = lower limit confidence interval, ULCI = upper limit confidence interval, confidence interval level = 95%.

[a] These effects are calculated by applying a simple mediation model with proportion of negative information as independent variable, plausibility as mediator and trustworthiness as dependent variable.

proportions of negative information are less plausible.[5] Furthermore, the coefficients of the direct effects for proportion of negative information on plausibility in the mediation model are stepwise decreasing. Taken together, this likewise hints at a linear relationship between proportion of negative information and plausibility. Thus, H2a is supported.

The direct effect of trustworthiness on CSR perception shows that participants who perceived the company as trustworthy also judged it as more socially responsible, which provides support for H3. In contrast to H4, there was no direct negative effect of proportion of negative information on CSR perception.

4.3. Additional analysis

Analyzing the total effect on CSR perception, which incorporates the positive indirect effect of trustworthiness, allowed us to assess how CSR perception develops when the proportion of moderately negative information is stepwise increased. The results show that each of the confidence intervals includes zero. This means that there was neither a positive nor a negative total effect on CSR perception for any of the groups that received negative information.

---

[5] Repeating the mediation analysis and treating proportion of negative information as a continuous variable (percentage values) showed a significant effect of proportion of negative information on plausibility (coeff. = −2.24, $p < 0.01$).

Table 4
ANCOVA with proportion of negative information as independent variable and trustworthiness as dependent variable.

I. Analysis of covariance

Source	Type III sum of squares	df	Mean square	F	Partial eta squared
Corrected model	96.0	6	16.00	17.88*	0.26
Intercept	71.94	1	71.94	80.36*	0.21
Proportion of negative information	7.46	3	2.49	2.78*	0.03
Plausibility	33.56	1	33.56	37.49*	0.11
Sensitivity to social performance	6.68	1	6.68	7.46*	0.02
Trust disposition	13.62	1	13.62	15.22*	0.05
Error	278.39	311	0.90		

II: Pairwise comparison

Group comparison[a]	Mean difference	SE	95% CI	
			Lower bound	Upper bound
0% vs. 20%	−0.27	0.15	−0.57	0.02
0% vs. 40%	−0.31*	0.16	−0.62	−0.01
0% vs.60%	−0.44*	0.16	−0.74	−0.13
20% vs. 40%	−0.04	0.16	−0.35	0.27
20% vs. 60%	−0.16	0.16	−0.47	0.14
40% vs. 60%	−0.12	0.15	−0.42	0.18

Note.
* $p < 0.05$.
[a] Percentage values refer to the proportion of negative information within the disclosure.

5. Discussion and conclusion

5.1. Theoretical implications and contributions

With respect to trustworthiness, we found two opposing effects caused by the disclosure of moderately negative CSR information. First, we found a direct positive effect, which is in line with attribution theory (Crowley & Hoyer, 1994) and previous research on two-sided messages (Eisend, 2006). While our results provide some indications that this effect might be linear, we cannot confirm it, as no significant differences between the various groups that received moderately negative information were found. There may be different reasons that we did not find such an effect. As mentioned, the effects of two-sided messages are often context-specific (Eisend, 2006; Golden, 1979). In the given case, the reason might have been the way of manipulating negative information. Unachieved objectives only represent moderately negative information, and thus the increase in trustworthiness is likely to be less pronounced. Second, the disclosure of negative information has a

Table 5
ANCOVA with proportion of negative information as independent variable and plausibility as dependent variable.

I. Analysis of covariance

Source	Type III sum of squares	df	Mean square	F	Partial eta squared
Corrected model	159.30	5	31.86	17.34*	0.22
Intercept	104.02	1	104.02	56.78*	0.15
Proportion of negative information	81.01	3	27.00	14.70*	0.12
Sensitivity to social performance	5.68	1	5.68	3.01	0.01
Trust disposition	47.48	1	47.48	25.84*	0.08
Error	573.32	312	1.84		

II: Pairwise comparison

Group comparison[a]	Mean difference	SE	95% CI	
			Lower bound	Upper bound
0% vs. 20%	0.33	0.22	−0.10	0.75
0% vs. 40%	0.93*	0.22	0.50	1.35
0% vs.60%	1.28*	0.21	0.87	1.70
20% vs. 40%	0.60*	0.22	0.17	1.03
20% vs. 60%	0.96*	0.22	0.53	1.38
40% vs. 60%	0.36	0.22	−0.07	0.79

Note.
* $p < 0.05$.
[a] Percentage values refer to the proportion of negative information within the disclosure.

negative indirect effect on trustworthiness. This is because the publication of negative information is perceived as implausible, which, in turn, negatively affects trustworthiness. For this effect, our data provides support for a linear relationship.

These two opposing effects have an essential implication. Both are based on similar mechanisms. The negative effect is mainly caused by plausibility, which was defined as deviance from previous expectations (Mercer, 2004). The positive effect rests on attribution theory. However, the attribution of a disposition only occurs if uncommon behavior is observed (Jones & Davis, 1965), such as the voluntary disclosure of negative information. Disentangling these effects enhances our understanding of two-sided messages and thus represents a noteworthy contribution to this research stream.

If the increase in trustworthiness, which might be linear according to attribution theory, is offset by a negative linear effect caused by missing plausibility, then higher proportions of negative information do not lead to evermore trustworthiness. Instead, when both effects overlap, the total effect might follow, in an ideal scenario, an inverted u-shape. This implies the existence of a threshold beyond which trustworthiness does not further increase if more negative information is issued. Thus, the identification of plausibility as further variable that affects trustworthiness is an important contribution, as it helps clarify previous findings that did not indicate an increase in credibility for higher proportions of negative information (Anderson & Golden, 1984; Golden & Alpert, 1987).

We found that trustworthiness positively affects CSR perception. This is in line with past empirical studies on this relationship (Alcañiz et al., 2010; Schmeltz, 2012; Swaen & Vanhamme, 2004). However, contrary to what we expected, we found neither a negative direct nor a negative total effect of any proportion of moderately negative information on CSR perception. One explanation for this rather surprising result might be found in our experimental material. Several authors found a buffer or insurance effect of a long CSR history and good previous CSR evaluations against negative events (Klein & Dawar, 2004; Vanhamme & Grobben, 2009). While the company description was intended to be moderately credible (Lafferty, 2007), it might have been perceived as sufficiently positive to trigger such an effect and thus prevent a negative CSR perception. This would imply that the reactions to two-sided CSR disclosures are more adverse in the case of companies that have low CSR reputations or if the information itself is more negative than that used in the current study.

In addition to the general understanding of two-sided messages, this study contributes in particular to the scarce literature investigating the effect of disclosing negative CSR-related information. While previous studies focused on investment decisions (Chan & Milne, 1999; Reimsbach & Hahn, 2015), our study takes a different angle because it focuses on trustworthiness and CSR perception. By doing so, we confirm that there are no adverse effects due to disclosing negative information and help to explain previous results (investors that do not perceive a company as less responsible are unlikely to decrease stock price assessments) (Reimsbach & Hahn, 2015). Trustworthiness is a crucial variable in two-sided advertising studies; however, it was previously neglected when investigating two-sided CSR disclosures despite its theoretical connection to CSR (Schwartz & Carroll, 2008).

5.2. Managerial implications

From a managerial point of view, different implications follow. First of all, in the given case the disclosure of moderately negative information does not adversely affect CSR perception, which should reduce managers' reluctance to disclose negative information. This is highly relevant in order to comply with sustainable reporting standards such as the Global Reporting Initiative (GRI) that, through its "Balance"

principle (GRI, 2016), demands companies to disclose positive as well as negative information in order to provide an unbiased picture of the organization's performance. Ideally, organizations should not only report unachieved objectives they should describe and explain reasons for non-achievements, and, additionally, announce how they are going to reach the objectives in the future (Hahn & Lülfs, 2014). Furthermore, the voluntary disclosure of moderately negative information might cause positive long-term effects through higher trustworthiness by enabling more reliable and stronger relationships with stakeholders, which, in turn, may result in financial success (Garcia-Castro, Ariño, & Canela, 2011; Schnackenberg & Tomlinson, 2016).

Moreover, as missing plausibility reduces the trustworthiness increase caused by the disclosure of moderately negative information, it seems useful to establish a comprehensive voluntary disclosure policy, which should include regular, e.g., quarterly, disclosure independent of current success or failure as well as immediate disclosures in case of uncommon events. Such a policy seems necessary in order to accustom readers to the company's willingness to disclose negative information. Hence, such a behavior would not any longer be unexpected (i.e., regarded implausible) and no adverse effect on trustworthiness occurs. Again, it is likely that such a policy primarily pays off in the long run, as it will take some time to accustom the readers to it. If plausibility and trustworthiness increase, credibility should rise. Credibility, in turn, is positively related to other crucial variables, such as legitimacy (Jahn, Eichhorn, & Brühl, 2017), trust (Brühl, Basel, & Kury, 2018) and attitude toward the brand or purchase intentions (Goldsmith, Lafferty, & Newell, 2000); thus, it is likely that there are further positive effects of two-sided messages that are not covered in the current study.

5.3. Limitations and future research

We measured the negative effects of negative information only in terms of CSR perception; however, there might be other factors inhibiting the frank disclosure of negative incidents, such as financial penalties for admitted ecological crimes. The investigation of these and other legal issues, such as a comparison between voluntary and mandatory disclosure, could be a valuable area for future research.

Furthermore, as operationalization of negative information, we referred to objectives that were not met by the company and thus referred to moderately negative information. While we feel confident that our operationalization is appropriate—the manipulation check confirms different perceptions of negativity—we are also aware that this is different from negative events such as oil spills. The communication of such information might cause different reactions and is therefore a valuable objective for future research.

As indicated above, investigating the effect of two-sided messages on further dependent variables such as legitimacy is worthwhile in order to test whether two-sided messages provide additional benefits caused by increased trustworthiness. Moreover, the situation of companies that disclose two-sided messages can be analyzed. Are the various effects stronger or weaker for companies with great (vs. poor) financial or social performance?

Appendix A. Letter to the stakeholders, translated version (40% moderately negative information)

Appendix B

Table B1
Mean, Standard Deviation, and loadings of the items, Cronbachs α and AVE for constructs.

Items	M	SD	Factor loading
Trustworthiness (α = 0.86, AVE = 0.82)			
I trust the LAINKO AG (Trustw1).	5.06	1.12	0.90
I do not believe what the LAINKO AG tells me (Trustw2).	5.20	1.56	0.54
			(continued on next page)

Table B1 (*continued*)

Items	M	SD	Factor loading
The LAINKO AG is honest (Trustw3).	5.02	1.20	0.93
The LAINKO AG makes truthful claims (Trustw4).	4.68	1.26	0.86
CSR perception (α = 0.92, AVE = 0.84)			
Lainko AG is trying to sponsor pro-environmental programmes (CSR1).	5.26	1.26	0.87
Lainko AG is trying to allocate resources to offer services compatible with the environment (CSR2).	4.88	1.35	0.84
Lainko AG is trying to carry out programmes to reduce pollution (CSR3).	5.45	1.16	0.84
Lainko AG is trying to protect the environment (CSR4).	5.27	1.22	0.92
Lainko AG is trying to recycle its waste materials properly (CSR5).	4.96	1.33	0.69
Trust disposition (α = 0.87, AVE = 0.87)			
I think most people keep their promises (DiTru1).	4.02	1.07	0.81
In general you can rely on other people (DiTru2).	4.18	1.28	0.89
In principle, other people can be trusted (DiTru3).	4.15	1.36	0.90
Sensitivity to social performance (α = 0.87, AVE = 0.64)			
I would be willing to pay a little more to buy a product from a company that has good environmental practices (Sens1).	5.12	1.55	0.92
I would be willing to pay a little more to buy a product from a company that has a good record in hiring and promoting minority groups (Sens5).	5.17	1.49	0.80
It would bother me to be employed by a company with a poor reputation for social responsibility (Sens6).	5.20	1.47	0.60
I would be willing to pay a little more to buy a product from a company that invests in and creates new jobs rather than downsizing (Sens8).	4.86	1.54	0.85

Table B2
Cross-loadings to assess discriminant validity.

Construct	Item	CSR	TW	S	TD	P
CSR	CSR1	**0.90**	−0.02	0.00	−0.01	0.01
	CSR2	**0.83**	0.03	0.01	0.01	0.10
	CSR3	**0.74**	0.11	0.05	−0.04	−0.06
	CSR4	**0.85**	0.06	0.02	0.01	−0.17
	CSR5	**0.77**	−0.09	−0.01	0.04	0.14
TW	Trustw1	0.01	**0.89**	0.01	0.00	−0.02
	Trustw2	−0.06	**0.59**	0.13	−0.08	0.09
	Trustw3	0.00	**0.94**	−0.01	0.01	0.04
	Trustw4	0.08	**0.77**	−0.02	0.07	−0.09
S	Sens1	0.03	−0.06	**0.91**	0.03	0.01
	Sens5	0.03	0.04	**0.80**	−0.05	−0.07
	Sens6	−0.08	0.04	**0.60**	0.01	0.21
	Sens8	0.00	0.04	**0.82**	0.03	−0.03
TD	DiTru1	0.04	0.06	0.02	**0.76**	0.11
	DiTru2	0.00	−0.04	−0.02	**0.92**	−0.03
	DitTru3	−0.03	0.03	0.04	**0.88**	−0.04
P	Plausibility	0.19	0.24	−0.03	0.18	**1**

Note. CSR = CSR perception; S = sensitivity to social performance; TW = trustworthiness; TD = trust disposition; P = plausibility. Bold faced values indicate which factors the items load on.

References

Alcañiz, E. B., Cáceres, R. C., & Pérez, R. C. (2010). Alliances between brands and social causes: The influence of company credibility on social responsibility image. *Journal of Business Ethics, 96*(2), 169–186. https://doi.org/10.1007/s10551-010-0461-x.

Alvarado-Herrera, A., Bigne, E., Aldas-Manzano, J., & Curras-Perez, R. (2017). A scale for measuring consumer perceptions of corporate social responsibility following the sustainable development paradigm. *Journal of Business Ethics, 140*(2), 243–262. https://doi.org/10.1007/s10551-015-2654-9.

Anderson, W. T., & Golden, L. L. (1984). Bank promotion strategy. *Journal of Advertising Research, 24*(2), 53–65.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory, 5*(3), 241. https://doi.org/10.1037/0278-7393.5.3.241.

Barton, J., & Mercer, M. (2005). To blame or not to blame: Analysts' reactions to external explanations for poor financial performance. *Journal of Accounting and Economics, 39*(3), 509–533. https://doi.org/10.1016/j.jacceco.2005.04.006.

Brown, T. J., & Dacin, P. A. (1997). The company and the product: Corporate associations and consumer product responses. *Journal of Marketing, 61*(1), 68–84. https://doi.org/10.2307/1252190.

Brühl, R., Basel, J. S., & Kury, M. F. (2018). Communication after an integrity-based trust violation: How organizational account giving affects trust. *European Management Journal, 36*(2), 161–170. https://doi.org/10.1016/j.emj.2017.08.001.

Chan, C. C., & Milne, M. J. (1999). Investor reactions to corporate environmental saints and sinners: An experimental analysis. *Accounting and Business Research, 29*(4), 265–279. https://doi.org/10.1080/00014788.1999.9729588.

Chebat, J.-C., & Picard, J. (1988). Receivers' self-acceptance and the effectiveness of two-sided messages. *Journal of Social Psychology, 128*(3), 353–362. https://doi.org/10.1080/00224545.1988.9713752.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Crowley, A. E., & Hoyer, W. D. (1994). An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research, 20*(4), 561–574. https://doi.org/10.1086/209370.

Donaldson, T. (1982). *Corporations and morality.* Englewood Cliffs, NJ: Prentice-Hall.

Eisend, M. (2006). Two-sided advertising: A meta-analysis. *International Journal of Research in Marketing, 23*(2), 187–198. https://doi.org/10.1016/j.ijresmar.2005.11.001.

Eisend, M. (2010). Explaining the joint effect of source credibility and negativity of information in two-sided messages. *Psychology and Marketing, 27*(11), 1032–1049. https://doi.org/10.1002/mar.20372.

Ellen, P. S., Webb, D. J., & Mohr, L. A. (2006). Building corporate associations: Consumer attributions for corporate socially responsible programs. *Journal of the Academy of Marketing Science, 34*(2), 147–157. https://doi.org/10.1177/0092070305284976.

Federal Statistical Office (2015). *Zensus 2011: Zensus Kompakt – Endgültige Ergebnisse.* Stuttgart: Federal Statstical Office.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research,* 39–50. https://doi.org/10.2307/3151312.

Friedman, M. (1970). *The social responsibility of business is to increase its profits. 1970*New York Times Magazine (September 13).

Garcia-Castro, R., Ariño, M. A., & Canela, M. A. (2011). Over the long-run? Short-run impact and long-run consequences of stakeholder management. *Business & Society, 50*(3), 428–455. https://doi.org/10.1177/0007650308315493.

Golden, L. L. (1977). Attribution theory implications for advertisement claim credibility. *Journal of Marketing Research, 14*(1), 115–117. https://doi.org/10.2307/3151069.

Golden, L. L. (1979). Consumer reactions to explicit brand comparisons in advertisements. *Journal of Marketing Research, 16*(4), 517–532. https://doi.org/10.2307/

3150812.

Golden, L. L., & Alpert, M. I. (1987). Comparative analysis of the relative effectiveness of one-and two-sided communication for contrasting products. *Journal of Advertising, 16*(1), 18–25. https://doi.org/10.1080/00913367.1987.10673056.

Goldsmith, R. E., Lafferty, B. A., & Newell, S. J. (2000). The impact of corporate credibility and celebrity credibility on consumer reaction to advertisements and brands. *Journal of Advertising, 29*(3), 43–54. https://doi.org/10.1080/00913367.2000. 10673616.

GRI (2016). *GRI 102: General disclosures.* Amsterdam: Global Reporting Initiative.

Hackston, D., & Milne, M. J. (1996). Some determinants of social and environmental disclosures in New Zealand companies. *Accounting, Auditing & Accountability Journal, 9*(1), 77–108. https://doi.org/10.1108/09513579610109987.

Hahn, R., & Lülfs, R. (2014). Legitimizing negative aspects in GRI-oriented sustainability reporting: A qualitative analysis of corporate disclosure strategies. *Journal of Business Ethics, 123*(3), 401–420. https://doi.org/10.1007/s10551-013-1801-4.

Hair, J., Black, W., Babin, A., & Anderson, R. (2014). *Multivariate data analysis* (7th ed.). New Jersey, Essex: Pearson.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York: Guilford Publications.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Holder-Webb, L., Cohen, J. R., Nath, L., & Wood, D. (2009). The supply of corporate social responsibility disclosures among US firms. *Journal of Business Ethics, 84*(4), 497–527. https://doi.org/10.1007/s10551-008-9721-4.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; Psychological studies of opinion change.* New Haven, CT: Yale University Press.

Hur, W.-M., Kim, H., & Woo, J. (2014). How CSR leads to corporate brand equity: Mediating mechanisms of corporate brand credibility and reputation. *Journal of Business Ethics, 125*(1), 75–86. https://doi.org/10.1007/s10551-013-1910-0.

Hutton, A. P., Miller, G. S., & Skinner, D. J. (2003). The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research, 41*(5), 867–890. https://doi.org/10.1046/j.1475-679X.2003.00126.x.

Jahn, J., Eichhorn, M., & Brühl, R. (2017). How do individuals judge organizational legitimacy? Effects of attributed motives and credibility on organizational legitimacy. *Business & Society.* https://doi.org/10.1177/0007650317717959 (online first).

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions the attribution process in person perception. In L. Berkowitz (Ed.). *Advances in experimental social psychology* (pp. 219–266). (2nd ed.). New York: Academic Press.

Kelley, H. H. (1971). *Attribution in social interaction.* Morristown, NJ: General Learning Press.

Klein, J., & Dawar, N. (2004). Corporate social responsibility and consumers' attributions and brand evaluations in a product–harm crisis. *International Journal of Research in Marketing, 21*(3), 203–217. https://doi.org/10.1016/j.ijresmar.2003.12.003.

Korschun, D., Bhattacharya, C. B., & Swain, S. D. (2014). Corporate social responsibility, customer orientation, and the job performance of frontline employees. *Journal of Marketing, 78*(3), 20–37. https://doi.org/10.1509/jm.11.0245.

Lafferty, B. A. (2007). The relevance of fit in a cause–brand alliance when consumers evaluate corporate credibility. *Journal of Business Research, 60*(5), 447–453. https://doi.org/10.1016/j.jbusres.2006.09.030.

Leonidou, C. N., & Skarmeas, D. (2017). Gray shades of green: Causes and consequences of green skepticism. *Journal of Business Ethics, 144*(2), 401–415. https://doi.org/10. 1007/s10551-015-2829-4.

Lin-Hi, N., & Müller, K. (2013). The CSR bottom line: Preventing corporate social irresponsibility. *Journal of Business Research, 66*(10), 1928–1936. https://doi.org/10. 1016/j.jbusres.2013.02.015.

Liu, T. C., Wang, C. Y., & Wu, L. W. (2010). Moderators of the negativity effect: Commitment, identification, and consumer sensitivity to corporate social performance. *Psychology and Marketing, 27*(1), 54–70. https://doi.org/10.1002/mar.20319.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review, 23*(3), 473–490. https://doi.org/10.5465/amr.1998.926622.

Mercer, M. (2004). How do investors assess the credibility of management disclosures? *Accounting Horizons, 18*(3), 185–196. https://doi.org/10.2308/acch.2004.18.3.185.

Mizerski, R. W., Golden, L. L., & Kernan, J. B. (1979). The attribution process in consumer decision making. *Journal of Consumer Research, 6*(2), 123–140. https://doi.org/10. 1086/208756.

Newell, S. J., & Goldsmith, R. E. (2001). The development of a scale to measure perceived corporate credibility. *Journal of Business Research, 52*(3), 235–247. https://doi.org/ 10.1016/S0148-2963(99)00104-6.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52. https://doi.org/10.1080/00913367.1990.10673191.

Parguel, B., Benoît-Moreau, F., & Larceneux, F. (2011). How sustainability ratings might deter 'greenwashing': A closer look at ethical corporate communication. *Journal of Business Ethics, 102*(1), 15–28. https://doi.org/10.1007/s10551-011-0901-2.

Perks, K., Farache, F., Shukla, P., & Berry, A. (2013). Communicating responsibility-practicing irresponsibility in CSR advertisements. *Journal of Business Research, 66*(10), 1881–1888. https://doi.org/10.1016/j.jbusres.2013.02.009.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method

biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879. https://doi.org/10.1037/0021-9010.88.5.879.

Reimsbach, D., & Hahn, R. (2015). The effects of negative incidents in sustainability reporting on investors' judgments—An experimental study of third-party versus self-disclosure in the realm of sustainable development. *Business Strategy and the Environment, 24*(4), 217–235. https://doi.org/10.1002/bse.1816.

Rim, H., & Song, D. (2016). "How negative becomes less negative": Understanding the effects of comment valence and response sidedness in social media. *Journal of Communication, 66*(3), 475–495. https://doi.org/10.1111/jcom.12205.

Scalet, S., & Kelly, T. F. (2010). CSR rating agencies: What is their global impact? *Journal of Business Ethics, 94*(1), 69–88. https://doi.org/10.1007/s10551-009-0250-6.

Schlegelmilch, B. B., & Pollach, I. (2005). The perils and opportunities of communicating corporate ethics. *Journal of Marketing Management, 21*(3–4), 267–290. https://doi.org/10.1362/0267257053779154.

Schmeltz, L. (2012). Consumer-oriented CSR communication: Focusing on ability or morality? *Corporate Communications: An International Journal, 17*(1), 29–49. https://doi.org/10.1108/13563281211196344.

Schnackenberg, A. K., & Tomlinson, E. C. (2016). Organizational transparency: A new perspective on managing trust in organization-stakeholder relationships. *Journal of Management, 42*(7), 1784–1810. https://doi.org/10.1177/0149206314525202.

Schwartz, M. S., & Carroll, A. B. (2008). Integrating and unifying competing and complementary frameworks. The search for a common core in the business and society field. *Business & Society, 47*(2), 148–186. https://doi.org/10.1177/0007650306297942.

Settle, R. B., & Golden, L. L. (1974). Attribution theory and advertiser credibility. *Journal of Marketing Research,* 181–185. https://doi.org/10.2307/3150556.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422.

Skarmeas, D., & Leonidou, C. N. (2013). When consumers doubt, watch out! The role of CSR skepticism. *Journal of Business Research, 66*(10), 1831–1838. https://doi.org/10. 1016/j.jbusres.2013.02.004.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*(1), 131–142. https://doi.org/10.1037/0033-2909.105.1.131.

Smith, R. E., & Hunt, S. D. (1978). Attributional processes and effects in promotional situations. *Journal of Consumer Research, 5*(3), 149–158. https://doi.org/10.1086/208725.

Swaen, V., & Vanhamme, J. (2004). See how 'good' we are: The dangers of using corporate social activities in communication campaigns. In B. E. Kahn, & M. F. Luce (Eds.). *Advances in consumer research* (pp. 302–303). (31st ed.). .

Van der Laan Smith, J., Adhikari, A., & Tondkar, R. H. (2005). Exploring differences in social disclosures internationally: A stakeholder perspective. *Journal of Accounting and Public Policy, 24*(2), 123–151. https://doi.org/10.1016/j.jaccpubpol.2004.12.007.

Vanhamme, J., & Grobben, B. (2009). "Too good to be true!" The effectiveness of CSR history in countering negative publicity. *Journal of Business Ethics, 85*(2), 273–283. https://doi.org/10.1007/s10551-008-9731-2.

Wagner, T., Lutz, R. J., & Weitz, B. A. (2009). Corporate hypocrisy: Overcoming the threat of inconsistent corporate social responsibility perceptions. *Journal of Marketing, 73*(6), 77–91. https://doi.org/10.1509/jmkg.73.6.77.

Wiedmann, K., Wüstefeld, T., & Klibert, T. (2011). *Finanzkommunikation börsennotierter Unternehmen.* Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.

Williams, P. A. (1996). The relation between a prior earnings forecast by management and analyst response to a current management forecast. *Accounting Review, 71*(1), 103–115.

Wood, R. E., Goodman, J. S., Beckmann, N., & Cook, A. (2008). Mediation testing in management research: A review and proposals. *Organizational Research Methods, 11*(2), 270–295. https://doi.org/10.1177/1094428106297811.

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197–206. https://doi.org/10.1086/651257.

Johannes Jahn is a research assistant at the Chair of Business Ethics and Management Control at ESCP Europe Business School, Berlin. He holds a doctoral degree in business administration (Dr. rer. pol., ESCP Europe Business School, Berlin) and his research interests include business ethics, corporate social responsibility and organizational legitimacy.

Rolf Brühl (Dr. rer. oec., Technical University Berlin) is full professor of Business Ethics and Management Control at ESCP Europe Business School, Berlin. His research interests focus on corporate social responsibility, social accounts, legitimacy and international management control systems. His articles have appeared in such journals as *European Journal of Innovation Management, European Management Journal, International Journal of Business Communication, Journal of Business Economics, Business & Society* and *Journal of Business Ethics.*