



The Effects of Counterfactual Attacks on the Morality and Leadership of Different Professionals

Mauro Bertolotti and Patrizia Catellani

Department of Psychology, Catholic University of Milan, Italy

Abstract: Past research has offered contrasting results regarding the effects of attacks on social judgments. In three experiments, we investigated the effects of counterfactual (“If only...”) and non-counterfactual attacks on the morality versus leadership of politicians versus entrepreneurs. First, participants rated morality as the most desirable, but least typical dimension of politicians, and leadership as the most desirable and most typical dimension of entrepreneurs (Study 1). Then, counterfactual attacks led to poorer evaluation of the attacked target and better evaluation of the attacking source as compared to non-counterfactual attacks, especially when counterfactuals were focused on the most desirable dimension for the professional category of the attacked target (Study 2). Similar results emerged when the typicality of the attacked dimension was manipulated (Study 3). Discussion focuses on the higher success of attacks on desirable personality dimensions and of counterfactual attacks as compared to other attacks.

Keywords: counterfactual communication, morality, leadership, criticism, social judgment

Imagine that during an interview to a Prime Minister regarding governmental action a journalist says something like: “If you had really acted in the interest of voters, the economic situation of our country *would be better* now.” A statement like this is an example of *counterfactual attack*, that is, an attack based on the mental simulation of a different behavior of an actor (in this case the Prime Minister) resulting in a better outcome than the real one. Past research has investigated the role of counterfactual thinking in several cognitive processes, such as causal explanation (Kahneman & Miller, 1986; Roese, 1997), attribution (Catellani, Alberici & Milesi, 2004; Mandel, Hilton, & Catellani, 2005), and self-regulation (Epstude & Roese, 2008; Roese & Epstude, 2017). Few studies, however, have analyzed its effects on impression formation (Tal-Or, Boninger, Poran, & Gleicher, 2004; Wong, 2010) or impression management (Bertolotti, Catellani, Douglas, & Sutton, 2013; Catellani & Bertolotti, 2014a, 2014b; Catellani & Covelli, 2013).

Attacks often backfire due to their negativity being rejected by recipients (Roese & Sande, 1993). Counterfactual attacks have an enhanced rhetorical potential in this regard, as they offer the possibility of putting the outcome of someone’s behavior in a negative light just by comparing it with a hypothetical more positive condition, rather than disputing its actual negativity. Counterfactual attacks can affect social judgments through two possible pathways

(Catellani & Bertolotti, 2014a, 2015). The first pathway is content-specific and is based on how counterfactuals affect recipients’ perception of past events and the actors involved in them. By describing how an actor could or should have behaved, counterfactual attacks refer to specific normative standards and promote congruent responsibility attributions. The second pathway is content-neutral and is based on how counterfactual messages and their sources are perceived. By describing a hypothetical scenario leading to a more positive scenario, a counterfactual attack (and its source) are perceived as less negative and less biased than other more explicitly negative forms of criticism.

Previous research (Catellani & Bertolotti, 2014a) has shown that counterfactual attacks on politicians’ morality affect the politicians’ reputation more than non-counterfactual attacks. In the present research, we further investigated the conditions under which counterfactual attacks can be effective tools for impression management, focusing on whether counterfactual attacks can be used to stress a target’s deviation from a relevant norm regarding the behavior of the professional category the target belongs to. We analyzed the effects of counterfactual and non-counterfactual attacks with different content (morality- vs. leadership-based) and against different targets (a politician vs. an entrepreneur). We expected counterfactual attacks to be most effective when based on a relevant content dimension

in the evaluation of the target as a member of a certain professional category.

Counterfactual Attacks

Attacks can be defined as messages providing negative information about specific targets and their behavior. Past research showed that when negative information is provided in the form of an attack, its impact is often limited. This is the case because attack messages often backfire (Carraro, Gawronski, & Castelli, 2010; Funk, 1996; Haddock & Zanna, 1997; Lau, Sigelman, & Rovner, 2007; Roese & Sande, 1993), as recipients can easily detect the persuasive intent of their source, and consequently discount the information provided with the message (Douglas & Sutton, 2010; Hornsey, 2005). To reduce the backlash effect, speakers sometimes resort to more indirect types of criticism (Hornsey, Robson, Smith, Esposo, & Sutton, 2008). Counterfactuals can be employed for this purpose, as some studies have shown (Catellani & Covelli, 2013; Wong, 2010).

Counterfactuals are mental simulations in which an antecedent of an event is mutated in order to hypothetically alter the outcome of the event (Kahneman & Tversky, 1982; Roese, 1997). For example, after a car crash one might state that the accident would have been avoided, “if the driver had stopped in time.” Several studies on counterfactual thinking have shown that individuals often use this type of reasoning to make causal and responsibility attributions (Markman & Tetlock, 2000; Nario-Redmond & Branscombe, 1996; Wells & Gavanski, 1989). Counterfactuals can be used also in communication as a form of indirect or subtle criticism, to imply that the actions of a given actor are the cause of a certain outcome, and that the actor is therefore responsible for it. According to Catellani and Bertolotti (2014a, 2015), counterfactual attacks can influence recipients’ social judgments along two distinct but converging pathways (resembling the dual-pathway model proposed by Epstude & Roese, 2008).

The first pathway regards the way recipients appraise the events discussed in a counterfactual attack, and the actors involved in them. Spontaneously generated counterfactuals tend to mutate controllable, rather than uncontrollable behaviors (Alicke, Buckingham, Zell, & Davis, 2008; Giroto, Legrenzi, & Rizzo, 1991), and to focus on the actor which is deemed responsible for the final outcome (Markman & Tetlock, 2000; Nario-Redmond & Branscombe, 1996; Wells & Gavanski, 1989). Besides, counterfactual thoughts usually focus on behaviors that are perceived as deviating from *normality* (Davis & Lehman, 1995; Kahneman & Miller, 1986). Initial research focused on deviations from *routine norms* (e.g., “If the driver had

taken the usual route, he/she would not have been involved in the accident”; Gavanski & Wells, 1989; Kahneman & Miller, 1986; Kahneman & Tversky, 1982; Klauer, Jacobsen, & Migulla, 1995; Wells, Taylor, & Turtle, 1987). Later research extended this effect to deviation from *social norms*, showing that people are more likely to generate counterfactuals when they perceive the actor’s behavior to deviate from what is expected from members of a given social group or category (e.g., “If the victim had not accepted a lift from a stranger, she would have not been sexually assaulted”; Branscombe, Crosby, & Weir, 1993; Branscombe & Weir, 1992; Catellani et al., 2004; Hegarty & Pratto, 2001). In consideration of the above, by using a counterfactual attack a speaker highlights that the target has intentionally deviated from a social norm. In turn, recipients exposed to a counterfactual attack are likely to assume that the negative outcome could have been prevented, if the target had acted in a different way.

The second pathway along which counterfactual attacks can influence social judgments regards the way recipients appraise attacks, and their source, when determining whether they can base their judgment on it. By comparing reality with a better hypothetical alternative, rather than stressing its negativity, counterfactual attacks have reduced illocutionary force (Thaler, 2012) compared to plain, explicit criticism, thus making the speaker’s communicative intention less apparent. This can significantly reduce backlash effect. Past research has indeed shown that recipients tend to perceive counterfactual attacks as less hostile and more appropriate messages than non-counterfactual attacks (Catellani & Bertolotti, 2013, 2014b), and are consequently less inclined to discount information coming from the source as biased (Bavelas, Black, Bryson, & Mullett, 1988; Fiedler & Mata, 2013; Hornsey et al., 2008).

Based on the above, one can reasonably expect recipients of a counterfactual attack to be, in general, more likely to consider the information it conveys as compared to recipients of a non-counterfactual attack. More specifically, they are likely to use the counterfactual antecedent as a reference standard to evaluate the behavior of the attacked target. The evoked reference norm regarding how one should have behaved, however, may or may not be relevant for the specific target of an attack. For example, an effective attack to an entrepreneur would be saying that if he or she had outsold competition, a company would have been more profitable, because pursuing market dominance is what one usually expects from an entrepreneur. Conversely, a less compelling attack would be saying that if an entrepreneur had sought cooperation with competing companies, his or her company would have been more profitable, because pursuing harmony with competitors is regarded as less relevant in entrepreneurial behavior (see Maak, Pless, & Voegtlin, 2016).

In the present research we compared the effects of counterfactual and non-counterfactual attacks on different professionals, investigating the interplay between the social norm used to evaluate the behavior of different professionals, and the content dimension on which the attack is based. We expected counterfactual attacks to be more effective than non-counterfactual attacks. In particular, we expected a content-specific effect, with counterfactual attacks leading to a negative evaluation of the target only when their content would regard a dimension of social judgment deemed relevant for the target's professional category. We also expected a content-neutral effect, with counterfactual attacks being perceived in general as more convincing and appropriate than non-counterfactual attacks.

Content Dimensions in Social Judgment

A host of research on social judgment has consistently showed that people evaluate other individuals along two main content dimensions: *agency*, which refers to the person's proactivity and ability to foster one's intentions, and *communion*, which refers to the person's relationships with other persons, as well as compliance with social norms governing these relationships (Abele & Wojciszke, 2007; Fiske, Cuddy, & Glick, 2007; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005).

Within these two main dimensions, further sub-dimensions have been identified (Abele et al., 2016; Bertolotti et al., 2013; Carrier, Louvet, Chauvin, & Rohmer, 2014). The agency dimension includes both *competence*, related to traits such as being competent and intelligent, and *dominance* or *leadership*, related to traits such as being decided and resolved. The communion dimension includes both *sociability*, related to traits such as being sociable and friendly, and *morality*, related to traits such as being honest and sincere (Leach, Ellemers, & Barreto, 2007).

The different content dimensions are not equally regarded in the impression formation process. Several studies found morality to be generally more important in forming impressions of other individuals and groups (Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Leach et al., 2007; Leach, Minescu, Poppe, & Hagendoorn, 2008; Pagliaro, Brambilla, Sacchi, D'Angelo, & Ellemers, 2013), although such primacy can be moderated by normative expectations regarding what is deemed as typical and desirable in a specific context.

Stereotypes provide individuals with information on how frequently members of a given group engage in behaviors,

and possess traits, related to different content dimensions. Several studies have investigated ambivalent stereotypes (Cuddy, Fiske, & Glick, 2008) regarding the typicality of the fundamental content dimensions of social judgment among members of different groups, including different professionals (Asbrock, 2010; Cislak, 2013; Imhoff, Woelki, Hanke, & Dotsch, 2013). Based on widespread stereotypical expectations, for example, one might consider atypical an entrepreneur described as incompetent and indecisive. Conversely, negative information regarding an entrepreneur's morality might not be used by recipients to determine whether the target is typical or atypical within the professional category.

Another kind of normative expectation regards how desirable certain behaviors and traits are for members of a social category. Although there are less studies on the desirability of the fundamental content dimensions in social groups, past research has found that what we expect to obtain from individuals and groups (i.e., our *self-interest*) often determines the desirability of traits related to the achievement of that goal (Scholer & Higgins, 2008; Wojciszke, 2005). This is especially evident when evaluating individuals in a professional setting, where we expect someone doing a certain job to have the traits that make them most suitable to perform well in that job, to the advantage of potential employers, employees, clients, or the whole community. In the case of entrepreneurs, a high level of leadership might be seen as not only typical, but also desirable for members of that professional group, as their job is mainly based on their ability to advance themselves (and their companies) in the market (Cottrell, Neuberg, & Li, 2007; Mollaret & Miraucourt, 2016; Nichols & Cottrell, 2014). If we consider politicians, a dimension that may be seen as particularly desirable is morality as their job is mainly based on their dedication to representing the interest of their voters, and pursuing the public good (Castelli, Carraro, Ghitti, & Pastore, 2009; Catellani & Bertolotti, 2015; Cislak & Wojciszke, 2006).

The case of politicians' morality is peculiar because it shows that the dimension perceived as most desirable for a professional category is not necessarily perceived also as the most typical. Even if morality-related traits (e.g., honest, sincere, loyal) are considered very desirable, and define the *ideal* profile of a politician, citizens perceive these traits as unfrequently possessed by *actual* politicians, and therefore very atypical (Catellani & Quadrio, 1991; McGraw, Lodge, & Jones, 2002).

In the present research, we considered members of two professional categories, namely a politician or an entrepreneur, as targets of attacks based on two content dimensions, namely morality and leadership. We chose these two professional categories because of their similarities,

as both politicians and entrepreneurs are tasked with organizations to manage, have generally high social status and a public profile. This would make the leadership dimension desirable for both professional categories. The two categories, however, also have a key difference, namely the fact that politicians are supposed to work for the public good, whereas entrepreneurs are supposed to work for a private (either their own or their company's) good. This would likely make the morality dimension more desirable for politicians than for entrepreneurs. Based on the above, we expected that the evaluation of politicians and entrepreneurs would be influenced by counterfactual attacks focussed on different content dimensions.

Counterfactual Attacks on Different Content Dimensions

So far, research on the effects that attacks may have on social judgments have yielded mixed results. Some studies have analyzed the effects of attacks as a function of the typicality of the attacked trait for the social or professional category the attacked target person belongs to (McGraw, 2003; Meeks, 2012; Schlenker & Pontari, 2000). However, attacks focused on stereotypical traits of the attacked target's category are rarely more successful than attacks focused on other traits (e.g., Eagly & Karau, 2002; Johns & Shepard, 2007). One reason for this inconsistency may be that these studies only considered the typicality of the attacked dimension, and not its desirability. To our knowledge, however, there is no systematic data available on whether attacks are more successful when they address a feature perceived as desirable versus typical in a target category. Another possibility is that the high likelihood of rejection and backlash of explicit attacks may have obscured the specific effects deriving from the typicality and desirability of the criticized behavior of the target. Such obstacle may be more easily overcome using a counterfactual attack. As discussed above, the upward hypothetical comparison proposed by a counterfactual attack would be particularly effective in inducing participants to compare the target's alleged behavior with the *ideal* behavior expected from members of a specific group.

In the present research, we tested the effects of counterfactual and non-counterfactual attacks against either the leadership or morality dimensions of a fictitious member of two professional categories. We expected counterfactual attacks to be most effective when the content of the allegations against the target matched the content dimension deemed more desirable for the members of the target's category.

Research Overview

In order to test the effectiveness of counterfactual attacks on the morality and leadership of members of two professional categories, namely politicians and entrepreneurs, we conducted three studies.

In Study 1, we tested to what extent morality and leadership were considered *desirable* and *typical* for the professional categories of politicians and entrepreneurs. As anticipated above, in the case of entrepreneurs we expected traits pertaining to the leadership dimension to be regarded both as more desirable and more typical than those pertaining to morality (Nichols & Cottrell, 2014). In the case of politicians, conversely, we expected traits pertaining to the morality dimension to be regarded as more desirable, but not more typical than those pertaining to leadership (Catellani & Quadrio, 1991; Castelli et al., 2009).

In Study 2, we tested the effects of counterfactual and non-counterfactual attacks on a fictional politician or entrepreneur. The attacks criticized either the leadership or the morality dimension of the target, either by stating that the present situation would have been better if the target had behaved in a more positive way in the past (e.g. "If you had kept the promises made to your voters, your party would be in a better condition now"; "If you had disclosed the fiscal disputes of your company, your finances would be in a better condition now") or by directly making allegations of past negative behaviors (e.g., "You broke the promises made to your voters"; "You hid the truth on the fiscal disputes of your company"). We expected counterfactual attacks to influence social judgments more than non-counterfactual attacks, especially when focused on the most desirable dimension for each professional category. This would be the case because a counterfactual attack implicitly contrasts the actual and the expected behavior of the attacked target, resulting in a more negative impression of the target than when the attack is directly focused on the negative behavior of the target, as in a non-counterfactual attack.

Finally, in Study 3, we again tested the influence on social judgment of the counterfactual attacks employed in Study 2, manipulating also the typicality of the morality and leadership dimensions within the professional category of the target. The aim of this study was to test whether a counterfactual attack against a desirable dimension would be equally effective when such dimension was presented as typical or atypical among members of the target's group. Such a result would further confirm that the success of a counterfactual attack is more related to the desirability of the attacked dimension than to its perceived typicality in the relevant social category.

Study 1

In Study 1, we explored to what extent morality and leadership were considered typical and desirable for the members of two different professional categories, politicians and entrepreneurs. According to the results of previous research on the perception of politicians (Catellani & Quadrio, 1991), we expected diverging assessments of the typicality of the two dimensions among the two professional categories: whereas leadership would be regarded as typical among both politicians and entrepreneurs (Nichols & Cottrell, 2014; Cislak & Wojciszke, 2008), we expected morality to be considered quite atypical among politicians.

Method

Participants were 145 students of the Catholic University of Milan (68.2% females, $M_{\text{age}} = 21.7$, $SD = 3.06$), who were asked during classes to fill out a questionnaire about their perception of different professional categories. Based on previous research on social judgment along different content dimensions (Abele, Uchrowski, Suitner, & Wojciszke, 2008), we anticipated medium to large effect sizes. We used G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to compute the required sample size with 80% power and an α level of .05 (two-tailed). The analysis indicated that the number of participants in this study was more than sufficient to find the anticipated effects.

We used a mixed 2 (assessment condition: typicality vs. desirability) \times 2 (professional category: politicians vs. entrepreneurs) \times 2 (evaluated content dimension: leadership vs. morality), with the former factor as a between-participant condition and the latter two factors as repeated measures. Therefore, all participants rated both politicians and entrepreneurs (in random order), but half of them did so in the *typicality assessment* condition and the other half in the *desirability assessment* condition. Participants in the *typicality assessment* condition were asked to think about a typical politician and a typical entrepreneur and to rate to what extent each professional typically possessed 6 personality traits, using a 7-point scale ranging from 1 (= *not at all*) to 7 (= *very much*). The traits were drawn from previous research (Abele et al., 2008, 2016; Catellani & Bertolotti, 2014a, 2014b) and covered two content dimensions: morality (*honest, sincere, loyal*) and leadership (*decided, resolved, tenacious*). Participants in the *desirability assessment* condition were asked to think about an ideal politician and an ideal entrepreneur, and to rate to what extent they possessed the same 6 traits employed in the typicality assessment condition. Mean indexes of politicians' leadership (Cronbach's $\alpha = .900$ across conditions) and morality ($\alpha = .979$), and entrepreneurs' leadership ($\alpha = .800$) and

morality ($\alpha = .880$) were then computed and used in the main analyses.

Results

Using the dimension indexes of each professional as dependent variables, we performed a 2 (assessment: typicality vs. desirability) \times 2 (professional category: politician vs. businessman) \times 2 (content dimension: leadership vs. morality) mixed analysis of variance (ANOVA), with the latter two factors as within-participant factors. A main effect of assessment condition emerged, $F(1, 147) = 262.52$, $p < .001$, $\eta^2 = .64$, with both leadership and morality ratings being on average significantly higher in the desirability assessment condition than in the typicality assessment condition. A main effect of professional category also emerged, $F(1, 147) = 14.97$, $p < .001$, $\eta^2_p = .09$, with entrepreneurs receiving more positive ratings ($M = 5.01$, $SD = 0.85$) than politicians ($M = 4.69$, $SD = 1.47$). The effect of the evaluated content dimension was also significant, $F(1, 147) = 166.73$, $p < .001$, $\eta^2_p = .53$, with the leadership dimension ($M = 5.56$, $SD = 0.94$) rated higher than the morality dimension ($M = 4.68$, $SD = 1.55$). Significant interaction effects between professional category and personality dimension, $F(3, 441) = 61.56$, $p < .001$, $\eta^2_p = .30$, and between professional category and reference norm, $F(1, 147) = 136.19$, $p < .001$, $\eta^2_p = .43$, were also found. These effects were further qualified by a three-way interaction, $F(3, 145) = 27.99$, $p < .001$, $\eta^2_p = .16$.

As shown in Figure 1, follow-up *t*-tests (Bonferroni-corrected for multiple comparisons) in the desirability assessment condition indicated that morality was perceived as more desirable ($M = 6.44$, $SD = 0.79$) than leadership ($M = 5.90$, $SD = 0.78$) for politicians, $t(74) = 4.78$, $p < .0001$, whereas leadership ($M = 6.10$, $SD = 0.76$) was perceived as more desirable than morality ($M = 5.01$, $SD = 1.16$) for entrepreneurs $t(74) = 6.69$, $p < .0001$. In the typicality assessment condition, leadership was perceived as more typical than morality both for politicians ($M = 4.25$, $SD = 1.35$ and $M = 2.60$, $SD = 1.27$, respectively), $t(73) = 11.21$, $p < .0001$, and entrepreneurs ($M = 5.65$, $SD = 0.91$ and $M = 3.68$, $SD = 1.14$, respectively), $t(73) = 13.32$, $p < .001$. Finally, although as stated above desirability was found to be higher than typicality across both target professions and dimensions, such difference was larger (and more significant) in the case of politicians' morality, $t(121.58) = 22.10$, $p < .0001$, than in the case of politicians' leadership, $t(116.49) = 9.16$, $p < .0001$, entrepreneurs' leadership, $t(147) = 3.28$, $p = .0014$, and entrepreneurs' morality, $t(147) = 7.00$, $p < .0001$.

Overall, findings from Study 1 indicated that participants held different norms regarding the typicality and desirability

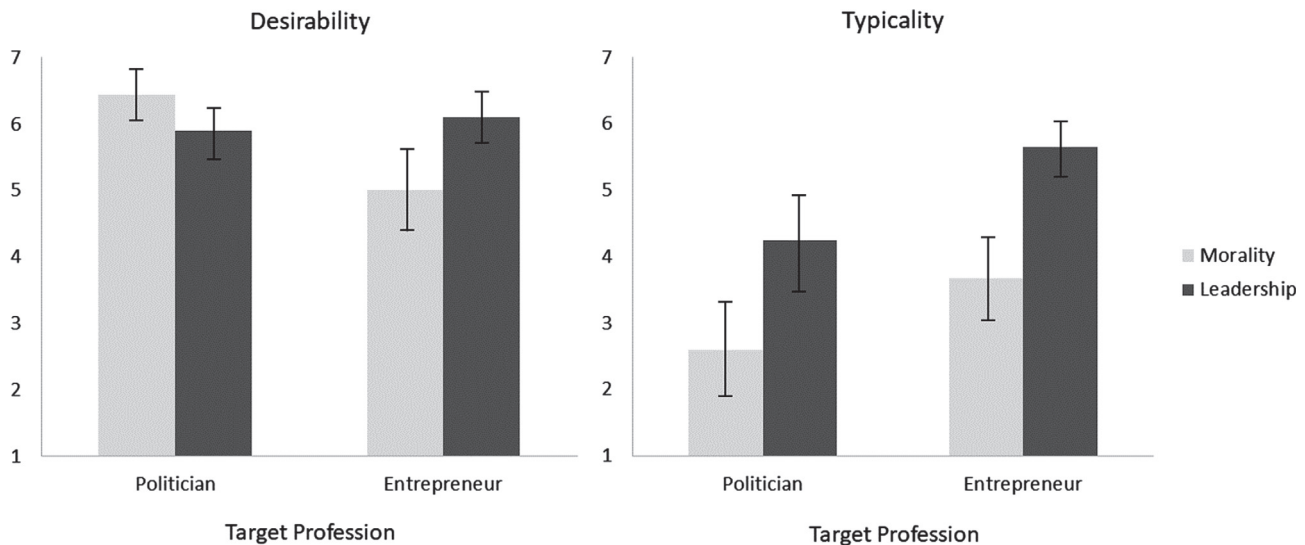


Figure 1. Desirability and typicality of morality and leadership traits as a function of target profession (Study 1).

of the personality dimensions of politicians and entrepreneurs. Leadership was considered the most typical dimension for both professional groups, as opposed to morality, which was considered the least typical dimension. When we considered the desirability of personality dimensions, however, an asymmetry between the two professional groups emerged. While leadership was considered the most desirable dimension for an entrepreneur, morality was considered the most desirable dimension for a politician. These findings therefore confirmed our initial expectations, according to which the two professional categories would differ as to the degree of desirability of morality and leadership attributed to their members.

Study 2

After we assessed the desirability and typicality of morality and leadership for politicians and entrepreneurs in Study 1, in Study 2 we analyzed the effects of attacks focused on each personality dimension and addressing members of either professional category. We designed a set of simulated interviews in which a journalist attacked either an entrepreneur or a politician, criticizing their past behavior as either lacking morality or lacking leadership. The attacks were formulated using either counterfactual statements (i.e., “If you had . . . things would be better now”) or non-counterfactual statements (i.e., “You did . . .”).

Consistent with the results of previous research (Catellani & Bertolotti, 2014b), we expected counterfactual attacks to yield a more positive evaluation of the attack and the attack source, and a more negative evaluation of the

target as compared to non-counterfactual attacks. We also expected that the latter effect would be especially strong when counterfactuals addressed the most desirable dimension for each professional category, namely, morality for politicians and leadership for entrepreneurs.

Method

Participants and Procedure

Participants were 157 undergraduate students (58 males, 92 females, 7 n.r., $M_{age} = 23.0$, $SD = 6.89$) of the Catholic University of Milan, who joined the study on a voluntary basis. Due to the limited availability of volunteers, we had a smaller than recommended sample size to detect the anticipated small- to medium-sized effects with the conventional two-tailed .05 α level and 80% power. Nevertheless, we were able to match the number of total participants of previous studies with similar manipulations and design (see Catellani & Bertolotti, 2014b; Funk, 1996; Kervyn, Dolterer, Mahieu, & Yzerbyt, 2010).

Participants received an email with the link to an online questionnaire in which they were asked to read a fictional interview published in a local newspaper. The content of the fictional interview was based on the one used by Catellani and Bertolotti (2014b) in which, after an initial exchange between a journalist and a politician, the journalist attacked the politician with a series of either counterfactual or non-counterfactual statements. We maintained an unnamed journalist as the ostensibly neutral source, in order to avoid potential confounds deriving from the introduction of a non-neutral source, such as a rival politician or entrepreneur. The sentences were manipulated also

according to the attacked dimension, including either morality-related or leadership-related allegations regarding the management of the local party in the politician condition, or the management of a small business in the entrepreneur condition.

For example, a counterfactual attack statement regarding the politician's [or entrepreneur's] morality was: "If you had told the truth regarding the financial situation of your party [company], things would be better now." The corresponding non-counterfactual version of the statement was: "You hid the truth regarding the financial situation of your party [company]." Similarly, a counterfactual attack regarding the politician's [entrepreneur's] leadership was: "If you had resisted pressures from your allies [shareholders], things would be better now." The corresponding non-counterfactual version of the statement was: "You yielded to pressures from your allies [shareholders]." The full text of the interviews used in the study manipulations is available from the Authors. The overall experimental design was therefore 2 (target profession: politician vs. entrepreneur) \times 2 (attacked dimension: morality vs. leadership) \times 2 (attack style: counterfactual vs. non-counterfactual), all between-participant factors.

Measures

Manipulation Check

Participants were first asked which personality dimension of the target was mainly discussed in the interview, with the target's morality or leadership as the response options.

Evaluation of the Attack Message

Participants were then asked to rate to what extent the journalist's final statement was: *balanced, polite, intelligent, convincing, relevant, clear, complete, and appropriate*, on a 7-point scale ranging from 1 (= *not at all*) to 7 (= *very*) (see Catellani & Bertolotti, 2014b). A message evaluation index was computed from the mean of the ratings, Cronbach's $\alpha = .870$.

Evaluation of the Attacking Source

Participants were then asked to rate to what extent the journalist was hostile to the target and to what extent criticism against the target was based on personal judgments rather than concrete facts, in both cases employing a 7-point scale ranging from 1 (= *not at all*) to 7 (= *very*). A source evaluation index was then computed from the mean of the two items, $r(155) = .451, p < .001$.

Evaluation of the Attacked Target

Participants were finally asked to give a general evaluation of the target on a 7-point scale ranging from 1 (= *very negative*) to 7 (= *very positive*), and to provide more specific ratings on the target's perceived ability to achieve results, keep

promises, and deserve trust, on a 7-point scale ranging from 1 (= *not at all*) to 7 (= *very much*). As the four ratings had moderate to strong correlations, $r(152) = .332$ to $r(152) = .615$, a target evaluation index was computed from the mean of the four items, Cronbach's $\alpha = .800$.

Results

Manipulation Check

Most participants recognized the personality dimension attacked by the journalist at the end of the interview. In particular, most participants in the morality attack condition correctly indicated morality as the attacked dimension of the politician (78.3%) and the entrepreneur (85.3%), and most participants in the leadership attack condition correctly indicated leadership as the attacked dimension of the politician (81.6%) and the entrepreneur (88.4%), $\chi^2(N = 157) = 128.48, p < .001$.

Evaluation of the Attacked Target

A 2 (target profession: politician vs. entrepreneur) \times 2 (attacked dimension: morality vs. leadership) \times 2 (attack style: factual vs. counterfactual) between-participant ANOVA was performed on the evaluation of the attacked target. A main effect of attack style was found, $F(1, 148) = 8.43, p = .004, \eta^2 = .046$, in line with previous findings on the varying effectiveness of counterfactual and non-counterfactual attacks. The attacked target was evaluated more negatively after a counterfactual attack ($M = 3.11, SD = 1.09$) than after a non-counterfactual attack ($M = 3.54, SD = 0.88$). No main effects of target profession or attacked dimension emerged, $F(1, 148) < 1.15, p > .29, \eta^2 < .01$, but the interaction between them was indeed significant, $F(1, 148) = 8.45, p = .004, \eta^2 = .046$, and this effect was verified further by a three-way interaction with attack style, $F(1, 148) = 15.23, p < .001, \eta^2 = .083$ (Figure 2). Consistent with our expectations, Bonferroni post hoc comparisons showed that the evaluation of the politician was more negative after a counterfactual attack on morality ($M = 2.43, SD = 1.10$) than after a counterfactual attack on leadership ($M = 3.68, SD = 0.89$), $p < .001$, while the evaluation of the entrepreneur was more negative after a counterfactual attack on leadership ($M = 2.77, SD = 1.11$) than after a counterfactual attack on morality ($M = 3.55, SD = 0.74$), $p = .006$. No differences of this type were found in the case of non-counterfactual attacks. After non-counterfactual attacks on either leadership or morality the evaluations of the politician did not differ significantly ($M = 3.62, SD = 1.11$ and $M = 3.25, SD = 0.82$, respectively), $p = .247$, and the same held true for the entrepreneur ($M = 3.61, SD = 0.71$ vs. $M = 3.68, SD = 0.85$), $p = .797$. No other significant interaction effects were found, $F(1, 148) < 2.30, p > .13, \eta^2 < .02$.

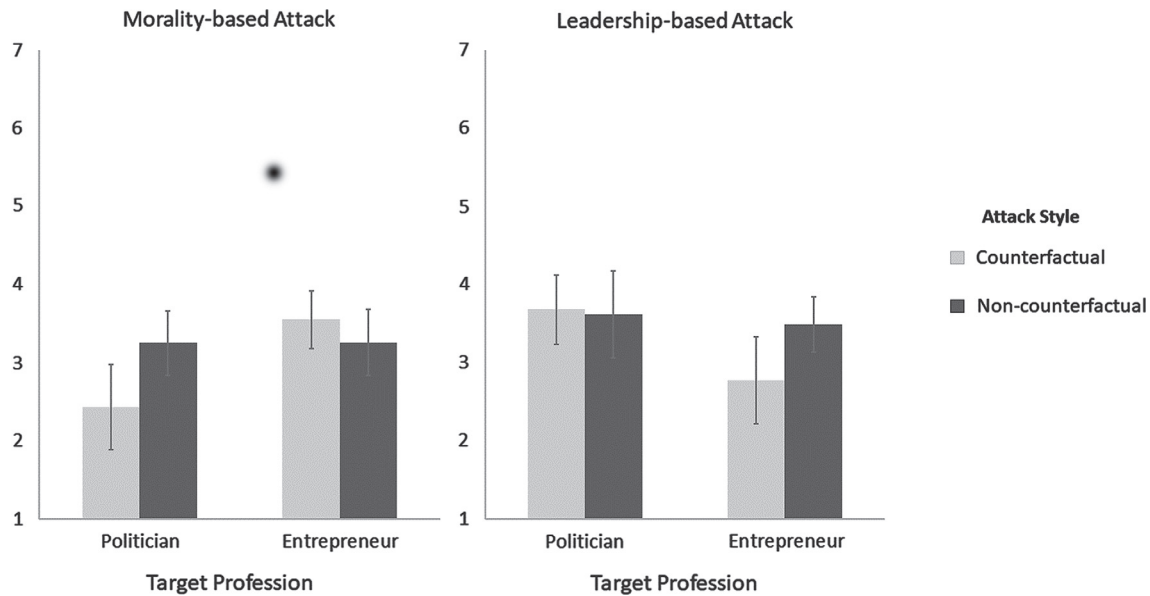


Figure 2. Evaluation of the target as a function of attacked dimension, attacked target, and attack style (Study 2).

Evaluation of the Attack Message and Attacking Source

The same 2 (target profession) \times 2 (attacked dimension) \times 2 (attack style) ANOVA was performed on the evaluation of the attack message. A main effect of attack style was found, $F(1, 148) = 12.42, p = .001, \eta^2 = .075$, with the counterfactual message being evaluated more positively ($M = 4.47, SD = 1.31$) than the non-counterfactual message ($M = 3.86, SD = 1.02$). No other main or interaction effects were found, $F(1, 149) < 2.96, p > .08, \eta^2 < .02$. The same 2 \times 2 \times 2 ANOVA was also performed on the evaluation of the attacking source. A main effect of attack style was again found, $F(1, 149) = 12.14, p = .001, \eta^2 = .072$, with the journalist using a counterfactual attack being generally evaluated as less biased ($M = 4.51, SD = 1.03$) than the journalist using a non-counterfactual attack ($M = 5.08, SD = 1.08$). No other main or interaction effects were found, $F(1, 149) < 2.87, ps > .09, \eta^2s < .02$.

To summarize, the findings of Study 2 supported our hypotheses regarding the differing effectiveness of counterfactual attacks according to attacked dimension and target profession. In particular, we found that counterfactual messages were most effective in attacking a politician's morality on the one hand, and an entrepreneur's leadership on the other. These results suggest that counterfactual attacks are particularly well suited to attack the dimension considered to be most desirable, that is, according to what emerged in Study 1, the morality dimension for politicians and the leadership dimension for entrepreneurs. Furthermore, the analyses of participants' reaction to the attack itself and to the attack source confirmed previous findings (Catellani & Bertolotti, 2014b) regarding the reduced

backlash triggered by counterfactual criticism compared to non-counterfactual criticism.

Study 3

Findings from Study 2 showed that professionals were evaluated more negatively after a counterfactual attack on the content dimension considered to be most desirable for their professional category than after attacks on another dimension. The results of Study 2 therefore suggested that the relevance of an attacked dimension depends more on its desirability than on its typicality, at least when counterfactual attacks are used.

To further test this hypothesis, in Study 3 we manipulated the typicality of the attacked dimension and tested whether this manipulation would influence the effectiveness of the attacks. Before asking participants to read the same stimulus text employed in Study 2, we asked them to read some introductory information regarding the context in which the interview took place, describing the politicians versus entrepreneurs living in the same area of the attacked target as generally high versus low in the dimension which was later attacked in the interview.

We expected the effectiveness of counterfactual attacks to be once again influenced by the desirability of the attacked dimension more than by the (manipulated) typicality of the dimension. Therefore, we expected a counterfactual attack on a politician's morality to yield a more negative evaluation of the politician than a counterfactual attack against the politician's leadership. Conversely,

we expected a counterfactual attack on an entrepreneur's leadership to yield a more negative evaluation of the entrepreneur than a counterfactual attack on the entrepreneur's morality.

Method

Participants and Procedure

Participants were 153 undergraduate students (104 males, 45 females, 4 n.r., $M_{\text{age}} = 23.1$, $SD = 8.60$) of the Catholic University of Milan, who joined the study on a voluntary basis.

The procedure was similar to the one used in Study 2. Depending on the experimental condition, the stimulus consisted in an interview of either a local politician or entrepreneur, and ended with an attack on either the target's leadership or the target's morality. In Study 3, however, attacks were only made in counterfactual terms and, before reading the interview, participants were given some additional information. In the *high typicality conditions*, participants were told that, according to a recent survey, the town where the attacked target lived was at the *top* of the national ranking of politicians' (or entrepreneurs') morality (or leadership). In the *low typicality conditions*, participants were instead told that, according to a recent survey, the town where the attacked target lived was at the *bottom* of the same national ranking. This manipulation aimed to induce the perception that the violation for which the target was attacked referred to a dimension more or less typical of the same professional group. The study therefore had a 2 (target profession: politician vs. entrepreneur) \times 2 (attacked dimension: morality vs. leadership) \times 2 (typicality of the attacked dimension: high vs. low) between-participant design.

Measures

Manipulation Checks

As in Study 2, participants were asked which personality dimension of the target was mainly discussed in the interview, with the target's morality and leadership as response options. To check for the effectiveness of typicality manipulation, in Study 3 participants were also asked to rate the average morality and leadership of local professionals (either politicians or entrepreneurs, depending on the experimental condition) using a 7-point scale ranging from 1 (= *low morality/leadership*) to 7 (= *high morality/leadership*).

Evaluation of the Attacked Target

To measure participants' evaluation of the attacked politician or entrepreneur, we used the same four items used in Study 2. A target evaluation index was again computed from the mean of the four items, Cronbach's $\alpha = .848$.

Evaluation of the Attack Message

Participants were asked to evaluate the final statement of the journalist, with eight items (*balanced, polite, intelligent, convincing, relevant, clear, complete, and appropriate*), on a 7-point scale ranging from 1 (= *not at all*) to 7 (= *very*). A message evaluation index was then computed from the mean of the ratings, Cronbach's $\alpha = .870$. No separate measure of the attack source evaluation was included in this study, as results of Study 2 showed it to be highly correlated with the evaluation of the attack message.

Severity of the Allegation

Two items measured the severity of the allegations made by the journalist. Participants were first asked: "To what extent are the allegations about the behaviour of the politician (or the entrepreneur) severe?", and replied on a 7-point scale ranging from 1 (= *not severe at all*) to 7 (= *very severe*). Then, participants were asked: "To what extent such behaviour is inappropriate for a politician (or entrepreneur)?", and replied on a 7-point scale ranging from 1 (= *not inappropriate at all*) to 7 (= *very inappropriate*). The ratings of the two items, $r(150) = .581$, $p < .001$, were averaged into a single allegation severity index.

Results

Manipulation Checks

As in Study 2, most participants recognized which personality dimension was attacked by the journalist. In the morality attack condition, 80.6% of participants indicated that the politician's morality was attacked and 81.6% indicated that the entrepreneur's morality was attacked. In the leadership attack condition, 92.1% of participants indicated that the politician's leadership was attacked and 74.4% indicated that the entrepreneur's leadership was attacked, $\chi^2(2, N = 151) = 94.65$, $p < .001$.

To check for the typicality manipulation, we performed a 2 (target profession: politician vs. entrepreneur) \times 2 (attacked dimension: morality vs. leadership) \times 2 (dimension typicality: low vs. high) between-participant ANOVA on participants' ratings of how typical the relevant dimension among local professionals was. A main effect of typicality emerged, $F(1, 140) = 28.87$, $p < .001$, $\eta^2 = .17$, indicating that participants in the high typicality condition rated the relevant dimension as more typical of local professionals ($M = 3.81$, $SD = 1.79$) compared to participants in the low typicality condition ($M = 2.42$, $SD = 1.33$). This result confirmed the success of our manipulation. A main effect of attacked dimension was also found, $F(1, 140) = 9.29$, $p = .003$, $\eta^2 = .06$, showing that morality was generally rated as less typical ($M = 2.67$, $SD = 1.51$) than leadership ($M = 3.49$, $SD = 1.80$). This result was consistent with what

already emerged in Study 1 regarding the tendency to rate morality as lower than leadership. No other significant main or interaction effects were found, $F(1, 140) < 1.77, p > .180, \eta^2 < .01$.

Evaluation of the Attacked Target

A 2 (target profession) \times 2 (attacked dimension) \times 2 (dimension typicality) ANOVA was conducted on the evaluation of the attacked target. No main effects were found, $F_s(1, 143) < 2.22, p > .140, \eta^2 < .02$, but an interaction effect between attacked dimension and professional category emerged, $F(1, 143) = 6.83, p = .01, \eta^2 = .043$. As in Study 2, the politician was evaluated more negatively after a morality attack ($M = 2.78, SD = 1.33$) than after a leadership attack ($M = 3.45, SD = 1.11$), while the entrepreneur was evaluated more negatively after a leadership attack ($M = 3.18, SD = 1.02$) than after a morality attack ($M = 3.66, SD = 1.43$). No further two-way or three-way interactions were found, $F_s(1, 143) < 2.95, p > .09, \eta^2 < .02$.

Evaluation of the Attack Message

The same 2 (target profession) \times 2 (attacked dimension) \times 2 (dimension typicality) ANOVA was conducted on participants' evaluation of the attack message. A main effect of dimension typicality was found, $F(1, 144) = 4.63, p = .033, \eta^2 = .033$, as participants rated the message more convincing when the members of the target's professional group were presented as low in the attacked dimension than when they were presented as high in the dimension ($M = 4.91, SD = 1.47$ vs. $M = 4.41, SD = 1.47$). An attacked dimension by target profession interaction was also found, $F(1, 144) = 3.93, p = .049, \eta^2 = .027$. In the case of the politician, the message was evaluated more positively in the morality attack condition ($M = 4.66, SD = 0.89$) than in the leadership attack condition ($M = 4.27, SD = 1.06$), whereas in the case of the entrepreneur, the attack was evaluated more positively in the leadership attack condition ($M = 5.51, SD = 1.17$) than in the morality attack condition ($M = 4.09, SD = 0.93$). No other main or interaction effects were found, $F_s(1, 144) < 2.71, p > .10, \eta^2 < .02$.

Severity of the Allegation

Finally, the same 2 \times 2 \times 2 ANOVA was conducted on the perceived severity of the allegation. Significant effects mirrored those emerged in the analysis of the attack target message. A main effect of dimension typicality was found, $F(1, 144) = 4.00, p = .047, \eta^2 = .023$, as participants rated the allegation as more severe when the members of the target's professional group were presented as low than when they were presented as high in the attacked dimension ($M = 5.39, SD = 1.17$ vs. $M = 4.21, SD = 1.01$). An attacked dimension by target profession interaction was also found, $F(1, 144) = 12.97, p < .001, \eta^2 = .074$. In the case of the

politician, the morality attack was considered to be more severe ($M = 5.57, SD = 1.17$) than the leadership attack ($M = 4.42, SD = 1.37$), whereas in the case of the entrepreneur, the leadership attack was considered to be more severe ($M = 5.51, SD = 1.17$) than the morality attack ($M = 5.21, SD = 1.17$). No other main or interaction effects were found, $F_s(1, 144) < 2.71, p > .10, \eta^2 < .021$.

To summarize, when the target's negative behavior was described as common, the allegation included in the attack was rated as more severe than when the negative behavior was described as uncommon, and the attack was considered more convincing. This suggests that the prevalence of negative behaviors in a certain context may activate harsher evaluative standards, as well as a better evaluation of the source of an attack. However, dimension typicality did not influence the evaluation of the attacked target. Counterfactual attacks on morality and leadership led to a worse evaluation of the attacked target only when either dimension was considered to be highly *desirable* for the professional target, regardless of how *typical* such dimension was of the target's professional group in a given context. Therefore, results from Study 3 further confirmed those of Study 2, showing that counterfactual attacks are particularly effective in leading recipients to compare a target's behavior with a hypothetical ideal behavior the target should have followed.

General Discussion

Our research showed that counterfactual attacks against members of two different professional categories are more effective than corresponding non-counterfactual attacks. Counterfactual attacks induce a more positive evaluation of the attack and of the attack source, as well as a more negative evaluation of the attacked target, as compared to non-counterfactual attacks. In addition, our research showed that counterfactual attacks are especially effective when they focus on a personality dimension that is regarded as highly desirable for the professional category the attacked target belongs to (i.e., morality for politicians and leadership for entrepreneurs).

These findings advance our understanding of the impact of attacks on targets' reputation from different points of view.

First, the findings of the present studies corroborate and extend previous findings regarding the effects of counterfactual criticism on social judgment. In line with past research (Catellani & Bertolotti, 2014b), we found that a counterfactual attack against a politician is more persuasive than a plain factual attack, leading to a poorer judgment of the attacked target and, conversely, to a better judgment of

the attack source. We found evidence that this is the case not only for politicians but, under certain conditions, also for other targets. This effect can be explained by recipients perceiving counterfactual attacks as appropriate statements according to the rules of good communication (Grice, 1975), as compared to non-counterfactual attacks, which are easily recognized as hostile attempts at damaging the target's reputation. Consequently, in the absence of further indications regarding the reliability of the source (see Catellani & Bertolotti, 2014b, Study 2), counterfactual attacks are less likely to be discarded than non-counterfactual attacks. In addition to being perceived as less hostile, in our research counterfactual attacks may have been more persuasive thanks to some of their linguistic and pragmatic features. Due to their additive structure (Roese, 1997), the statements used in our counterfactual attack conditions focused on alleged *inactions* (e.g., failing to disclose information about one's party or company), whereas the statements used in the non-counterfactual attack conditions focused on alleged negative *actions* (e.g., hiding the truth). Past research has found that this distinction can affect how people evaluate a past event, as inaction is often evaluated more negatively than action, and leads to more negative emotional responses (e.g., regret, Zeelenberg, Van den Bos, Van Dijk, & Pieters, 2002). Future research might test the differential perception of past behavior presented in terms of either action or inaction, for instance comparing the effectiveness of additive and subtractive counterfactual attacks. As for the pragmatic features of the attacks, classic research on the rhetoric use of presumption (Walton, 1993) suggests that by presenting the allegation as a mere antecedent to a negative outcome, counterfactual attacks shift recipients' attention from the allegation itself to its consequence. This might have induced our recipients to focus on whether the outcome could or could not have been better, taking the allegation for granted (Cariani & Rips, 2017; Kervyn, Bergsieker, & Fiske, 2012). Further research might investigate this hypothesis, for example by measuring to what extent recipients of counterfactual attacks focus on either part of the sentence, deem it plausible, and generate counterarguments.

Second, our data show that counterfactual attacks are particularly effective when focusing on a personality dimension that is considered desirable for the attacked target's social category. Counterfactual attacks regarding a very desirable dimension induced a more negative evaluation of the attack target than counterfactual attacks regarding a less desirable dimension. Evidently, using a counterfactual attack implicitly pointing at a violation of a shared social norm is an effective way to highlight the contrast between what *happened* and what *should have happened*, and thus influence judgment on who has allegedly violated this norm. Past research

showed that counterfactual thinking is naturally triggered by events and behaviors deviating from both routine (Gavanski & Wells, 1989; Kahneman & Miller, 1986; Kahneman & Tversky, 1982) and social norms (Branscombe et al., 1993; Catellani et al., 2004). Therefore, we might say that counterfactual attacks are processed by recipients as a "natural" form of criticism, because they reproduce in communication the comparison between ideal and actual behavior that individuals spontaneously perform when evaluating a target. This raises the question on whether the inferential process initiated by counterfactual communication is automatic or controlled (Strack & Deutsch, 2004) by the individual exposed to it. Past research on upward counterfactual thinking and its consequences in terms of attitudes, intentions, and behaviors (Epstude & Roese, 2011) suggests that whereas the elicitation of counterfactual thoughts is often automatic, their subsequent processing can be checked, stopped, and reconsidered with conscious effort (Goldinger, Kleider, Azuma, & Beike, 2003). Individuals' ability to control this process is particularly important when we consider the potential consequences of counterfactual communication coming from an external source. Reassuringly, our findings indicate that a counterfactual attack readily evokes a social norm in recipients' mind, but it ultimately affects their judgments only when the norm is deemed relevant by recipients themselves.

Third, our research also contributes to understanding which content dimensions of social judgment are most vulnerable to attacks. In our studies, counterfactual attacks were more effective when they focused on the dimension considered particularly desirable for the social category to which the attacked target belonged. The dimension of morality (sincerity, honesty, reliability) emerged as very desirable for the category of politicians, and consistently an attacked politician was evaluated more negatively when the attack hinted at inadequate behavior related to this dimension. Likewise, the dimension of leadership (determination, decisiveness, tenacity) emerged as very desirable for the category of entrepreneurs, and consistently an entrepreneur was judged more negatively when the attack hinted at inadequate behavior related to this dimension.

According to our results, the *desirability* of the attacked dimension matters more in the evaluation of a target, whereas its *typicality* only affects the perceived severity of the allegation and the overall credibility of the attack. This might reflect recipients' consideration that criticism is in general more plausible when it focuses on more frequent, rather than exceptional, violations. Such consideration, however, does not seem to affect the evaluation of the specific attacked target, which is based on a comparison with the ideal, rather than actual, behavior of the social category the target belongs to.

Future research should further strengthen the study of the effects of counterfactual attacks on different content dimensions and different social and professional groups. It would be also useful to extend the analysis to other social categories where the desirability and typicality of different dimensions may or may not overlap. Several studies have shown that morality has a peculiar relevance in social judgments (Brambilla et al., 2011; Leach et al., 2007). It is possible that expectations regarding morality are relatively more likely to clash with the perception of individuals' and groups' actual behavior. If this is the case, the contrast between desirability and typicality should generally be more evident for morality than for other dimensions.

Understanding which dimensions are perceived as most desirable for different social categories is an essential premise to understanding not only which attacks will most likely be successful, but also how attacked targets may respond to such criticism. Past findings on the effects of defensive communication are conflicting (Kim, Dirks, Cooper, & Ferrin, 2006; Kervyn et al., 2012). We already know that counterfactual defences, like counterfactual attacks, are more successful than non-counterfactual defences because they are indirect (Bertolotti et al., 2013; Catellani & Bertolotti, 2014a). However, it remains to be investigated whether counterfactual defences can further increase their success when they concern personality dimensions that are considered very desirable for the social category to which the defendant belongs.

Future research may also address some limitations of our research. In particular, the limited number of participants in our studies did not allow us to test for potential moderators of the effectiveness of counterfactual attacks, such as recipients' endorsement of the norms evoked by counterfactual attacks (including stereotypes regarding the behavior of different groups, and beliefs regarding its desirability). Future studies may also explore the external validity of our findings, by testing our hypotheses on participants with a different background (e.g., adults and professionals) and nationality. The latter point could be particularly interesting, as it could provide some insight on the extent to which the social norms we measured in Study 1 are influenced by cultural and historical contingencies.

In conclusion, our results contribute to extending our knowledge of the conditions under which counterfactual attacks can be successful, corroborating the idea of a dual pathway through which counterfactual communication affects social judgment. On the one hand, by evoking an appropriate comparison standard, a speaker using a counterfactual attack can prompt the audience to compare what a target person did with what should have been done instead. On the other hand, the same speaker can exert such influence without overtly criticizing the target, but

rather proposing an ostensibly innocuous "what if..." scenario.

References

- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness-communion with warmth and morality. *Frontiers in Psychology, 7*, 1–17. <https://doi.org/10.3389/fpsyg.2016.01810>
- Abele, A. E., Uchroński, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology, 38*, 1202–1217. <https://doi.org/10.1002/ejsp.575>
- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*, 751–763. <https://doi.org/10.1037/0022-3514.93.5.751>
- Alicke, M. D., Buckingham, J., Zell, E., & Davis, T. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin, 34*, 1371–1381. <https://doi.org/10.1177/0146167208321594>
- Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology, 41*, 76–81. <https://doi.org/10.1027/1864-9335/a000011>
- Bavelas, J. B., Black, A., Bryson, L., & Mullett, J. (1988). Political equivocation: A situational explanation. *Journal of Language and Social Psychology, 7*, 137–145. <https://doi.org/10.1177/0261927X8800700204>
- Bertolotti, M., Catellani, P., Douglas, K. M., & Sutton, R. M. (2013). The "Big Two" in political communication: The effects of attacking and defending politicians' leadership and morality in two European countries. *Social Psychology, 44*, 117–128. <https://doi.org/10.1027/1864-9335/a000141>
- Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology, 41*, 135–143. <https://doi.org/10.1002/ejsp.744>
- Branscombe, N. R., Crosby, P., & Weir, J. A. (1993). Social inferences concerning male and female homeowners who use a gun to shoot an intruder. *Aggressive Behavior, 19*, 113–124. [https://doi.org/10.1002/1098-2337\(1993\)19:2<113::AID-AB2480190204>3.0.CO;2-B](https://doi.org/10.1002/1098-2337(1993)19:2<113::AID-AB2480190204>3.0.CO;2-B)
- Branscombe, N. R., & Weir, J. A. (1992). Resistance as stereotype-inconsistency: Consequences for judgments of rape victims. *Journal of Social and Clinical Psychology, 11*, 80–102. <https://doi.org/10.1521/jscp.1992.11.1.80>
- Cariani, F., & Rips, L. J. (2017). Conditionals, context, and the suppression effect. *Cognitive Science, 41*, 540–589. <https://doi.org/10.1111/cogs.12336>
- Carraro, L., Gawronski, B., & Castelli, L. (2010). Losing on all fronts: The effects of negative versus positive person-based campaigns on implicit and explicit evaluations of political candidates. *The British Journal of Social Psychology, 49*, 453–470. <https://doi.org/10.1348/014466609x468042>
- Carrier, A., Louvet, E., Chauvin, B., & Rohmer, O. (2014). The primacy of agency over competence in status perception. *Social Psychology, 45*, 347–356. <https://doi.org/10.1027/1864-9335/a000176>

- Castelli, L., Carraro, L., Ghitti, C., & Pastore, M. (2009). The effects of perceived competence and sociability on electoral outcomes. *Journal of Experimental Social Psychology*, 45, 1152–1155. <https://doi.org/10.1016/j.jesp.2009.06.018>
- Catellani, P., Alberici, A. I., & Milesi, P. (2004). Counterfactual thinking and stereotypes: The nonconformity effect. *European Journal of Social Psychology*, 34, 421–436. <https://doi.org/10.1002/ejsp.206>
- Catellani, P., & Bertolotti, M. (2013). Political communication, social cognitive processes, and voters' judgments. In J. Forgas, O. Vincze, & E. J. Laszlo (Eds.), *Social cognition and communication* (pp. 283–296). New York, NY: Psychology Press.
- Catellani, P., & Bertolotti, M. (2014a). The effects of counterfactual defences on social judgements. *European Journal of Social Psychology*, 44, 82–92. <https://doi.org/10.1002/ejsp.1995>
- Catellani, P., & Bertolotti, M. (2014b). The effects of counterfactual attacks on social judgments. *Social Psychology*, 45, 371–381. <https://doi.org/10.1002/ejsp.1995>
- Catellani, P., & Bertolotti, M. (2015). The perception of politicians' morality: Attacks and defences. In J. P. Forgas, K. Fiedler, & W. Crano (Eds.), *Social psychology and politics* (pp. 112–128). New York, NY: Psychology Press.
- Catellani, P., & Covelli, V. (2013). The strategic use of counterfactual communication in politics. *Journal of Language and Social Psychology*, 32, 480–489. <https://doi.org/10.1177/0261927x13495548>
- Catellani, P., & Quadrio, A. (1991). Ideal and real in the representation of politics. *Revue Internationale de Psychologie Sociale*, 4, 231–255.
- Cislak, A. (2013). Effects of power on social perception. *Social Psychology*, 44, 138–146. <https://doi.org/10.1027/1864-9335/a000139>
- Cislak, A., & Wojciszke, B. (2006). The role of self-interest and competence in attitudes toward politicians. *Polish Psychological Bulletin*, 4, 203–212.
- Cislak, A., & Wojciszke, B. (2008). Agency and communion are inferred from actions serving interests of self or others. *European Journal of Social Psychology*, 38, 1103–1110. <https://doi.org/10.1002/ejsp.554>
- Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, 92, 208–231. <https://doi.org/10.1037/0022-3514.92.2.208>
- Cuddy, A., Fiske, S., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS Map. *Advances in Experimental Social Psychology*, 40, 61–149. [https://doi.org/10.1016/s0065-2601\(07\)00002-0](https://doi.org/10.1016/s0065-2601(07)00002-0)
- Davis, C. G., & Lehman, D. R. (1995). Counterfactual thinking and coping with traumatic life events. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 353–374). Mahwah, NJ: Erlbaum.
- Douglas, K. M., & Sutton, R. M. (2010). By their words ye shall know them: Language abstraction and the likeability of describers. *European Journal of Social Psychology*, 40, 366–374. <https://doi.org/10.1002/ejsp.634>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109, 573. <https://doi.org/10.1037/0033-295x.109.3.573>
- Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, 12, 168–192. <https://doi.org/10.1016/bs.aesp.2017.02.001>
- Epstude, K., & Roese, N. J. (2011). When goal pursuit fails. *Social Psychology*, 42, 19–27. <https://doi.org/10.1027/1864-9335/a000039>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/bf03193146>
- Fiedler, K., & Mata, A. (2013). The art of exerting verbal influence through powerful lexical stimuli. In J. P. Forgas, J. Laszlo, & O. Vincze (Eds.), *Social cognition and communication* (pp. 43–62). New York, NY: Psychology Press.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Funk, C. L. (1996). The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior*, 18, 1–24. <https://doi.org/10.1007/bf01498658>
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314–325. [https://doi.org/10.1016/0022-1031\(89\)90025-5](https://doi.org/10.1016/0022-1031(89)90025-5)
- Giroto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica*, 78, 111–133. [https://doi.org/10.1016/0001-6918\(91\)90007-m](https://doi.org/10.1016/0001-6918(91)90007-m)
- Goldinger, S. D., Kleider, H. M., Azuma, T., & Beike, D. R. (2003). “Blaming the victim” under memory load. *Psychological Science*, 14, 81–85. <https://doi.org/10.1111/1467-9280.01423>
- Grice, H. P. (1975). Logic and conversation. In A. P. Martinich (Ed.), *Philosophy of language* (pp. 165–175). New York, NY: Oxford University Press.
- Haddock, G., & Zanna, M. (1997). Impact of negative advertising on evaluations of political candidates: The 1993 Canadian federal election. *Basic and Applied Social Psychology*, 19, 205–223. <https://doi.org/10.1207/15324839751037066>
- Hegarty, P., & Pratto, F. (2001). The effects of social category norms and stereotypes on explanations for intergroup differences. *Journal of Personality and Social Psychology*, 80, 723. <https://doi.org/10.1037/0022-3514.80.5.723>
- Hornsey, M. J. (2005). Why being right is not enough: Predicting defensiveness in the face of group criticism. *European Review of Social Psychology*, 16, 301–334. <https://doi.org/10.1080/10463280500436040>
- Hornsey, M. J., Robson, E., Smith, J., Esposito, S., & Sutton, R. M. (2008). Sugaring the pill: Assessing rhetorical strategies designed to minimize defensive reactions to group criticism. *Human Communication Research*, 34, 70–98. <https://doi.org/10.1111/j.1468-2958.2007.00314.x>
- Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in Psychology*, 4, 386. <https://doi.org/10.3389/fpsyg.2013.00386>
- Johns, R., & Shepard, M. (2007). Gender, candidate image and electoral preference. *British Journal of Political International Relations*, 9, 434–460. <https://doi.org/10.1057/bp.2008.8>
- Judd, C. M., James-Hawkins, L., Yzerbyt, V. Y., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, 89, 899–913. <https://doi.org/10.1037/0022-3514.89.6.899>
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136–153. <https://doi.org/10.1037/0033-295x.93.2.136>
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, E. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York, NY: Cambridge University Press.
- Kervyn, N., Bergsieker, H. B., & Fiske, S. T. (2012). The innuendo effect: Hearing the positive but inferring the negative. *Journal*

- of *Experimental Social Psychology*, 48, 77–85. <https://doi.org/10.1016/j.jesp.2011.08.001>
- Kervyn, N., Dolderer, M., Mahieu, T., & Yzerbyt, V. Y. (2010). Atypicality and the two fundamental dimensions: Applying the negativity effect on warmth to group perception. *European Journal of Social Psychology*, 40, 484–489. <https://doi.org/10.1002/ejsp.638>
- Kim, P. H., Dirks, K., Cooper, C., & Ferrin, D. (2006). When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. *Organizational Behavior and Human Decision Processes*, 99, 49–65. <https://doi.org/10.1016/j.obhdp.2005.07.002>
- Klauer, K. C., Jacobsen, T., & Migulla, G. (1995). Counterfactual processing: Test of an hierarchical correspondence model. *European Journal of Social Psychology*, 25, 577–595. <https://doi.org/10.1002/ejsp.2420250508>
- Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *Journal of Politics*, 69, 1176–1209. <https://doi.org/10.1111/j.1468-2508.2007.00618.x>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93, 234–249. <https://doi.org/10.1037/0022-3514.93.2.234>
- Leach, C. W., Minescu, A., Poppe, E., & Hagendoorn, L. (2008). Generality and specificity in stereotypes of out-group power and benevolence: Views of Chechens and Jews in the Russian federation. *European Journal of Social Psychology*, 38, 1165–1174. <https://doi.org/10.1002/ejsp.577>
- Maak, T., Pless, N. M., & Voegtlin, C. (2016). Business statesman or shareholder advocate? CEO responsible leadership styles and the micro-foundations of political CSR. *Journal of Management Studies*, 53, 463–493. <https://doi.org/10.1111/joms.12195>
- Mandel, D. R., Hilton, D. J., & Catellani, P. (Eds.). (2005). *The psychology of counterfactual thinking*. New York, NY: Routledge.
- Markman, K. D., & Tetlock, P. E. (2000). “I couldn’t have known”: Accountability, foreseeability and counterfactual denials of responsibility. *The British Journal of Social Psychology*, 39, 313–325. <https://doi.org/10.1348/014466600164499>
- McGraw, K. M. (2003). Political impressions: Formation and management. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 394–432). New York, NY: Oxford University Press.
- McGraw, K. M., Lodge, M., & Jones, J. M. (2002). The pandering politicians of suspicious minds. *Journal of Politics*, 64, 362–383. <https://doi.org/10.1111/1468-2508.00130>
- Meeks, L. (2012). Is she “man enough”? Women candidates, executive political offices, and news coverage. *Journal of Communication*, 62, 175–193. <https://doi.org/10.1111/j.1460-2466.2011.01621.x>
- Mollaret, P., & Miraucourt, D. (2016). Is job performance independent from career success? A conceptual distinction between competence and agency. *Scandinavian Journal of Psychology*, 57, 607–617. <https://doi.org/10.1111/sjop.12329>
- Nario-Redmond, M., & Branscombe, N. (1996). It could have been better or it might have been worse: Implications for blame assignment in rape cases. *Basic and Applied Social Psychology*, 18, 347–366. https://doi.org/10.1207/s15324834basps1803_6
- Nichols, A. L., & Cottrell, C. A. (2014). What do people desire in their leaders? The role of leadership level on trait desirability. *The Leadership Quarterly*, 25, 711–729. <https://doi.org/10.1016/j.leaqua.2014.04.001>
- Pagliaro, S., Brambilla, M., Sacchi, S., D’Angelo, M., & Ellemers, N. (2013). Initial impressions determine behaviours: Morality predicts the willingness to help newcomers. *Journal of Business Ethics*, 117, 37–44. <https://doi.org/10.1007/s10551-012-1508-y>
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121, 133–148. <https://doi.org/10.1037/0033-2909.121.1.133>
- Roese, N. J., & Epstude, K. (2017). The functional theory of counterfactual thinking: New evidence, new challenges, new insights. *Advances in Experimental Social Psychology*, 56, 1–79. <https://doi.org/10.1016/bs.aesp.2017.02.001>
- Roese, N. J., & Sande, G. N. (1993). Backlash effect in attack politics. *Journal of Applied Social Psychology*, 23, 632–653. <https://doi.org/10.1111/j.1559-1816.1993.tb01106.x>
- Schlenker, B. R., & Pontari, B. A. (2000). The strategic control of information: Impression management and self-presentation in daily life. In A. Tesser, R. B. Felson, & J. M. Suls (Eds.), *Psychological perspectives on self and identity* (pp. 199–232). Washington, DC: American Psychological Association.
- Scholer, A. A., & Higgins, E. T. (2008). People as resources: Exploring the functionality of warm and cold. *European Journal of Social Psychology*, 38, 1111–1120. <https://doi.org/10.1002/ejsp.509>
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247. https://doi.org/10.1207/s15327957pspr0803_1
- Tal-Or, N., Boninger, D. S., Poran, A., & Gleicher, F. (2004). Counterfactual thinking as a mechanism in narrative persuasion. *Human Communication Research*, 30, 301–328. <https://doi.org/10.1111/j.1468-2958.2004.tb00734.x>
- Thaler, V. (2012). Mitigation as modification of illocutionary force. *Journal of Pragmatics*, 44, 907–919. <https://doi.org/10.1093/hcr/30.3.301>
- Walton, D. (1993). The speech act of presumption. *Pragmatics & Cognition*, 1, 125–148. <https://doi.org/10.1075/pc.1.1.08wal>
- Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, 56, 161–169. <https://doi.org/10.1037/0022-3514.56.2.161>
- Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology*, 53, 421–430. <https://doi.org/10.1037/0022-3514.53.3.421>
- Wojciszke, B. (2005). Morality and competence in person and self-perception. *European Review of Social Psychology*, 16, 155–188. <https://doi.org/10.1080/10463280500229619>
- Wong, E. M. (2010). It could have been better: The effects of counterfactual communication on impression formation. *European Journal of Social Psychology*, 40, 1251–1260. <https://doi.org/10.1002/ejsp.719>
- Zeelenberg, M., Van den Bos, K., Van Dijk, E., & Pieters, R. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology*, 82, 314–327. <https://doi.org/10.1037/0022-3514.82.3.314>

Received July 5, 2017

Revision received January 16, 2018

Accepted January 16, 2018

Published online June 4, 2018

Mauro Bertolotti

Department of Psychology
Catholic University of Milan
Largo Gemelli, 1
20123 Milan
Italy
mauro.bertolotti@unicatt.it