Changing Impressions:

Moral Character Dominates Impression Updating

Marco Brambilla

University of Milano-Bicocca, Italy

Luciana Carraro, Luigi Castelli

University of Padova, Italy

Simona Sacchi

University of Milano-Bicocca, Italy

Abstract

Research suggests that morality, sociability, and competence exert different effects on impression formation and that morality forms the primary basis for the global evaluation of others. However, prior work has almost exclusively focused on "first" impressions, overlooking that social interactions require flexible updating of initial evaluations. Three experiments tested whether impression updating is influenced by morality, sociability, and competence characteristics to the same extent. Participants were asked to revise their impressions of an individual in light of new and inconsistent information pertaining to his morality, sociability or competence. Results showed that morality was perceived as more informative of interpersonal intentions; therefore a greater impression change occurred when moral information (vs. sociability or competence information) was added to what was previously learned about an individual. Our findings reveal that the key role of morality in social cognition goes beyond the formation of initial evaluations by influencing the updating of such first impressions.

Keywords: Impression Formation, Impression Updating, Morality, Sociability, Competence

Changing Impressions:

Moral Character Dominates Impression Updating

Imagine being at the airport waiting for your flight surrounded by a couple of people that you have never seen before. You look around and start a conversation with a man seated next to you. At the beginning of the conversation you find out that this person is very talkative and friendly. In a similar vein, you notice that he is very clever and smart and that he is pursuing a doctorate in chemistry at a prestigious university. Learning such details about this person would lead you to like him and have a positive impression. In the end, however, it turns out that this man is dishonest as a couple of minutes before boarding he attempted to steal a wallet from the bag of another passenger. This new detail would probably dramatically change your impression about this person and lead you to overtly dislike him. This example suggests that people are a dynamic source of social information who may act inconsistently in social interactions. As a consequence, impression formation is a dynamic process, and our impressions of other people are continually updated in light of new information that might be evaluatively inconsistent with prior information (Cone & Ferguson, 2015; Cone, Mann, & Ferguson, 2017; Mann & Ferguson, 2017; Mann & Ferguson, 2015; Mende-Siedlecki, Baron, & Todorov, 2013; Mende-Siedlecki, Cai, & Todorov, 2013; Reeder & Coovert, 1986; Rydell & McConnell, 2006; Wyer, 2010).

Extensive research has sought to understand our ability to update social impressions in light of behavioral inconsistencies. Most of this work has addressed the ease with which different impressions can be changed as a function of the amount and frequency of counterattitudinal behaviors (for a review, Cone et al., 2017). However, less is known about the specific person characteristics that can promote or disrupt impression change. Complementing and extending prior research evidence, here we argue that impression updating is influenced by the content characteristics that describe our fellow interaction partners and that trait-content information that refers to moral character has a primary role in this sense. Specifically, we tested the hypothesis that

morality should have a leading role over other basic dimensions of human social cognition (i.e., sociability and competence) in the impression-updating process.

Moral Character and Impression Formation

When evaluating other individuals or groups, we are faced with the task of accurately assessing whether someone's intentions are beneficial or harmful, that is, whether they represent an opportunity or a threat. In a similar vein, we need to know others' capabilities, that is, whether they are able to carry out their intentions (Cuddy, Fiske, & Glick, 2008). These two evaluations map onto the dimensions of warmth and competence, respectively (Abele & Wojciszke, 2007; 2014; Cuddy et al., 2008; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005; Wojciszke, 2005). The warmth dimension (also called communion) pertains to benevolence in social relations and involves qualities such as friendliness, honesty, cooperativeness, and trustworthiness. By contrast, the competence dimension (also called agency) refers to qualities that relate to goal-attainment, such as being intelligent or capable (Abele & Wojciszke, 2007; Judd et al., 2005; Wojciszke, 2005).

Two-dimensional models of person and group perception have been extremely influential and have been employed to understand a wide range of social cognitive processes, including impression formation (Abele & Bruckmuller, 2011; Rosenberg, Nelson, & Vivekananthan, 1968; Wojciszke, 1994; 2005), and stereotyping of social groups (Fiske, Cuddy, Glick, & Xu, 2002; Phalet & Poppe, 1997; Poppe, & Linssen, 1999). Specifically, the two-dimensional framework traces its origins to the seminal works on impression formation conducted by Asch (1946) and Rosenberg et al. (1968). Building on their findings, more recent research has shown that warmth and competence account for 82% of the variance in the global impressions of well-known others (Wojciszke, Bazinska, & Jaworski, 1998) and that three-quarters of over 1,000 personally experienced past events are framed in terms of either warmth or competence (Wojciszke, 1994, 2005). At the group level, research on the Stereotype Content Model (Fiske et al., 2002) has revealed that stereotypes differ not only in valence but also in content. Thus, stereotypes are not uniformly positive or negative, but rather can be simultaneously positive on warmth and negative

on competence, or vice versa (Cuddy et al., 2008; Fiske et al., 2002). Taken together, these findings suggest that warmth and competence are basic dimensions that underlie human social cognition and shape interpersonal and group perception (Abele & Wojciszke, 2014).

Extending these two-dimensional models of social cognition, a growing body of research has shown that the warmth dimension encompasses two distinct evaluative components: sociability and morality. Sociability means cooperating and forming connections with others and is exemplified by traits such as friendliness, likeability, and kindness. Morality is linked to the perceived correctness of social targets and is exemplified by traits such as honesty, sincerity, and trustworthiness (for reviews, Brambilla & Leach, 2014; Goodwin, 2015; see also Abele, Hauke, Peters, Louvet, Szymkow, & Duan, 2016).

Building on this distinction, it has been shown that morality forms the primary basis for the global evaluation of others. Thus, research has shown that people perceive facial trustworthiness after as little as a 100-ms exposure to novel faces (Todorov, Said, Engell, & Oosterhof, 2008; Wills & Todorov, 2006) and show a memory advantage for faces that vary in honesty and trustworthiness over faces that vary on non-moral characteristics (Rule, Slepian, & Ambady, 2012). Further, moral information is more decisive than information about sociability or competence in determining the overall impression that people form of other individuals and groups (Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Brambilla, Sacchi, Pagliaro, & Ellemers, 2013; Brambilla, Sacchi, Rusconi, Cherubini, & Yzerbyt, 2012; Cottrell, Neuberg, & Li, 2007; Ellemers & Van den Bos, 2012; Goodwin, Piazza, & Rozin, 2014; Leach, Ellemers, & Barreto, 2007). For instance, individuals rate trustworthiness as the most desirable characteristic for an ideal person to possess (Cottrell et al., 2007) and individual and group self-concept is strongly predicted by self-ascribed morality (Leach et al., 2007; Rodriguez Mosquera, Manstead, & Fischer, 2002). In a similar vein, when individuals are asked to judge either a stranger or a known person, their overall impressions are more strongly predicted by the moral qualities of the target than by sociability or competence characteristics (Brambilla et al., 2011; 2012; Goodwin et al., 2014). Research has demonstrated that morality is

also the primary determinant of the likelihood that people will approach and help others, instead of avoiding them (Brambilla, Sacchi, Menegatti, & Moscatelli, 2016; Brambilla et al., 2013; Iachini, Pagliaro, & Ruggiero, 2015).

According to a functional approach to social perception, "perceiving is for doing" (Fiske, 1992) and its primary purpose is to guide people in establishing others' intentions (Dunning, 2004; Heider, 1958; Zebrowitz & Collins, 1997). As said, in social interactions, people are primarily interested in defining whether someone's intentions are beneficial or harmful to the self and whether it is safe to approach a social target (Cuddy et al., 2008; Ybarra, Chan, & Park, 2001). Morality, which comprises traits such as honesty and trustworthiness, provides important information to infer the intentions of social targets (Brambilla & Leach, 2014). Indeed, morality and the experience of threat are inherently linked (Brambilla, Biella, & Freeman, 2018; Brambilla et al., 2012; 2013). Recent work has shown that the more a social target is perceived as lacking honesty and trustworthiness, the more such a target is believed to pose a threat to the stability and integrity of the entire community. At the group level, ingroup members who lack moral qualities are perceived as threatening to the image of their group (Brambilla et al., 2013; Leach et al., 2007; van der Toorn, Ellemers, & Doosje, 2015), while immoral outgroup members are perceived as posing a real and concrete danger to the ingroup's survival possibilities and represent a threat to the group's safety (Brambilla et al., 2012; 2013; Leidner & Castano, 2012). Consistent with these findings, functional neuroimaging studies have shown that detection of trustworthiness in a face is a spontaneous and automatic process linked to activity in the amygdala (Winston, Strange, O'Doherty, & Dolan, 2002), a subcortical brain structure that is implicated in the detection of potentially dangerous and threatening stimuli (Engell, Haxby, & Todorov, 2007; Todorov, Mende-Siedlecki, & Dotsch, 2013; Todorov, Said, Oosterhof, & Engell, 2011). Taken together, these findings corroborate the claim that a target's morality establishes it as beneficial or harmful to the self. Thus, it makes sense that we are oriented to others' morality and that moral information drives the impression-formation process.

Moral Character and Impression Updating

The primacy of morality in shaping first impressions raises the question of whether morality also drives the updating of such first impressions. Indeed, impression formation is a dynamic and evolving process, as other individuals are an endless source of social information. As a case in point, other individuals may enact inconsistently (Mann & Ferguson, 2015; Mende-Siedlecki, Cai et al., 2013). As a consequence, social interactions require a continuous and flexible updating of our initial impressions. Consider the last time you changed your mind about someone in your life, for instance, a longtime trusted partner who cheated on you or a severe boss who surprised you with an empathetic attitude. In each of these instances, you may have felt that your impression about that person was incorrect and that a different impression was warranted instead.

A growing body of research on impression updating has tapped the processes implied in impression change and their neural bases (Brannon & Gawronski, 2017; Cone & Ferguson, 2015; Mann & Ferguson, 2017; Mann & Ferguson, 2015; Mende-Siedlecki, Baron et al., 2013; Mende-Siedlecki, Cai et al., 2013; Mende-Siedlecki & Todorov, 2016; Reeder & Coovert, 1986; Rydell & McConnell, 2006; Wyer, 2010; see also, Gawronski & Bodenhausen, 2006). This research reveals that both explicit (Rydell & McConnell, 2006; Rydell, McConnell, Strain, Claypool, & Hungenberg, 2007) and implicit (Cone & Ferguson, 2015; Mann & Ferguson, 2017; Mann & Ferguson, 2015; Wyer, 2010) impressions about others can be updated in light of new information. In particular, it has been shown that the revision of first impressions is stronger when the information is subjectively assessed as diagnostic and important (Cone & Ferguson, 2015). For instance, Cone and Ferguson (2015) asked participants to form an impression of an individual person by using large amounts of positive information. Next, participants received either one piece of highly diagnostic negative information or neutral information. Results revealed that participants who received a single piece of diagnostic negative information showed a revision of their initial deliberative and implicit impressions. In addition, such a revision emerged mainly when the target

person was personally responsible for the counterattitudinal behavior rather than merely incidentally associated with a negative act.

Going beyond diagnosticity, research reveals that the revision of first impressions occurs more easily when individuals can elaborate on the earlier information about a social target (Mann & Ferguson, 2015; Wyer, 2010). In these studies, participants formed a negative impression of an individual who enacted various negative acts. After forming that impression, participants received additional positive information about the target person. Results showed that the revision of impressions tends to occur when the additional information that was provided dramatically reversed the meaning of the previous acts performed by the target person and offered a reinterpretation of what was previously learned.

Departing from this body of work, we investigated whether impression updating is influenced by the moral characteristics that are ascribed to an individual target person and whether moral trait-content information has a primary role in this sense over information pertaining to other basic dimensions of social cognition (i.e., sociability and competence). This might help to extend prior findings on the factors that promote impression updating and the work on the role of morality in the impression-formation process. Indeed, prior research evidence on impression updating has shown that impression change is stronger when new information is subjectively assessed as diagnostic and important (Cone & Ferguson, 2015), but has not defined the specific person characteristics that may enhance or diminish impression change. In a similar vein, the studies that considered moral information in impression updating (Mende-Siedlecki & Todorov, 2016; Reeder & Coovert, 1986) did not test whether moral information is more relevant than information that pertains to other basic dimensions of social cognition in promoting the revision of first impressions. In addition, most studies on the key role of morality in the impression-formation process have almost exclusively focused on "first" impressions, overlooking impression updating (Brambilla & Leach, 2014; Goodwin, 2015).

Adopting an additive perspective (Anderson, 1962), one might expect that the key role of morality in predicting initial evaluations is merely confirmed in the subsequent stages of the impression-formation process. In this case, first impressions would be adjusted incrementally by additional pieces of information (Hogarth & Einhorn, 1992). However, research has shown that first impressions and belief updating are two distinct phases of the impression-formation process that involve different brain areas (Mende-Siedlecki & Todorov, 2016). In a similar vein, research has revealed that the same information might be treated differently at each stage of the impression-formation process. As such, first information is key to form an initial evaluation; at the impression updating stage, all pieces of evidence are reinterpreted considering the social perceiver's expectations (Trope & Liberman, 1996). Accordingly, Asch (1946) argued that the first information might produce a "context effect" that alters the meaning of the following information.

Consistent with this reasoning, a large body of work on confirmation bias has shown that first impressions operate as filters through which we interpret new information (Klayman, 1995; Nickerson, 1998). In other words, first impressions can create expectations that influence our subsequent information processing. By learning that someone is an introvert, for instance, we might establish an expectation that can lead us to be more sensitive to additional information that confirms such a trait and pay less attention to other information. As a case in point, we do not know whether morality more strongly changes initial judgments based on other meaningful dimensions (such as sociability or competence).

Thus, here, we investigated whether impression updating is influenced to the same extent by morality, sociability and competence-trait information. Considering that morality has a leading role in establishing the intentions of other individuals (Brambilla et al., 2013; Willis & Todorov, 2006; for a review, Brambilla & Leach, 2014), there is good reason to expect that morality could more strongly influence the impression-updating process than sociability and competence. Given that the identification of the intentions of our fellow interaction partners is the main driver of the information-formation process (Wojciszke, 2005; Zebrowitz & Collins, 1997), we expected that

moral behaviors would be interpreted as highly informative in this sense and would therefore be key in promoting impression revisions. In other words, we expected a greater impression change when moral information (rather than nonmoral information) is added to what was previously learned about a target person. We tested these predictions in three studies where participants were asked to form an initial impression about a target person and subsequently revise their first impression in light of new information about that person. In Experiment 1 we manipulated morality and sociability information, while in Experiment 2 and Experiment 3 morality was crossed with competence information. In Experiment 3 we further tested the mediating mechanism that may drive the hypothesized effect. The studies that are reported in this paper were approved by the local ethics committees and were conducted according to the guidelines that were established in the Declaration of Helsinki. In the three experiments we report all measures, manipulations, and exclusions. Moreover, in the three experiments sample sizes were determined before any data analysis.

Experiment 1

Experiment 1 was designed as a first test of our hypothesis that morality drives the impression-updating process. To do so, we asked participants to form a first impression of a target person based on information pertaining to either the moral character or sociability of that individual. Subsequently, we asked participants to revise their impressions in light of new and inconsistent information. The study employed a within-participants design.

Participants

Forty Italian students (25 females, 15 males; $M_{age} = 23.67$, $SD_{age} = 2.68$) volunteered to participate in the study. We advertised the study on campus and all the students who responded within 4 weeks were involved in the study. A sensitivity analysis conducted with G*Power (Faul, Erdfelder, Buchner, & Lang, 2007) showed that our sample was sufficient to detect small-to-medium effects of $f = 0.20$, assuming an α of 0.05, and power of 0.80 for a within-participants ANOVA (observed correlation among repeated measures, $r = .61$).

Material and Procedure

Participants were asked to participate in an experiment on impression formation. When participants arrived at the laboratory, they were comfortably seated in a chair that was positioned approximately 60 cm away from a 22-in LCD computer monitor (Asus® VW226; Resolution: 1680 pixels × 1050 pixels; Refresh rate: 59 Hz). Stimuli presentation and response registration were controlled by the E-Prime 2.1 software.

After receiving instructions, participants were presented with the picture of a male target accompanied by a short sentence describing his behavior (e.g. "*He has lied to his parents*"). In the following screen, participants were asked to report their initial impression of the target (i.e., '*What is your global impression of this individual*?' – T1) by using a seven-point scale that ranged from 1 (*extremely negative*) to 7 (*extremely positive*) (see De Bruin & Van Lange, 1999; Wojciszke et al., 1998). This first behavior varied for dimension (morality vs. sociability) and valence (positive vs. negative). Then, participants were directed to the next screen where they were presented with additional information on the impression target's behavior (e.g. "*He has been friendly with a colleague*"). This second behavior varied for dimension (morality vs. sociability), whereas its valence was always inconsistent with the valence of the first behavior. Thus, for instance, if the first behavior was positive and morality-related, then the second behavior was either morality- or sociability-related but negative. Next, in light of the new behavior, participants were asked to report their impression of the target (T2) along the same seven-point scale that ranged from 1 (*extremely negative*) to 7 (*extremely positive*). After this second answer, participants were presented with the next target (for the experiment flow, see the supplementary materials). The exposure time of the information was regulated by participants who were asked to tap on the spacebar of the computer keyboard to continue.

In sum, the experiment employed a 2 (first behavior dimension: morality vs. sociability) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. sociability) design with all the factors varying within participants. We selected 3 positive morality-related behaviors, 3

negative morality-related behaviors, 3 positive sociability-related behaviors and 3 negative sociability-related behaviors that were carefully balanced for their content relatedness and favorability[1] (to rule out that our findings might be due to a general effect of valence). The combination of the set of behaviors according to our experimental design (see the supplementary materials for the list of behaviors and their combination) resulted in a total of 72 trials. The target's image and the pair of statements on his behavior were randomly combined. The target images were balanced for image quality and expression neutrality and were drawn from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015). At the end of the experiment, participants were asked personal information (age, gender, and nationality), thanked and fully debriefed.

<div align="center">

Results

</div>

First Impressions

As a first step, the impression of the person after the exposure to the first behavior (T1) was submitted to a 2 (first behavior dimension: morality vs. sociability) $\times$ 2 (valence: positive vs. negative) within-participants ANOVA.

The analysis showed the expected main effect of valence, $F(1, 39) = 218.67$, $p < .001$, $\eta_p^2 = .85$. Participants judged the target more negatively in the negative condition ($M = 3.11$, $SD = .56$) than in the positive condition ($M = 5.26$, $SD = .81$). The main effect of the first behavior dimension was not significant, $F(1, 39) = .11$, $p = .74$, $\eta_p^2 = .003$. However, the ANOVA yielded an interaction effect between the first behavior dimension and valence, $F(1, 39) = 70.17$, $p < .001$, $\eta_p^2 = .64$. Positive moral behaviors led to more positive impressions ($M = 5.53$, $SD = .82$) than positive sociability behaviors ($M = 4.98$, $SD = .81$), $p < .001$, $t(39) = 5.57$, $p < .001$, $d = .88$, 95% CI [.51, 1.24]. By contrast, negative moral behaviors ($M = 2.82$, $SD = .62$) led to more negative impressions than negative sociability behaviors ($M = 3.41$, $SD = .49$), $p < .001$, $t(39) = 6.35$, $p < .001$, $d = 1$, 95% CI [.61, 1.38]. This result is consistent with prior research that shows the primacy of morality in predicting first impressions (see Brambilla & Leach, 2014).

Impression Updating

As a next step, we computed an index of impression updating by subtracting the impression score that was reported after the exposure to the first behavior (T1) from the impression score that was reported after the second behavior (T2). Thus, the greater the index – either in the positive or the negative direction – the greater was the impression change after being exposed to the new piece of information.

Then a 2 (first behavior dimension: morality vs. sociability) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. sociability) within-participants ANOVA was computed on the index of impression updating (see Figure 1).

The analysis yielded the expected main effect of valence, $F(1, 39) = 201.81$, $p < .001$, $\eta_p^2 = .84$. Specifically, the impression improved when the first behavior was negative and the second behavior was positive ($M = 1.17$, $SD = .69$), whereas it worsened when the first behavior was positive and the second behavior was negative ($M=-1.62$, $SD = .83$).

Importantly for the aim of our research, the ANOVA revealed an interaction effect between the second behavior dimension and valence, $F(1, 39) = 95.69$, $p < .001$, $\eta_p^2 = .71$. The impression updating was greater when the second positive and inconsistent piece of information referred to morality ($M = 1.46$, $SD = .67$) rather than to sociability ($M = .88$, $SD = .63$), $t(39) = 7.60$, $p <.001$, $d =1.20$, 95% CI [.79, 1.60]. The impression worsened more strongly when the second negative and inconsistent piece of information referred to morality ($M = -1.95$, $SD = .83$) rather than to sociability ($M = -1.28$, $SD = .69$), $t(39) = 7.63$, $p <.001$, $d =1.20$, 95% CI [.79, 1.61]. The analysis yielded no main effect of the second behavior dimension, $F(1, 39) = .64$, $p = .43$, $\eta_p^2 = .02$, and did not yield a three-way interaction effect, $F(1, 39) = .61$, $p = .44$, $\eta_p^2 = .01^2$.

We conducted additional analyses by computing a 2 (first behavior dimension: morality vs. sociability) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. sociability) × 2 (time: T1 vs. T2) within-participants ANOVA. The analysis yielded the predicted three-way interaction among the second behavior dimension, valence, and time, $F(1, 39) = 95.69$, $p < .001$, $\eta_p^2 = .71$. Results further confirmed that our effects were not due to unexpected differences

in the first impressions (T1) between two analogous conditions that referred to the same dimension (all *ps* were not significant; see the supplementary materials for the full set of analyses). Taken together, these findings confirm our predictions and revealed that morality has a leading role over sociability in the impression-updating process.

Experiment 2

Experiment 2 was designed to replicate and extend the findings of Experiment 1 by considering morality and competence information. Thus, participants formed a first impression of a target person based on information that pertained to either morality or competence. Subsequently, we asked participants to revise their impressions after considering new and inconsistent information about that target person. The study employed a within-participants design.

Participants

We aimed at collecting the same number of participants employed in Experiment 1. Therefore, we recruited forty Italian students that were not involved in Experiment 1 ($M_{age} = 23.82$ $SD_{age} = 6.04$; 14 males). A sensitivity analysis conducted with G*Power showed that our sample was sufficient to detect small-to-medium effects of $f = 0.21$, assuming an α of 0.05, and power of 0.80 for a within-participants ANOVA (observed correlation among repeated measures, $r = .58$).

Material and Procedure

The experimental design and the procedure mirrored the design and procedure that were employed in Experiment 1. The only change concerned the dimension that was compared with morality: in Experiment 1, this dimension was sociability, whereas in Experiment 2 it was competence. Thus, the experiment employed a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) design with all the factors varying within participants. We selected 3 positive competence-related behaviors (e.g., *"He put a lot of effort to achieve a challenging goal"*) and 3 negative competence-related behaviors (e.g., *"He did not get good marks at the university"*), and we employed the same morality-related behaviors used in Experiment 1. All the behaviors were

carefully balanced for their content relatedness and favorability (See footnote 1). The combination of the different behaviors according to our experimental design (see the supplementary materials) resulted in a total of 72 trials.

Results

First Impressions

As a first step, the impression on the social target after the exposure to the first behavior (T1) was submitted to a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) within-participants ANOVA.

The analysis showed the expected main effect of valence, $F(1, 39) = 60.71$, $p < .001$, $\eta_p^2 =$ .61. Participants judged the target more negatively in the negative condition ($M = 3.90$, $SD = .59$) than in the positive condition ($M = 4.56$, $SD = .56$). We also found the main effect of the first behavior dimension, $F(1, 39) = 18.30$, $p < .001$, $\eta_p^2 = .32$. Thus, participants judged the targets more positively after receiving clues (both positive and negative) on their competence ($M = 4.33$, $SD =$ .56) rather than on their morality ($M = 4.12$, $SD = .60$). Most importantly, the ANOVA yielded an interaction effect between the first behavior dimension and valence, $F(1, 39) = 20.45$, $p < .001$, η_p^2 = .34. Indeed, positive moral behaviors ($M = 4.58$, $SD = .58$) and positive competence behaviors ($M = 4.54$, $SD = .54$) equally affected the participants' impressions, $t(39) = .70$, $p = .48$, $d = .11$, 95% CI [-.20, .142]; however, immoral behaviors ($M = 3.67$, $SD = .62$) led to more negative impressions than incompetent behaviors ($M = 4.13$, $SD = .57$), $t(39) = 5.22$, $p < .001$, $d = .82$, 95% CI [.46, 1.18]. Thus, consistent with Experiment 1, we found that morality – and especially immorality - has a leading role in driving first impressions.

Impression Updating

As a next step, we computed an index of impression updating by subtracting the impression score that was reported after the exposure to the first behavior (T1) from the impression score that was reported after the second behavior (T2).

Then a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) within-participants ANOVA was computed on the index of impression updating (see Figure 2).

As in Experiment 1, the analysis yielded a main effect of valence, $F(1, 39) = 216.93$, $p < .001$, $\eta_p^2 = .85$. Specifically, the impression improved when the first piece of information was negative and the second piece of information was positive ($M = 1.38$, $SD = .84$), whereas it worsened when the first piece of information was positive and the second piece of information was negative ($M = -1.78$, $SD = .87$).

Importantly, consistent with the first study, the ANOVA revealed an interaction effect between the second behavior dimension and valence, $F(1, 39) = 19.57$, $p < .001$, $\eta_p^2 = .33$. The impression improved to a greater extent when the second positive and inconsistent piece of information referred to morality ($M = 1.48$, $SD = .76$) rather than to competence ($M = 1.28$, $SD = .81$), $t(39) = 2.40$, $p = .02$, $d = .37$, 95% CI [.05, .69]. By contrast, when the second and inconsistent piece of information was negative, the impression worsened more in the morality condition ($M = -2.01$, $SD = .82$) than in the competence condition ($M = -1.55$, $SD = .84$), $t(39) = 3.91$, $p < .001$, $d = .61$, 95% CI [.27, .95]. The analysis did not yield any other significant effects.

As in Experiment 1, we conducted additional analyses by computing a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) × 2 (time: T1 vs. T2) within-participants ANOVA. The ANOVA yielded a three-way interaction among the second behavior dimension, valence, and time, $F(1, 39) = 19.57$, $p < .001$, $\eta_p^2 = .33$. Results further confirmed that our effects were not due to unexpected differences in the first impressions (T1) between two analogous conditions that referred to the same dimension (all ps were not significant; see the supplementary materials for the full set of analyses). Taken together, these findings confirm that morality has a leading role over competence in the impression-updating process as the impressions were more polarized when the second behavior referred to morality rather than to competence.

Experiment 3

Experiment 3 aimed at replicating and extending the findings of the previous two experiments by uncovering one likely mediating mechanism that may drive the greater power of moral behaviors in modifying first impressions. A good deal of work has indicated that the primary purpose of impression formation is to establish other people's intentions (Cuddy et al., 2008; Dunning, 2004; Heider, 1958; Ybarra, Chan, & Park, 2001; Zebrowitz & Collins, 1997). In the last decade, a growing body of work has revealed that moral-trait content has greater informational power than sociability- and competence-trait content in inferring the intentions of social targets (for reviews, Brambilla & Leach, 2014; Goodwin, 2015). As such, immoral individuals are perceived as harmful while moral individuals are seen as beneficial (Brambilla & Leach, 2014). That is, moral and immoral individuals are seen as fundamentally good and bad, respectively. Given that the identification of others' intentions is the main driver of the information-formation process (Wojciszke, 2005; Zebrowitz & Collins, 1997) and considering that moral information is more relevant than other information in establishing whether someone is fundamentally good or bad (Brambilla & Leach, 2014; Goodwin, 2015), we expected that morality would promote a greater impression change because such information would be interpreted as more diagnostic of a person's intentions. To do so, we asked participants to revise their first impressions about a target person after considering new and inconsistent information (which referred either to morality or competence) about such an individual. We further measured the extent to which participants viewed the actions as informative of the intentions of the individual.

To increase the validity and the robustness of our findings, Experiment 3 employed a slightly different methodology. First, we used a larger set of morality-related and competence-related behaviors. Second, we adopted a between-subjects design that was similar to most of the studies on impression updating (see Cone et al., 2017).

Participants

Two hundred and sixty-seven young adults ($M_{age} = 26.61$, $SD_{age} = 8.17$; 48 males) volunteered to complete an online experiment. A sensitivity analysis conducted with G*Power showed that our sample was sufficient to detect small-to-medium effects of $f= 0.17$, assuming an α of 0.05, and power of 0.80 for a between-participants ANOVA with eight groups.

Material and Procedure

Participants were asked to participate in an online experiment on impression formation. Participants were presented with the picture of a male target (named Fabio) accompanied by a short sentence that described his behavior (e.g. "*Fabio did not give back the change in excess he got at the supermarket*"). This first behavior varied for dimension (morality vs. competence) and valence (positive vs. negative). In our first two studies, we measured our criterion variable through a single item. To increase the validity of our findings, Experiment 3 relied on a measure of the first impression of the target that involved three evaluative items. Thus, participants were asked to report their behavioral dispositions (T1) toward Fabio (i.e., I would like to: *meet Fabio*, *interact with Fabio, talk to Fabio* – α = .85).

Then, participants were directed to the next screen where they were presented with additional information on the impression target's behavior (e.g., "*Fabio has made a patent*"). This second behavior varied for dimension (morality vs. competence), whereas its valence was always inconsistent with the valence of the first behavior. Thus, for instance, if the first behavior was negative and morality-related, then the second behavior was either morality- or competence-related but positive.

Participants were then asked to assess how much such an additional behavior was informative of the target's intentions by means of two items (i.e., *"How much is this behaviour useful to determine Fabio's intentions?"; How much is this behaviour useful to determine Fabio's purposes?"*- α = .66; $r=.49$, $p<.001$).

For exploratory purposes, we further included the assessment of the perceived frequency of the behaviours as an additional potential mediator (Mende-Siedlecki et al., 2013; Rothbart & Park,

1986). Indeed, it might be possible that morality drives its effects on impression updating because moral behaviors are perceived as less frequent than other behaviors. Thus, participants were asked to indicate the extent to which the additional behavior is frequent (i.e., *How much do you think the behavior described is widespread among the general population?, How much do you think the described behavior is rare?, How much do you think the behavior described is frequent?, How likely are you to witness similar behavior in everyday life?"-* α = .93).

Next, in light of the new piece of information, participants were further asked to report their behavioral dispositions toward the target (T2) along the same items that were used in T1 (α = .89).[3] Participants answered all the questions by using 7-point scales that ranged from 1 (*not at all/strongly disagree*) to 7 (*very much/strongly agree*). In sum, the experiment employed a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) design with all the factors varying between participants. Participants were randomly assigned to one of the eight experimental conditions. To manipulate the morality and competence, we used 24 behaviors: we employed the same 12 behaviors that were used in Experiment 2, plus 12 additional behaviors (3 positive morality-related behaviors, 3 negative morality-related behaviors, 3 positive competence-related behaviors and 3 negative competence-related behaviors). The entire new set of 24 behaviors was carefully balanced for content relatedness and favorability[4]. Participants were exposed to one behavior in T1 and one behavior in T2 that were randomly selected from the above list of 24 behaviors (see the supplementary materials for the list of behaviors). At the end of the experiment, they were asked personal information (age, gender, and nationality), thanked and fully debriefed.

Results

First Impressions

As a first step, the behavioral dispositions after the exposure to the first behavior (T1) were submitted to a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) between-participants ANOVA. The analysis yielded the main effect of valence, $F(1, 263)$

= 11.03, $p = .001$, $\eta_p^2 = .04$. Participants showed more negative dispositions in the negative

condition ($M = 2.85$, $SD = 1.23$) than in the positive condition ($M=3.30$, $SD = 1.15$). We also found

the main effect of the first behavior dimension, $F(1, 263) = 9.21$, $p = .003$, $\eta_p^2 = .03$. Thus,

participants judged the social target more positively after receiving clues (both positive and

negative) on his competence ($M = 3.28$, $SD = 1.16$) rather than on his morality ($M=2.85$, $SD =$

1.23). More importantly, the ANOVA yielded an interaction effect between the first behavior

dimension and valence, $F(1, 263) = 6.02$, $p = .015$, $\eta_p^2 = .02$. Thus, positive moral behaviors ($M =$

3.26, $SD = 1.19$) and positive competence behaviors ($M = 3.35$, $SD = 1.11$) equally affected the

participants' dispositions, $t(128) = .41$, $p = .68$, $d=.07$, 95% CI [-.27, .42]; however, immoral

behaviors ($M = 2.44$, $SD = 1.14$) led to more negative dispositions than incompetent behaviors ($M =$

3.22, $SD = 1.20$), $t(135) = 3.90$, $p < .001$, $d=.66$, 95% CI [.32, 1.01]. These results replicated the

results of Experiment 2 by revealing that morality – and especially immorality - has a leading role

in driving first impressions.

Impression Updating

As a next step, we computed an index of impression updating by subtracting the behavioral

dispositions that were reported after the exposure to the first behavior (T1) from the behavioral

dispositions that were reported after the second behavior (T2). Then a 2 (first behavior dimension:

morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension:

morality vs. competence) between-participants ANOVA was computed on the index of impression

updating (Figure 3).

The analysis revealed a main effect of valence, $F(1, 259) = 49.72$, $p < .001$, $\eta_p^2 = .16$. The

behavioral dispositions improved when the first piece of information was negative and the second

piece of information was positive ($M = .27$, $SD = .90$), whereas they worsened when the first piece

of information was positive and the second piece of information was negative ($M=-.51$, $SD = .95$).

Importantly, the ANOVA revealed an interaction effect between the second behavior dimension and

valence, $F(1, 259) = 21.43$, $p < .001$, $\eta_p^2 = .08$. The dispositions improved to a greater extent when

the second positive and inconsistent piece of information referred to morality ($M = .51$, $SD = 1.01$) rather than to competence ($M = .04$, $SD = .71$), $t(135) = 3.13$, $p = .002$, $d = .53$, 95% CI [.19, .87]. By contrast, when the second and inconsistent piece of information was negative, the dispositions worsened more in the morality condition ($M = -.77$, $SD = .98$) than in the competence condition ($M = -.22$, $SD = .83$), $t(128) = 3.44$, $p = .001$, $d = .60$, 95% CI [.25, .96]. The ANOVA did not yield any other significant effects, $Fs(1, 259) < 1.14$, $ps > .29$.

We conducted additional analyses by computing a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) × 2 (time: T1 vs. T2) ANOVA with the first three factors between-participants and the last factor within-participants. Consistent with prior findings, the analysis yielded the three-way interaction among the second behavior dimension, valence, and time, $F(1, 259) = 21.43$, $p < .001$, $\eta_p^2 = .08$. Results confirmed that our effects were not due to unexpected differences in the first impressions (T1) between two analogous conditions that referred to the same dimension (all ps were not significant; see the supplementary materials for the full set of analyses). Taken together, these findings confirmed our predictions and the findings reported in Experiment 1 and Experiment 2 by revealing that morality has a leading role in the impression-updating process.

Mediational Analysis

To test whether morality promoted a greater impression change because such information is interpreted as more informative of a person's intentions, the index of person's intentions was submitted to a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) between-participants ANOVA. The analysis yielded the main effect of the second behavior dimension, $F(1, 259) = 47.14$, $p < .001$, $\eta_p^2 = .15$: in line with the hypotheses, the second behavior was judged to be more informative of the target's intentions when it was related to morality ($M = 3.94$, $SD = 1.41$) than to competence ($M = 2.85$, $SD = 1.17$).

Next, we submitted the index of perceived frequency of the behaviors to a 2 (first behavior dimension: morality vs. competence) × 2 (valence: positive vs. negative) × 2 (second behavior dimension: morality vs. competence) between-participants ANOVA. The analysis yielded a main effect of valence, $F(1, 259) = 142.81$, $p < .001$, $\eta_p^2 = .36$. Indeed, positive behaviors were perceived to be more frequent ($M = 5.21$, $SD = 1.29$) than negative behaviors ($M = 3.50$, $SD = 1.24$). Moreover, we found a main effect of the second behavior dimension, $F(1, 259) = 15.20$, $p < .001$, $\eta_p^2 = .06$. Moral behaviors were perceived to be more frequent ($M = 6.64$, $SD = 1.37$) than competence behaviors ($M = 4.02$, $SD = 1.50$).

We further computed a mediation model (Model 4, 5000 bootstrap resampling, PROCESS; Hayes, 2013) by using the second behavior dimension as independent variable (competence = 0; morality = 1), person's intentions and perceived frequency as the mediators, and the index of impression updating in absolute value as the dependent variable. The mediation model was significant when person's intentions were considered, B = .18, SE= .05, CI [.09, .31]. By contrast, perceived frequency of the behaviors did not mediate our effects, B = .01, SE= .02, CI [-.05, .03]. Taken together, these findings suggest that morality has a leading role in promoting the impression change because moral qualities are perceived to be more informative of the intentions of social targets than competence.

General Discussion

Three studies provided consistent support for our hypothesis that morality has a leading role over the other basic dimensions of human social cognition (i.e., sociability and competence) in driving the impression-updating process. Experiment 1 confirmed prior insights by showing that moral information is more decisive than sociability information in determining the initial impression about an individual person (Brambilla & Leach, 2014; Goodwin, 2015). Thus, moral individuals elicit more positive impressions than competent social targets. By contrast, immoral individuals are perceived more negatively than unsociable targets. Going beyond first impressions, Experiment 1 revealed that morality plays a key role when people are asked to revise their initial

evaluations in light of new and inconsistent information about an unknown other person. Indeed, we found a greater impression change when moral information (vs. sociability information) was added to what was previously learned about a target person. Impressions more strongly improved when positive moral (vs. sociability) qualities were added to previous negative qualities that described an unknown other person. By contrast, impressions more strongly worsened when negative moral (vs. sociability) qualities were added to previous positive qualities that described a target person.

Experiment 2 corroborated these findings in a design that enabled us to disentangle the effects of perceived morality from competence as another important evaluative dimension. Thus, morality (especially negative information) predicted more strongly than competence the initial evaluation of an unknown other. Moreover, consistent with Experiment 1, participants displayed a greater impression change when moral information (vs. competence information) was added to what was previously learned about a target person.

Finally, Experiment 3 replicated prior findings by further uncovering the mediating mechanism that drives the greater power of moral behaviors in modifying first impressions. Thus, we showed that morality promoted a greater impression change because such information was interpreted to be more informative of person's intentions. Given that establishing people's intentions is a key driver of impression formation (Dunning, 2004; Heider, 1958; Zebrowitz & Collins, 1997) it makes sense that morality has a greater power over other basic dimensions when we are called to revise our initial evaluations.

Together, these findings make a novel contribution to the literature on impression change. Indeed, prior work in this area has investigated the neural basis (Mende-Siedlecki, Cai et al., 2013; Mende-Siedlecki & Todorov, 2016) and the role of diagnosticity (Cone & Ferguson, 2015) and elaboration (Mann & Ferguson, 2015, 2017) in promoting impression change. These studies reveal that people update their impressions when new information about a target person is subjectively assessed to be diagnostic and important. In a similar vein, impression change tends to occur when the additional information offers a reinterpretation of what was previously learned about our

interaction partners. However, prior research did not define the specific person characteristics that are more likely to promote impression change. Our findings show that person characteristics actually influence the revision of first impressions. We further showed that person characteristics are not all alike and that morality has an exclusive and distinctive role in this sense. As a case in point, we showed that information about the moral character of an individual promoted a stronger impression change beyond the information that referred to the other basic dimensions of social cognition (i.e., sociability and competence). Taken together, these findings extend and complement prior research by showing that the counterattitudinal behaviors performed by social targets promote a revision of first impressions, especially when such behaviors have a moral content.

In a similar vein, our data also show that such a key role of morality in driving impression updating goes over and beyond the perceived frequency of the behaviors. As such, rarity/frequency did not mediate our effects from moral behaviors to impression updating (Experiment 3). Moreover, our data speak against the role of perceived extremity in promoting impression updating. As such, the behaviors that we employed were carefully balanced for evaluative extremity. Moreover, given that perceived extremity and rarity tend to be highly correlated (Cone & Ferguson, 2015), the results we found in Experiment 3 – which showed no role of perceived frequency in driving our effects – suggest that other mechanisms rather than extremity may drive impression updating. Although Experiment 1 revealed that moral and immoral behaviors predicted more extreme impressions at T1 than sociable and unsociable behaviors, Experiment 2 and Experiment 3 showed that immoral behaviors predicted more extreme impressions than incompetent behaviors at T1. By contrast, positive moral and competence information did not predict different impressions at T1. Given that Experiments 2 and 3 revealed that impressions changed more strongly when both positive and negative moral information was added at T2, we can further rule out that our findings are due to the fact that moral information is more extreme to begin with.

Our findings also make a novel contribution to the literature on the implication of morality for the impression-formation process. Most studies in this area have shown that morality forms the

primary basis for the global evaluation of other people (Brambilla & Leach, 2014; Goodwin, 2015). However, these studies have considered first impressions and have overlooked that social interactions require a continuous and flexible updating of our initial evaluations. Thus, extending prior research we show that the key role of morality in shaping social cognition goes beyond the formation of initial evaluations by influencing the updating of such first impressions. Our findings are consistent with prior insights that suggest that morality provides important information to infer the intentions of social targets (for a review, Brambilla & Leach, 2014). By considering that morality is key in determining the intentions of our interaction partners, it makes sense that moral behaviors are more salient when we are called to revise our initial evaluations. The mediational model reported in Experiment 3 supports this claim.

As they stand, our findings extend prior evidence on the debate concerning the basic dimensions that underlie social cognition. In particular, our data complement recent work showing that morality and sociability represent two distinct characteristics of the general dimension of warmth/communion (Abele et al., 2016; Brambilla & Leach 2014; Leach et al. 2007; see also Brambilla et al., 2011, Brambilla et al., 2013). By showing that moral information has a leading role over sociability and competence in predicting impression change, we show that the distinction between sociability and morality is useful at different stages of person perception and impression formation.

Our findings also suggest several avenues for future research. We focused on explicit and deliberative impressions. Prior research on impression change has shown that explicit evaluations change rapidly in response to small amounts of new information. By contrast, implicit evaluations tend to change only after exposure to large amounts of counterattitudinal information (Gregg, Seibt, & Banaji, 2006; Rydell, McConnell, Strain, Claypool, & Hugenberg, 2007; for a discussion, Brannon & Gawronski, 2017). However, more recent studies show the relative ease with which implicit evaluations can also change (Cone & Ferguson, 2015; Mann & Ferguson, 2015). Based on these mixed findings, an intriguing avenue for future research would be to test the relative

importance of moral trait-content information to promote a change of both explicit and implicit impressions. In a similar vein, future work might investigate the ease with which deliberative and automatic impressions change across different contexts and situations. Indeed, the relevance of moral information in promoting impression change might vary depending on the purpose of the interaction. Moral behaviors could be taken as more relevant clues to a person's intentions by default, but under some conditions competence-related or sociability-related behaviors might be more relevant clues to establish a person's intentions.

A second area that deserves a closer inspection pertains to the (a)symmetry of our results on impression change. According to prior research on the confirmability and disconfirmability of trait content (Reeder & Brewer, 1979; Skowronski & Carlston, 1987; Trafimow & Trafimow, 1999), one can expect a greater effect of negative moral information rather than positive information on impression updating. As assumed by the model of implicational schemata (Reeder & Brewer, 1979; Skowronski & Carlston, 1987), the moral dimension is asymmetric, and negative moral behaviors are rarer and more informative than positive moral acts. As a consequence, people more heavily value negative moral behaviors when they are asked to form an impression about an unknown other because they are perceived as more diagnostic of the underlying moral nature of a social target (Brambilla & Leach, 2014; Mende-Siedlecki, Baron et al., 2013; Sanbonmatsu, Mazur, Behrends, & Moore, 2015). In some previous studies of impression updating, this valence asymmetry has also been evident. In these studies, impressions tended to undergo greater change following the addition of an (inconsistent) immoral rather than moral behavior (Reeder & Coovert, 1986; see also Skowronski & Carlston, 1992).

In contrast, our results show that both positive and negative moral information elicited impression change. Moreover, our mediational analyses did not support the claim that morality drives impression updating because (im)moral behaviors are perceived as rarer than other behaviors as prior research has argued (Mende-Siedlecki, Baron et al., 2013). Thus, our data suggested that the key role of morality in driving belief revisions goes over and beyond the statistical principle of

frequency-derived diagnosticity. Our data are in line with some previous work on inconsistency perception (Brannon, Sacchi, & Gawronski, 2017). These works did not show valence asymmetry in eliciting expectancy-violation effects. These studies show that positive and negative information elicited equally strong expectancy-violation effects. Similarly, prior research on ERP has also found no evidence for valence asymmetries in that participants showed similar neural activity in response to expectancy-violating information regardless of whether this information was positive or negative (Cacioppo, Crites, Berntson, & Coles, 1993).

The presence or absence of valence asymmetries represents an interesting nuance that is worthy of further consideration. It may be possible that the negativity effect of morality might involve only specific aspects of impression formation and not every aspect of such a process. In the case of impression updating, it would be worthwhile to systematically investigate – by even considering different measures and social targets – the factors that might promote or suppress the negativity effect on morality when studying the development of our impressions over time. These points considered, our findings provide a better understanding of the social consequences of being moral and of the role of moral character in everyday life more generally.

References

Abele, A. E., & Bruckmuller, S. (2011). The bigger one of the "Big Two"? Preferential processing of communal information. *Journal of Experimental Social Psychology*, *47*, 935-948.

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*.

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.

Abele, A.E. & Wojciszke, B. (2014). Communal and agentic content in social cognition: A Dual Perspective Model. *Advances in Experimental Social Psychology, 50*, 195-255.

Asch, S. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 1230-1240.

Anderson, N. H. (1962). Application of an additive model to impression formation. *Science, 138*, 817-818.

Brambilla, M., Biella, M., & Freeman, J.B. (2018). The influence of visual context on the evaluation of facial trustworthiness. *Journal of Experimental Social Psychology, 78*, 34-42.

Brambilla, M., & Leach, C.W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition, 32*, 397-408.

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.

Brambilla, M., Sacchi, S., Menegatti, M., & Moscatelli, S. (2016). Honesty and dishonesty don't move together: Trait content information influences behavioral synchrony. *Journal of Nonverbal Behavior, 40,* 171-186.

Brambilla, M., Sacchi, S., Pagliaro, S., & Ellemers, N. (2013). Morality and intergroup relations: Threats to safety and group image predict the desire to interact with outgroup and ingroup members. *Journal of Experimental Social Psychology, 49*, 811-821.

Brambilla, M., Sacchi, S., Rusconi, P., & Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, *51*, 149-166.

Brannon, S. M., & Gawronski, B. (2017). A second chance for first impressions? Exploring the context-(in) dependent updating of implicit evaluations. *Social Psychological and Personality Science*, *8*, 275-283.

Brannon, S. M., Sacchi, D. L., & Gawronski, B. (2017). (In) consistency in the eye of the beholder: The roles of warmth, competence, and valence in lay perceptions of inconsistency. *Journal of Experimental Social Psychology, 70*, 80-94.

Cacioppo, J. T., Crites, S. L., Jr., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science, 4*, 108-112.

Cone, J., & Ferguson, M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology, 108*, 37-57.

Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Changing our implicit minds: How, when, and why implicit evaluations can be rapidly revised. I*n Advances in Experimental Social Psychology* (Vol. 56, pp. 131-199). Academic Press.

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.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The Stereotype Content Model and the BIAS Map. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 61-149). San Diego: Academic

Press.

De Bruin, E. N. M., & Van Lange, P. A. M. (1999). Impression formation and cooperative

behavior. *European Journal of Social Psychology*, *29*, 305-328.

Dunning, D. (2004). On the motives underlying social cognition. In M. Brewer & M. Hewstone

(Eds.), *Emotion and motivation* (pp. 137-164). Oxford: Blackwell.

Ellemers, N., & Van den Bos, K. (2012). On the social regulatory functions of right and wrong.

Social and Personality Psychology Compass, 6, 878-889.

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic

coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19*,

1508-1519.

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.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from Daguerreotype to

laserphoto. *Journal of Personality and Social Psychology, 63*, 877-889.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype

content: Competence and warmth respectively follow from perceived status and

competition. *Journal of Personality and Social Psychology*, *82*, 878-902.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in

evaluation: An integrative review of implicit and explicit attitude change. *Psychological

Bulletin, 132*, 692-731.

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological

Science, 24*, 38-44.

Goodwin, G.P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception

and evaluation. *Journal of Personality and Social Psychology, 106*, 148-168.

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

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment

model. *Cognitive psychology, 24*, 1-55.

Iachini, T., Pagliaro, S., & Ruggiero, G. (2015). Near or far? It depends on my impression: Moral

information and spatial behavior in virtual interactions. *Acta Psychologica, 161*, 131-136.

Judd, C.M., James-Hawkins, L., Yzerbyt, V., & 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.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of learning and motivation, 32*, 385-

418.

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.

Leidner, B., & Castano, E. (2012). Morality shifting in the context of intergroup violence. *European

Journal of Social Psychology, 42*, 82-91.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of

faces and norming data. *Behavior Research Methods, 47*, 1122-1135

Mann, T. C., & Ferguson, M. J. (2015). Can we undo our first impressions? The role of

reinterpretation in reversing implicit evaluations. *Journal of Personality and Social

Psychology, 108*, 823-849.

Mann, T. C., & Ferguson, M. J. (2017). Reversing implicit first impressions through reinterpretation

after a two-day delay. *Journal of Experimental Social Psychology, 68,* 122-127.

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric

updating of impressions in the morality and ability domains. *The Journal of Neuroscience,

33*, 19406-19415.

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person

impressions. *Social Cognitive and Affective Neuroscience, 8*, 623-631.

Mende-Siedlecki, P., & Todorov, A. (2016). Neural dissociations between meaningful and mere inconsistency in impression updating. *Social Cognitive and Affective Neuroscience, 11,* 1489-1500.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology, 2*, 175-220.

Phalet, K., & Poppe, E. (1997). Competence and morality dimensions of national and ethnic stereotypes: a study in six eastern-European countries. *European Journal of Social Psychology, 27*, 703-723.

Poppe, E., & Linssen, H. (1999). In-group favouritism and the reflection of realistic dimensions of difference between national states in Central and Eastern European nationality stereotypes. *British Journal of Social Psychology, 38*, 85-102.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review, 86*, 61-79.

Reeder, G. D., & Coovert, M. D. (1986). Revising an impression of morality. *Social Cognition, 4*, 1-17.

Rodriguez Mosquera, P. M., Manstead, A. S. R., & Fischer, A. H. (2002). The role of honour concerns in emotional reactions to offences. *Cognition and Emotion, 16,* 143-163.

Rosenberg, S., Nelson, C., & Vivekananthan, P.S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*, 283-294.

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*, 131-142.

Rule, N.O., Slepian, M.L., & Ambady, N. (2012). A memory advantage for untrustworthy faces. *Cognition, 125*, 207-218.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*, 995-1008.

Rydell, R. J., McConnell, A. R., Strain, L. M., Claypool, H. M., & Hugenberg, K. (2007). Implicit and explicit attitudes respond differently to increasing amounts of counterattitudinal information. *European Journal of Social Psychology, 37,* 867-878.

Sanbonmatsu, D. M., Mazur, D., Behrends, A. A., & Moore, S. M. (2015). The role of the baserate frequency of correspondent behavior and trait stereotypes in attribution: building on Rothbart and Park (1986). *Social Cognition, 33*, 255-283.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52*, 689-699.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131-142.

Skowronski, J. J., & Carlston, D. E. (1992). Caught in the act: When impressions based on highly diagnostic behaviours are resistant to contradiction. *European Journal of Social Psychology, 22*, 435-452.

Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology, 23*, 373-380.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12*, 455-460.

Todorov, A., Said, C. P., Oosterhof, N. N., & Engell, A. D. (2011). Task-invariant brain responses to the social value of faces. *Journal of Cognitive Neuroscience, 23*, 2766-2781.

Trafimow, D., & Trafimow, S. (1999). Mapping imperfect and perfect duties onto hierarchically and partially restrictive trait dimensions. *Personality and Social Psychology Bulletin, 25*, 686-695.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), S*ocial psychology: Handbook of basic principles* (pp. 239- 270). New York: Guilford Press.

van der Toorn, J., Ellemers, N., & Doosje, B. (2015). The threat of moral transgression: The impact of group membership and moral opportunity. *European Journal of Social Psychology, 45*, 609-622.

Willis, J., & Todorov, A. (2006). First impressions : Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592-598.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*, 277-283.

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence or morality. *Journal of Personality and Social Psychology, 67*, 222- 232.

Wojciszke, B. (2005). Morality and competence in person- and self-perception. *European Review of Social Psychology, 16,* 155-188.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin, 24*, 1251-1263.

Wyer, N. A. (2010). You never get a second chance to make a first (implicit) impression: The role of elaboration in the formation and revision of implicit impressions. *Social Cognition, 28*, 1-19.

Ybarra, O., Chan, E., & Park, D. (2001). Young and old adults' concerns about morality and competence. *Motivation and Emotion*, *25*, 85-100.

Zebrowitz, L., & Collins, M. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review*, *1*, 204-223.

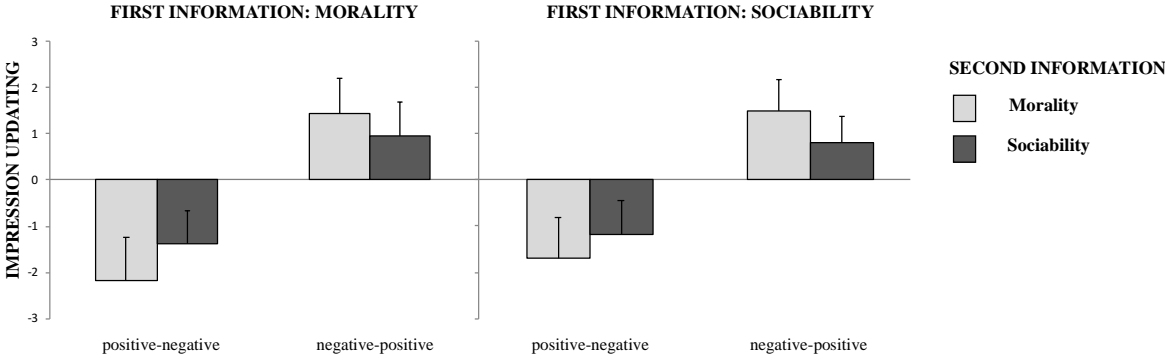*Figure 1*. Impression Updating - Experiment 1.

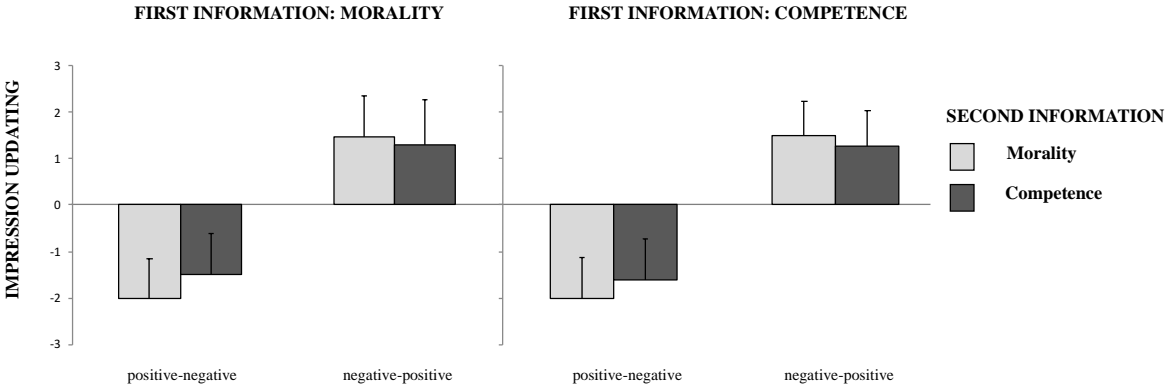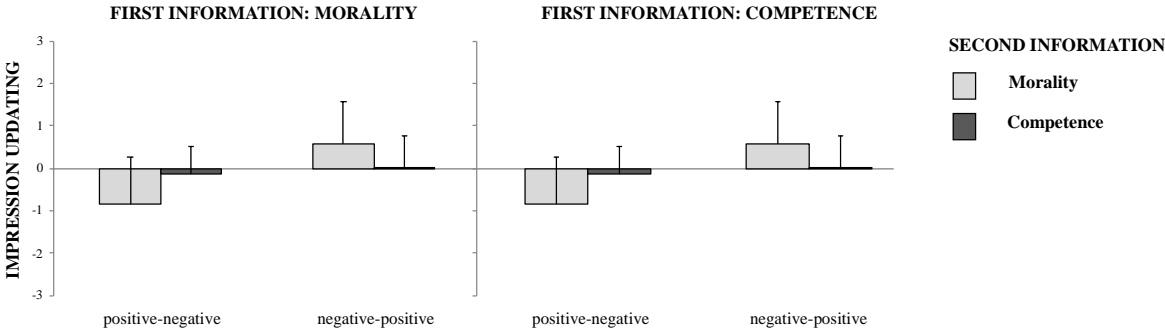*Figure 2*. Impression Updating - Experiment 2.

Figure 3. Impression Updating - Experiment 3.

Open Practices

Supplementary materials provided details of the measures employed, so that independent

researchers can reproduce the methodology.

Appendix A. Supplementary Materials

Supplementary data to this article and the materials for all experiments can be found online at

Footnotes

[1] To ascertain that the selected behaviours were unambiguously classified in only one of the three content dimensions, 32 students rated the behaviours employed in Experiment 1 and Experiment 2 on their competence-, sociability-, and morality-relatedness on scales ranging from 1 (*not at all*) to 7 (*extremely*). As expected, the within-subjects ANOVA yielded a significant interaction effect, $F(4, 124) = 101.18$, $p < .001$, $\eta_p^2 = .76$. Thus, moral behaviors (positive and negative) were rated as more related to morality than to sociability ($p<.001$) and competence ($p<.001$). Similarly, sociability behaviours were rated as higher on sociability than on competence ($p<.001$) and morality relatedness ($p<.001$). Finally, competence behaviours were rated as higher on competence than on morality ($p<.001$) and sociability relatedness ($p<.001$). Importantly, only moral behaviours were rated above the scale mean on morality relatedness, $t(31)=14.69$, $p<.001$. In a similar vein, only sociability behaviours were rated above the scale mean on sociability relatedness, $t(31)=12.69$, $p<.001$. Finally, only competence behaviours were rated above the scale mean on competence relatedness, $t(31)=10.91$, $p<.001$. Thus, the selected behaviours were unambiguously classified in only one content dimension. To exclude that valence would drive our findings, we conducted an additional pilot study. 111 students were asked to rate the valence of the behaviors on a scale ranging from -3 (very negative) to +3 (very positive). To verify that positive and negative behaviors were equally polarized, we computed a 3 (dimension: morality, sociability, competence) × 2 (valence: positive vs. negative) between-subjects ANOVA using the valence ratings in absolute value. The ANOVA did not yield the interaction effect between dimension and valence, $F<1$, $p=.51$. Thus, the behaviors we employed were equivalent in evaluative extremity.

[2] The ANOVA showed the unexpected interaction effect between the first and the second behavior dimensions, $F(1, 39) = 12.03$, $p = .001$, $\eta_p^2 = .24$. Whereas the second clue related to sociability had the same impact on the impression based on morality ($M = -.20$, $SD = .27$) and on sociability ($M = -.20$, $SD = .30$), $t(39) = .02$, $p =.98$, the second clue related to morality had a

greater effect on the impression based on morality ($M = -.37$, $SD = .38$) rather than sociability ($M = -.11$, $SD = .43$), $t(39) = 3.05$, $p = .004$. Such a pattern of data suggests that the social judgment based on moral clues is less stable. This result is consistent with the model of the hierarchically restrictive schemata applied to the moral domain suggesting that the moral image might be more mutable than that sociable one, as sociability is a more symmetric dimension (Reeder & Brewer, 1979; Skowronski & Carlston, 1987). This process could also explain the main effect of the first behavioral dimension, $F(1, 39) = 4.83$, $p = .03$, $\eta_p^2 = .11$, and the interaction between the first behavioral dimension and valence, $F(1, 39) = 11.12$, $p = .002$, $\eta_p^2 = .22$. Indeed a positive piece of information had the same impact on impression updating when following a morality-related negative clue ($M = 1.20$, $SD = .73$) or a sociability-related negative clue ($M = 1.14$, $SD = .53$), $t(39) = .83$, $p = .41$. In contrast, a negative piece of information had a weaker negative impact, $t(39) = 3.47$, $p = .001$, when following a sociability-related positive clue ($M = -1.46$, $SD = .76$) than a morality-related positive clue ($M = -1.78$, $SD = .77$) which should be easier to disconfirm according to the previous model.

[3] Importantly, while in the first two experiments the screen displayed both the initial and the secondary information before participants expressed their final evaluation, in Experiment 3 the earlier information was out of sight during the T2 evaluation. This helped us to test our hypothesis in a more conservative way by preventing that participants aggregated two simultaneous sources of information during the impression-updating task.

[4] To ascertain that the 24 behaviours employed in Experiment 3 were unambiguously classified in only one of the two content dimensions, 35 students rated the behaviours on their competence-, and morality-relatedness on scales ranging from 1 (*not at all*) to 7 (*extremely*). As expected, the within-subjects ANOVA yielded a significant interaction effect, $F(1,33) = 300.15$, $p < .001$. Thus, moral behaviors (positive and negative) were rated as more related to morality than to competence ($p<.001$). By contrast, competence behaviours were rated as higher on competence than

on morality ($p<.001$). Importantly, only moral behaviours were rated above the scale mean on morality relatedness, $t(34)=15.99$, $p<.001$. In a similar vein, only competence behaviours were rated above the scale mean on competence relatedness, $t(33)=13.21$, $p<.001$. Thus, the selected behaviours were unambiguously classified in only one content dimension.

To exclude that valence would drive our findings, we conducted an additional pilot study. 40 students were asked to rate the valence of the behaviors employed in Experiment 3, on a scale ranging from -3 (very negative) to +3 (very positive). To verify that positive and negative behaviors were equally polarized, we computed a 2 (dimension: morality vs. competence) × 2 (valence: positive vs. negative) between-subjects ANOVA using the valence ratings in absolute value. The ANOVA did not yield the interaction effect between dimension and valence, $F<1$, $p=.46$. Thus, the behaviors we employed were equivalent in evaluative extremity.