Why So Serious?

A Laboratory and Field Investigation of the Link between Morality and Humor

Kai Chi Yam*
National University of Singapore
bizykc@nus.edu.sg

Christopher M. Barnes
University of Washington
chris24b@uw.edu

Keith Leavitt
Oregon State University & Bocconi University
keith.leavitt@bus.oregonstate.edu

Wu Wei
Wuhan University
weiwumail@163.com

Tsz Chun Lau
National University of Singapore
obrltc@nus.edu.sg

Eric Luis Uhlmann
INSEAD
eric.uhlmann@insead.edu

*Corresponding author

Author Notes: We thank Scott J. Reynolds and seminar participants at various institutions for constructive comments on earlier drafts of this manuscript. This research was funded by a Tier 1 research grant (R317000132) by the Singapore Ministry of Education awarded to the first author, the National Social Science Key Fund of China under grant 18AGL006 awarded to the fourth author, and a R&D grant from INSEAD awarded to the last author.
Abstract

Previous research has identified many positive outcomes resulting from a deeply held moral identity, while overlooking potential negative social consequences for the moral individual. Drawing from Benign Violation Theory, we explore the tension between moral identity and humor, and the downstream workplace consequence of such tension. Consistent with our hypotheses, compared to participants in the control condition, participants whose moral identities were situationally activated (Study 1a) or chronically accessible (Study 1b) were less likely to appreciate humor and generate jokes others found funny (Study 2), especially humor that involved benign moral violations. We also found that participants with a strong moral identity do not generally compensate for their lack of humor by telling more jokes that do not involve moral violations (Study 3). Additional field studies demonstrated that employees (Study 4) and leaders (Study 5) with strong moral identities and who display ethical leadership are perceived as less humorous by their coworkers and subordinates, and to the extent that this is the case are less liked in the workplace. Study 5 further demonstrated two competing mediating pathways – leaders with strong moral identities are perceived as less humorous but also as more trustworthy, with differentiated effects on interpersonal liking. Although having moral employees and leaders can come with many benefits, our research shows that there can be offsetting costs associated with an internalized moral identity: reduced humor and subsequent likeability in the workplace.

Keywords: Humor; Benign Violation Theory; Behavioral Ethics; Morality
Why So Serious?

A Laboratory and Field Investigation of the Link between Morality and Humor

“There is no humor in heaven”
(Mark Twain, 1910)

Behavioral ethics research has advanced the very useful theory of moral identity, with the construct of moral identity defined as “a self-conception organized around a set of moral traits” (Aquino & Reed, 2002, p. 1423). Indeed, a large empirical base provides strong support for this theory, demonstrating how a strong moral identity leads to desirable ethical behavior, including reduced unethical behavior and increased ethical leadership (Detert, Trevino, & Sweitzer, 2008; Mayer, Aquino, Greenbaum, & Kuenzi, 2012; Reynolds & Ceramic, 2007; for a review, see Shao, Aquino, & Freeman, 2008). The uniformly positive view of moral identity, however, overlooks the possibility that there are potential social costs to having a strongly internalized moral identity. While highly moral people are often viewed positively for displaying admirable character traits, they may also be disliked to the extent to which they are viewed as sanctimonious, prudish, or unrelatable (Monin, Sawyer, & Marquez, 2008).

Accordingly, we draw from Benign Violation Theory (BVT; McGraw & Warren, 2010) to identify a potential risk associated with a chronically accessible or situationally activated moral identity: a diminished sense of humor. Briefly stated, BVT is a theory that explains the construction of humor (i.e., what makes things funny), while also drawing boundaries around absurdities which would be considered taboo or offensive. McGraw and Warren (2010) proposed that humor often emerges when a norm is violated, yet the violation is simultaneously appraised as benign. Drawing from BVT, we propose that individuals with strong (chronic or situationally activated) moral identities will be less likely to engage in the benign moral violations common to the display of humor, as condemnatory moral judgments serve to uphold valued social norms and
create distress for individuals who even contemplate such violations (Tetlock, Kristel, Elson, Green, & Lerner, 2000). As the opening quote implies, we suggest that there exists a tension between being moral and being humorous, in that entertaining some forms of humor are considered proscribed by individuals with a deeply-held, internalized sense of morality. Moreover, we posit that this will have negative downstream implications for the way in which employees with strong moral identities are perceived by their colleagues. In particular, a growing literature suggests that humor is useful in building and maintaining likeability at work. Research indicates that most people prefer to interact with humorous than non-humorous others (Bressler, Martin, & Balshine, 2006), and that humor is a valued attribute within social exchanges (Lundy, Tan, & Cunningham, 1998; Mesmer-Magnus, Glew, & Viswesvaran, 2012). Although highly moral individuals should be no less likely to engage in forms of humor that do not involve any moral violation (e.g., innocuous linguistic puns), their avoidance of more off-color jokes that challenge moral norms may lead them to be seen as less humorous overall. We therefore propose that employees with strong moral identities are less likely to be perceived as humorous, which in turn can lead to reduced likability in the workplace. To investigate these ideas, we conducted six studies and employed multiple methodologies including two controlled experiments, two lab studies, a multi-source field study conducted in China, and a three-wave field study conducted in the United States.

Through these studies, we make a number of important theoretical contributions to research on moral identity and social judgments in the workplace. By drawing on recent developments in our theoretical understanding of humor (McGraw & Warren, 2010), we build theory and provide both experimental and field evidence that moral identity and certain forms of humor are negatively related as a function of proscriptions around moral violations. We also
extend research on humor by examining its antecedents, identifying psychological processes (i.e., accessibility of moral identity) and categories of individuals (i.e., individuals for whom moral identity is chronically accessible) that are least likely to be humorous. Finally, this research provides evidence regarding one of the reasons why individuals with strong moral identities are sometimes liked less, offering a complementary explanation to existing accounts of unfavorable perceptions of moral individuals (Monin et al., 2008; Parks & Stone, 2010; Trevino & Victor, 1992). Taken together, our findings suggest a previously unanticipated negative consequence of reinforcing morality—those with strong moral identities eschew some forms of humor, and because of this they may be liked less.

MORAL IDENTITY, HUMOR, AND BENIGN VIOLATION THEORY

Philosophers and psychologists have long hinted at a link between morality and humor. At the dawn of Western philosophy, both Plato and Aristotle conceptualized humor as a form of mockery (Aristotle, 1939). They argued that humor emerges when people exert superiority over others or their former selves, a theory later formalized as superiority theory of humor (Gruner, 1997). In contemporary psychological research, laughing at others’ misfortune or intentionally exerting superiority by making fun of others (e.g., a boss ridicules his/her followers publicly) is often characterized as deviant or even unethical behavior (Bennett & Robinson, 2000). Thus, the superiority theory hints at a potentially negative relationship between morality and humor.

Other conceptualizations, however, challenge this superiority account of humor. Freud (1928) proposed the relief theory of humor, suggesting that the main function of humor is to relieve sublimated desires and emotions in socially acceptable ways. For example, telling a sexual joke is proposed to release repressed hostile sexual desires. Thus, according to this theory humor is used to express socially unacceptable desires in a playful manner. Another early theory
of humor is incongruity theory (Suls, 1972). This account of humor suggests that humor emerges when unexpected things happen. For example, stand-up comedians often create an expectation at the beginning of a joke (e.g., A man in the library goes up to the desk and asks for a burger. The librarian says, “Sir this is a library.”) and then violate it at the end (“Oh I’m sorry, and leans over and whispers, “Can I get a burger?”) to generate humor (for a review, see Martin, 2001).

Although these theories (particularly incongruity theory) have dominated Western thoughts on humor and humor research for many years (for a review, see Mulder & Nijholt, 2002), McGraw and Warren (2010) argued that these theories often erroneously predict humor. For example, “unintentionally killing a loved one would be incongruous, assert superiority, and release repressed aggressive tension, but is unlikely to be funny” (pp. 1-2; also see Warren & McGraw, 2016). Therefore, McGraw and colleagues (McGraw & Warner, 2014; McGraw & Warren, 2010; McGraw, Williams, & Warren, 2014; McGraw, Warren, Williams, & Leonard, 2012) proposed Benign Violation Theory, which better predicts a wide-range of humor while also drawing careful boundaries that exclude events which are simply esoteric, tragic, or perverse. BVT makes three interrelated predictions about the generation of humor. First, a norm violation must occur, and that the violation can be physical or symbolic in nature, or both. For example, humans often laugh when an expected norm is violated (e.g., “I once farted in an elevator. It was wrong on so many levels;” McGraw & Warren, 2010). Second, the violation must be perceived as benign. Third, the first two conditions must occur simultaneously (McGraw & Warren, 2010). Empirical support for this theory comes from multiple studies (e.g., Warren & McGraw, 2016). For example, McGraw et al. (2013) found that jokes about Hurricane Sandy’s destruction during the crisis were considered offensive and malicious, thus not funny. However,
the same jokes became more innocuous over time, and people found those jokes to be funny about a month after the crisis.

Critically, BVT does not suggest that all humor is rooted in benign violations (leaving room for other sources of humor), but rather argues that the broadest domain of humor is produced by leveraging benign violations. Accordingly, we next argue that an individual’s moral identity is inhibitory of his/her overall level of humor, because it restricts the domain of what is treated as benign, and consequently limits instances of humor expression and appreciation. Importantly, we suggest that an internalized sense of morality is associated with less total humor, but only because it diminishes appreciation and generation of humor involving moral violations specifically. Individuals with a strong moral identity should be no less likely to engage in humor that does not involve transgressing moral norms (e.g., innocuous linguistic puns).

The Impact of Moral Identity on Humor

As noted above, moral identity involves a self-conception organized around a set of moral traits (e.g., caring, honest, fair, etc.; Aquino & Reed, 2002). Across a variety of samples, people who highly value these traits are especially likely to engage in ethical and refrain from unethical behavior. Subsequent research has replicated and expanded upon this result using a variety of research designs (e.g., Detert et al., 2008; Mayer et al., 2012, Reynolds & Ceramic, 2007). For example, a strong moral identity predicts an enhanced concern for the interests of a wide range of out-groups (e.g., individuals from another race or country; Reed & Aquino, 2003) and increased cooperative and pro-social behavior directed at others (Aquino, Freeman, Reed, Lim, & Felps, 2009; Aquino, McFerran, & Laven, 2011).

Critically, Aquino and colleagues (2009) proposed and found strong support for a social-cognitive model of moral identity, arguing that moral identity is a self-relevant schematic
knowledge structure which informs cognition and behavior, and that the activation of moral identity (either temporary or chronic) drives the pursuit of moral goals and moral behavior. Because the working self-concept is populated by both chronically accessible and situationally-relevant identities within a given situation (Markus & Kunda, 1986), situations which activate moral facets of the self in the working self-concept lead people for whom moral identity is less central to behave similarly to those for whom moral identity is especially self-defining across contexts (i.e., chronically active). Aquino and colleagues (2009) thus identified the underlying mechanism of the effects of a moral identity on behavior as the momentary accessibility of the moral self-knowledge structure, and found that ostensibly unrelated cues (e.g., reading a religious text) were sufficient to activate a moral identity among those for whom a moral identity was not centrally self-defining. Accordingly, regardless of whether an individual strongly identifies with morality across situations (“trait-like” moral identity), or the context itself brings to mind temporarily active moral self-schema (“state-like” moral identity), the activation of a moral identity in the working self-concept has been shown to be sufficiently powerful to motivate the pursuit of moral goals (Aquino et al., 2009).

Drawing from BVT, we suggest that individuals with active moral identities are less likely to generate or appreciate humorous content that entails violations of moral norms. Indeed, morality often functions to constrain and restrict behavior to maintain social standing in the eyes of a broader moral community (Tetlock, 2002). As such, even minor violations of such sacred values generally cause psychological distress and trigger self-focused moral cleansing (Tetlock et al., 2000). Moreover, empirical research has also identified a “mere contemplation effect,” in that even thinking about a normative transgression that violates one’s sacred values is sufficient to trigger cognitive rigidity and motivate the individual to reaffirm their virtue (Tetlock, 2002;
Tetlock et al., 2000). We propose that individuals would be especially motivated to avoid morally-questionable cognitions when their moral identity is active, as previous research has shown that the activation of moral identity makes moral concerns especially salient (Aquino et al., 2009). Thus, when one’s moral identity is active in the working self-concept, showing public appreciation for norm-violating humor (such as laughing at a sexist joke told by a colleague) may trigger self-regulatory concerns that the individual is not living up to important moral standards. Thus, individuals not only limit their own public behavior to uphold their moral standards, but actually censor and sanction their own cognition to avoid violating moral standards (Tetlock et al., 2000). Accordingly, activation of moral identity should increase motivation to censor thoughts that undermine valued social norms, as moral goals become especially salient when the moral identity is active (Aquino et al., 2009). We propose that when the moral identity is active in the working self-concept, it should reduce the individual’s motivation to violate moral norms and willingness to engage in the type of proscribed cognition necessary for generating norm-violating humor in the first place. Thus, moral identity should diminish appreciation of norm-violating humor, as well as impair the ability to think about the type of norm violations required for the production of humor.

Of course, not all norms are morally-laden. For example, violations of innocuous linguistic norms (e.g., puns) can be perceived as benign and should thus remain humorous even to individuals whose moral identities are strong. However, we suggest that individuals with strong moral identities are less likely to be humorous overall, because they are less likely to engage in the broader swath of benign moral violations often involved in the display and appreciation of humor. Thus, an active moral identity places a restriction around the types of norms which can be permissibly violated as a tool for producing humor and sensitizes the
individual to moral violations, thereby limiting the sources of humor available to those for whom a moral identity is highly accessible.

**Implications for Interpersonal Liking**

We argue that to the extent individuals with strong moral identities are perceived as less humorous, this will lead to a negative downstream social consequence for them in the workplace, namely reduced likability. A handful of empirical studies have identified some circumstances in which exceptionally competent and morally-outstanding individuals elicit negative reactions from others (Trevino & Victor, 1992). Work to date has largely focused on disadvantageous upward comparisons as the causes of the negative reactions to highly moral individuals. For example, participants who refused to participate in a racist task were viewed less favorably by others who had gone through the task and failed to object to it (Monin et al., 2008). Likewise, participants who were the most selfless and prosocial were more likely to be expelled from a group than participants who were less selfless (Parks & Stone, 2010), and this relationship is in part driven by peers’ upward social comparisons of morality (Minson & Monin, 2012; Monin et al., 2008; Parks & Stone, 2010).

Beyond social comparison processes, however, there may be additional reasons why highly moral people are sometimes disliked by others. We suggest that the negative association between moral identity and humor is likely to have implications for how individuals with strong moral identities are treated by others. If individuals with strong moral identities are less humorous as we hypothesize, and a sense of humor generally increases interpersonal liking (Cooper, 2005), then we suggest that there may be an indirect effect from an individual’s moral identity to reduced likability, via reduced humor. One’s likeability has significant implications at the workplace. For example, how much a follower is being liked is positively associated with
subjective ratings of job performances (Wayne & Ferris, 1990). Likewise, Heilman and Okimoto (2007) found that being liked or behaving in communal ways can mitigate the penalties that women incur for success in traditionally male areas. All in all, likability is an important factor for employees to be successful at work.

Although humor has been theorized as a double-edged sword (Malone, 1980), decades of research generally suggests that humorous individuals are often liked in social interactions. For example, humorous individuals are perceived as more socially adept by the opposite sex (Bressler & Balshine, 2006) and a sense of humor predicts mating success (Greengross & Miller, 2011). Humor can also be conceptualized as part of the glue of friendship (Kalbfleischl, 2013). In the workplace, humor is theorized as an effective ingratiation tactic for gaining social visibility and capital (Cooper, 2005). These findings can also be explained through research in person perception, as humorless individuals may be categorized with other negative traits in the perception formation stage (Cantor & Mischel, 1979). Research thus converges on the positive effects of humor on forming and maintaining social relationships. Therefore, if an individual’s moral identity is negatively related to humor as we hypothesize, then it is likely that to the extent they are seen as less humorous their likeability is also negatively affected.

**RESEARCH OVERVIEW**

We conducted six studies testing these ideas. In Studies 1a and 2, we experimentally manipulated the temporary accessibility of participants’ moral identities and examined its effects on humor appreciation and production in order to establish a causal relationship. We conceptually replicated these findings with a trait measure of chronic moral identity in Study 1b. An additional goal of these three studies was to examine the role of moral violations as a mechanism that underlies the negative relationship between moral identity and humor through both moderation (Studies 1a and 1b) and mediation (Study 2) approaches (Spencer, Zanna, &
Fong, 2005). After we provide evidence that individuals with strong moral identities are generally less humorous because they are unwilling to tell or appreciate humor that involves moral violations, we conducted Study 3 to examine whether individuals with strong moral identities compensate this deficit in humor by telling more jokes that do not involve moral violations. Finally, we conducted two additional field studies to examine the implications of a lack of humor on interpersonal liking for employees (Study 4) and leaders (Study 5) who are perceived by others as ethical. Study 5 further demonstrated the incremental validity of our proposed mechanism (i.e., reduced humor) above and beyond upward social comparisons with the moral leader (Parks & Stone, 2010).

**STUDY 1A**

In our first study, we examined whether temporarily activating a moral identity would negatively affect humor appreciation. We chose to examine this dependent variable, broadly defined as an ability to understand and enjoy messages containing humor, because being able to appreciate and recognize humor is an important aspect in a person’s overall sense of humor (Thorson & Powell, 1993). Indeed, it would be quite difficult to share humor with someone else if a person is unable to even spot it himself/herself.

**Participants**

We recruited 300 adults from Prolific for the study (Peer, Brandimarte, Samat, & Acquisti, 2017). We dropped 14 participants because they failed to write at least one full sentence in the manipulation task (described later), leading to a final sample size of 286 ($M_{age} = 35.33$; 43.4% female; 78.7% White). All participants received a small cash payment.

**Procedures**

Participants were randomly assigned to either a moral identity activation or control condition. In the moral identity condition, we employed a structured recall task leading
participants to bring to mind instances of themselves as moral people, with the assumption that they would recall instances that best matched their own prototypes for morality. Participants were specifically instructed to recall a time in which they themselves behaved in an especially moral way. After the recall task, participants were asked to write two to three sentences to describe what happened. In the control condition, participants were asked to recall what they did the day before. This manipulation has been used in research on social identity and is particularly useful in activating a specific aspect of one’s identity (e.g., Leavitt, Reynolds, Barnes, Schilpzand, & Hannah, 2012; Libby, Eibach, & Gilovich, 2005; Peters & Gawronski, 2011).

After the manipulation, participants were provided with 18 jokes and written captions and were asked to rate them on a 7-point scale (1 not funny to 7 = very funny; see the online supplement for the materials). We selected these jokes based on an archive of jokes varying in degree of funniness as rated by thousands of subjects (Wiseman, 2002). Nine of the jokes/captions contained moral violations (most of them involve hurting someone, either physically or emotionally, or jokes that involve sexual content), whereas the other nine did not. All jokes/captions were presented in random order. In a separated data collection, a total of 60 participants recruited from Amazon Mechanical Turk ($M_{age} = 29.48$, 59.6% male, 70.7% White) who were blind to both experimental condition and our hypotheses were asked to rate the degree of moral violation for each joke and caption (1 = no moral violation to 5 = definitely contains moral violations). Jokes and captions in the moral violation category were evaluated by these independent raters as containing a significantly higher degree of moral violation ($M = 2.89$, $SD = .53$) than jokes/captions that do not involve moral violations ($M = 1.71$, $SD = .46$), $t(59) = 14.21$, $p < .001$.\(^1\) We therefore examined the underlying process using a moderation approach,

\(^1\) To triangulate this check, we asked five independent researchers, with an average of over 15 years of experience researching behavioral ethics, to follow the same procedures and to rate the degree of moral violation for each joke
hypothesizing that activating a moral identity would generally reduce humor appreciation, and especially reduce appreciation for jokes that involved a moral violation (Spencer et al., 2005).

**Results**

We first examined the main effect of activating a moral identity on humor appreciation across both types of jokes/captions. Participants in the moral identity condition were less likely to appreciate humor ($M = 3.88, SD = 1.00$) than participants in the control condition ($M = 4.16, SD = 1.04$), $t(284) = -2.35, p = .02, d = -.27$. We then tested our hypothesized interaction with a mixed factorial experimental condition (between) X violation type (within) ANOVA. The expected interaction effect emerged between the experimental condition and appreciation of the two different types of humor, $F(284) = 13.57, p < .001, \eta^2 = .05$. Participants in the moral identity condition were less likely to appreciate humor associated with a moral violation ($M = 3.56, SD = 1.21$) than participants in the control condition ($M = 4.06, SD = 1.19$), but participants in the moral identity condition were no less likely to appreciate humor that contained no moral violation (see Figure 1).

---

**STUDY 1B**

Given that our theorizing does not make different predictions for the temporary (state-like) versus chronic (trait-like) accessibility of moral identity, we sought to conceptually replicate Study 1a’s findings with a trait measure of chronic moral identity.

**Participants**

and caption. Jokes and captions in the moral violation category were again rated as containing a significantly higher degree of moral violation ($M = 2.38, SD = 1.00$) than jokes/captions that do not involve moral violations ($M = 1.13, SD = .14$), $t(4) = 3.16, p = .03$. 
We recruited 120 adults (M_{age} = 30.03; 53.3\% male; 73.3\% White) from Amazon Mechanical Turk for the study. All participants were compensated a small cash payment.

**Procedures**

We first asked participants to complete the five-item internalization dimension of the moral identity scale (Aquino & Reed, 2002). This measure assessed the extent to which being “moral” is believed to be particularly self-defining and central to one’s self-concept, and is thus appropriate as assessing the extent to which moral content was chronically accessible to participants (Aquino et al., 2009). Critically, research on moral identity has identified two modestly correlated (and independently assessed) facets of moral identity: internalization and symbolization (Aquino & Reed, 2002). Whereas internalization represents the extent to which an individual finds moral traits to be especially valued and self-defining, symbolization reflects the extent to which an individual is motivated to outwardly display such traits (e.g., clothing choices; Aquino & Reed, 2002). The internalization facet of moral identity was chosen for two reasons. First, internalization aligns with the notion of morality motivating self-regulation of proscribed cognition, in that internalization of an identity describes the extent to which the individual believes a set of (moral) traits to be particularly self-defining and important to their own self-concept (Aquino & Reed, 2002; Erikson, 1964) with implications for how they think and feel (Kihlstrom & Klein, 1994). To wit, the internalization (but not the symbolization) facet of moral identity is significantly correlated with implicit associations between categories of “self” and “moral” (Aquino & Reed, 2002). Second, previous research has generally found that the internalization facet is a more reliable predictor of focal actor behavior for most relevant outcomes (Mayer, Aquino, Greenbaum, & Kuenzi, 2012), whereas symbolization appears more relevant for predicting the perceptions of observers (Desai & Kouchaki, 2017). Thus,
symbolization reflects what others see directly, but internalization reflects the filter through which the focal individual makes self-judgments and thus navigates the world.

Participants were first provided with a list of positive moral traits (e.g., caring, fair) and were asked to visualize the kind of person who has these characteristics and imagine how that person would think, feel, and act. Afterward, participants were asked to complete the five-item internalization of moral identity scale (see online supplement for the full scale). A sample item is “I strongly desire to have these characteristics” (1 = strongly disagree to 5 = strongly agree; α = .87). After completing this survey, participants were directed to complete the same humor appreciation tasks as in Study 1a.

Results

We first examined the main effect of a strong moral identity on humor appreciation across both types of jokes/captions. As expected, moral identity was associated with reduced humor appreciation in general (β = -.24, p = .008, R² = .05). We then tested our hypothesized interaction with two regression models, given that we have two dependent variables (i.e., humor that involves moral violations vs. humor that does not involve moral violations). A strong moral identity was associated with reduced humor appreciation for humor that involves moral violations (β = -.27, p = .003, R² = .07), but it was not associated with reduced humor appreciation for humor that involves no moral violations (β = -.05, p = .63, R² = .006). We also tested this interaction effect more directly using a single regression equation. Specifically, we tested the interaction by regressing the difference in humor appreciation for the two types of jokes on moral identity (Judd, McClelland, & Ryan, 2017). The coefficient for moral identity in this equation thus represents the interaction, and was significant (β = -.22, p = .02, R² = .04).

STUDY 2
Studies 1a and 1b were useful in establishing an initial link between moral identity and humor, but took a somewhat narrow approach to measuring humor. In Study 2, we broaden our approach by examining whether or not participants with a temporarily active moral identity would generate less humorous content than participants in the control condition. We hypothesized that participants in the moral identity condition would be less likely to generate humorous content because they would be less likely to engage in the benign moral violations often used to elicit humor.

**Participants**

We recruited 300 adults from Prolific for the study. We dropped 17 participants from the sample because they failed to write more than one complete sentence for the manipulation of moral identity (described below). Thus, our final sample size was 283 ($M_{age} = 30.84; 49.1\%$ female; $68.6\%$ White). All participants were compensated with a small cash payment.

**Procedures**

Participants were randomly assigned to either a moral identity activation or control condition. In the moral identity condition, we used a different method to manipulate moral identity activation compared to Study 1a. Specifically, participants were provided with a list of morally-laden trait words from Aquino and Reed’s (2002) moral identity scale (e.g., honest, fair, kind), whereas participants in the control condition were provided with morally irrelevant words (e.g., elephant). We then instructed participants to write a story using each of the words at least once, with the goal of activating versus not activating a moral identity, depending on the experimental condition. Prior research has demonstrated that this manipulation can successfully activate a moral identity (Aquino et al., 2011; Aquino et al., 2009). As a manipulation check, at the end of the study, participants were asked to what extent the stories they wrote involved
morality (1 = not at all to 5 = very much). Participants in the moral identity activation condition indicated that their stories involved significantly more moral components ($M = 3.09$, $SD = 1.35$) than participants in the control condition ($M = 1.76$, $SD = 1.02$), $t(281) = 9.35$, $p < .001$.

After the manipulation, participants proceeded to an ostensibly unrelated task in which they were presented with two different photos and asked to create the funniest captions they could (see online supplement). Generating humorous captions has been validated as a measure of humor production in previous studies (Feingold & Mazzella, 1991; Greengross & Miller, 2011; Mickes, Walker, Parris, Mankoff, & Christenfeld, 2012), which captures a relatively spontaneous style of humor production in the absence of human interactions.

Because this procedure produced a large number of written captions to be coded (283 subjects x 2 captions each = 566 captions), we used a crowdsourcing approach (Benoit, Conway, Lauderdale, Laver, & Mikhaylov, 2016; Eitan et al., in press) and recruited a total of 286 coders through Prolific. These independent raters, who were blind to the experimental conditions and research hypotheses, coded the level of humor for each written caption (1 = not funny at all to 5 = very funny). We asked each coder to rate the level of humor for approximately 30 captions. We built in a screening syntax such that coders could not code for more than 30 captions. On average, each caption was rated by 13.29 unique coders. Alpha coefficients for the ratings of the two captions were .85 and .91, respectively. Following a similar procedure, we also asked 286 additional independent coders to code whether or not the captions participants created contained a moral violation (1 = no moral violations to 5 = contain a great deal of moral violations). Alpha coefficients for the two captions were .85 and .90, respectively. An average score (across the two captions) was used for both ratings of humor and moral violation.

**Results and Discussion**
Participants in the moral identity activation condition displayed significantly less humor in the captions they generated ($M = 1.82, SD = .34$) than participants in the control condition ($M = 1.98, SD = .37$), $t(281) = -3.79, p < .001, d = -.45$. In other words, an active moral identity was negatively related to humor production. We also found that participants in the moral identity activation condition were less likely to create captions that involved moral violations ($M = 1.23, SD = .23$) than participants in the control condition ($M = 1.29, SD = .23$), $t(281) = -2.07, p = .04, d = -.26$. We followed the recommendations of Preacher and Hayes (2008; see also Shrout & Bolger, 2002) to examine the indirect effect through a bootstrapping procedure (for a summary of the advantages of using this procedure to test for indirect effects, see Hayes, 2009). Both the experimental condition ($\beta = -.21, p < .001$) and ratings of moral violation were significantly associated with humor ($\beta = .12, p = .04$). A bias corrected bootstrapping analysis (Hayes, 2013, Preacher & Hayes, 2008) revealed that the effect of an active moral identity on humor was mediated by a reduced likelihood of writing captions that involved moral violations ($coefficient = -.01, SE = .01, 95\% CI = -.03$ to $-.001$). Thus, results suggest that individuals with active moral identity are less likely to produce humorous content because they are less likely to commit moral violations in the process.\(^2\)

**STUDY 3**

Thus far, our theory and data (Studies 1a, 1b, and 2) suggest that people with activated moral identities are less likely to be humorous because they have fewer tools to be funny (i.e., appreciate or tell jokes that involve moral violation). However, it is possible that these people

\(^2\) We conceptually replicated the findings of Study 2 in two earlier studies, with smaller sample sizes, using both a student and an Amazon Mechanical Turk sample. These studies were replaced per reviewer request, as they used a less rigorous coding technique with only two trained undergraduate assistants, resulting in low inter-rater reliability. The results from these two earlier studies provided convergent support for our hypotheses, but are omitted for brevity in favor of the higher-powered Study 2.
will compensate for this effect by telling more jokes that do not involve any moral violations. We conducted Study 3 to test this idea.

Participants

We recruited 150 adults from Prolific for the study. \((M_{\text{age}} = 35.37; 60.7\% \text{ male}; 78.0\% \text{ White})\). All participants were compensated with a small cash payment for their time.

Procedure

Participants first completed the internalization dimension of the moral identity survey (Aquino & Reed, 2002; \(\alpha = .72\)) as in Study 1b. We then asked them to imagine themselves in their place of employment, interacting with their coworkers in a common social area. We asked them to imagine that they are chatting with these coworkers. We then presented them with the same set of nine jokes we used in Studies 1a/1b, five of which entail a moral violation and four of which do not. We did not include the nine captions used in Studies 1a/1b because captions cannot be told in a face-to-face social interaction. For each joke, we asked the participants to indicate the degree to which they would be willing to tell the joke in their workplace (1 = definitively not willing to 7 = definitively willing).

Results

We found that those high in moral identity internalization reported a lower level of willingness to tell jokes that involve moral violations \((r = .25, p = .002)\). Interestingly, we did not find a compensatory effect – individuals high in moral identity did not report a higher level of willingness to tell jokes that contain no moral violations \((r = .03, p = .71)\). In other words, it appears that the reduced humor associated with a strong moral identity is due to fewer jokes available to these individuals (both to tell and to laugh at), and that those with a strong moral
identity do not make up for this reduced humor by attempting to tell more jokes that do not involve moral violations.

**STUDY 4**

Studies 1-3 provided convergent support for our hypotheses that individuals with temporarily active or chronically accessible moral identity are less humorous, because they are less likely to appreciate humor that involves benign moral violations, or engage in the benign moral violations often used to generate humor. In our subsequent two field studies (Studies 4 and 5), we examined the downstream implication of the negative relationship between a moral identity and humor. Specifically, we hypothesized that employees with a chronically accessible moral identity would be perceived as less humorous by their coworkers, and to the extent they are seen as less humorous their likeability would be negatively affected.

**Participants and Procedures**

Participants were 70 dyads (54.3% male [two did not report], 55.7% between the ages of 26 to 35; two dyads were dropped for failing the attention check question) from multiple organizations located in central China. Participants worked in a variety of industries, including construction, sales, telecommuting, and human resource management. We contacted Executive Master of Business Administration students from a large public university and asked these managers to randomly recruit two subordinates who worked in the same department or work team that have daily interactions. To avoid selection biases, we asked the managers to choose the two employees based on the first alphabet of their last names. We then randomly assigned one employee as the focal employee and the other one as the coworker. The focal employees completed a measure of moral identity (Aquino & Reed, 2002), whereas their coworkers rated the focal participants’ humor and indicated how much they liked him/her.

**Measures**
Because the scales we used were originally developed in English, we followed Brislin’s (1980) back-translation procedure to develop the Chinese versions of the measures (see online supplement for all scales used).

**Moral identity.** To measure the focal employees’ moral identities, we asked them to complete the internalization dimension of the moral identity scale (Aquino & Reed, 2002) as in Studies 1b and 3. This measure assessed the extent to which being “moral” is believed to be particularly self-defining and central to one’s self-concept, and is thus appropriate as assessing the extent to which moral content was chronically accessible to participants (Aquino et al., 2009; 1 = strongly disagree to 5 = strongly agree; α = .90).

**Perceptions of peer’s humor.** We measured perceptions of humor with a seven-item scale adapted from Thorson and Powell (1993). Because the original scale was designed for measuring recognition of oneself as a humorous person, we reworded the items to reflect perceptions of others’ humor for the purpose of this study following past research (Yam, Christian, Wei, Liao, & Nai, 2018). A sample item is “my coworker says things in such a way as to make people laugh” (1 = strongly disagree to 7 = strongly agree; α = .96).

**Interpersonal liking.** We measured interpersonal liking with a four-item scale (Wayne & Ferris, 1990). A sample item is “I get along well with my coworker” (1 = strongly disagree to 5 = strongly agree; α = .92).

**Results and Discussion**

Descriptive statistics for the key variables are presented in Table 1.

We first conducted an ordinary least squares (OLS) regression. Focal employees’ moral identity was negatively associated with coworkers’ humor ratings for that person (β = -.27, p = .02, R² = .06). This result provides further support for the negative relationship between a
moral identity and a sense of humor: peers perceived colleagues with strong moral identities to be less humorous.\(^3\) We then used a bootstrapping procedure to test for the indirect effect (Hayes, 2013). Results indicated that the indirect effect of moral identity on interpersonal liking via perceptions of peer’s humor was significant (\textit{coefficient} = -.07, SE = .04, 95\% CI = -.19 to -.01; see Figure 2). In sum, these results suggested that individuals with chronically accessible moral identities were perceived as less humorous by their coworkers, and to the extent they are seen as less humorous their likeability is negatively affected.\(^4\)

\begin{center}
Insert Table 1 and Figure 2 Here
\end{center}

However, unexpectedly, we did not find a direct negative effect of moral identity on interpersonal liking (\textit{direct effect} = -.04, SE = .10, 95\% CI = -.23 to .16, \(p = .71\); Figure 2). Preacher and Hayes (2008) noted that a possible explanation for significant indirect effects that are not accompanied by significant direct effects is that there are actually opposing indirect effects that offset each other in the direct effect. Thus, in Study 5 we considered whether or not there could be such an opposing indirect effect that would eliminate the direct effect in the context of our model. Given that moral identity is positively associated with honesty (Hertz & Krettenauer, 2016) and can serve as a positive signal to others regarding one’s own morality (Aquino & Reed, 2002), in Study 5 we examined trust as a potential opposing indirect effect.

\(^3\) Our results in Studies 4 and 5 remain significant in the presence or absence of demographic control variables (i.e., age, gender, race, and dyadic work tenure). We presented the results without any control variables for purpose of presentational parsimony. The statistical analyses with all the control variables can be found in the online supplement.

\(^4\) We also collected data on individual differences in moral identity symbolization in Study 4. Moral identity symbolization was unrelated to humor ratings (\(\beta = -.01, p = .94, R^2 = -.02\)). This is consistent with our theorizing, as moral identity symbolization reflects the desire to demonstrate moral traits to others, whereas internalization reflects the individual’s aspirational self-beliefs (it is personally important for me to be this way; Aquino & Reed, 2002). As we are interested in the individual’s own personal responses to what they view as humorous, moral identity internalization, rather than symbolization, should be the key predictor of humor related outcomes.
Notably, research on ethical leadership suggests that leaders who are moral often experience positive interpersonal outcomes. For example, studies have generally shown that ethical leaders enjoy better social relationships with their subordinates compared to unethical leaders (Gu, Tang, & Jiang, 2015). We suggest that while individuals’ moral identities may lead them to be less liked via reduced humor, there should also be a positive pathway which leads moral individuals to be more liked by others. In the workplace, we suggest that moral employees or leaders are more likely to be trusted than their counterparts, and to the extent this is the case are more liked by others. Meta-analytic studies further indicate that when an employee or leader is perceived as trustworthy, it leads to more favorable evaluations of that person (Dirks & Ferrin, 2001; Ng & Feldman, 2015).

Overall, this provides a richer model in which moral identity has opposing negative and positive indirect effects on likeability through separate mediators of humorlessness and trustworthiness. Whereas others have found that highly moral (and likely, trustworthy) individuals are at times viewed as unlikeable (Monin et al., 2008), separating the more cognitive drivers of liking (i.e., beliefs about the target’s benevolence; Solomon, 1960) from more affective drivers of liking (i.e., perceptions of humorlessness) may reconcile seemingly incompatible findings in the literature. We test these hypotheses in Study 5.

STUDY 5

Study 5 was different from Study 4 in five key respects. First, although we found support for our hypotheses in a Chinese field setting, we wanted to extend the generalizability of these findings to a Western work setting. We thus recruited a sample of employees in the United States. Second, we examined employees’ perceptions of their leaders in Study 5. This allowed us to assess whether or not our findings are generalizable to people with different levels of status. Third, we increased temporal precedence by employing a three-wave design in order to minimize
common-method bias concerns (i.e., a limitation in Study 4 was that the perceptions of humor and interpersonal liking was measured by the same source at the same time point; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In other words, the independent variable, mediators, and dependent variables were all temporally separated in Study 5. Fourth, we tested the positive pathway of ethical leadership (via trust) in parallel with its negative pathway (via humor). Fifth and finally, because prior research has suggested that upward social comparison is the major reason why individuals tend to dislike moral individuals (e.g., Monin et al., 2008; Parks & Stone, 2010), we controlled for this factor in order to demonstrate the unique role of reduced humor on the likeability of individuals with strong moral identities.

**Participants and Procedures**

We contacted 700 full-time employees to participate in this research through Qualtrics, a third-party online survey administration company in the United States (for recent examples of data collection using Qualtrics, see Owens, Yam, Bednar, Mao, & Hart, in press; DeCelles, DeRue, Margolis, & Ceranic, 2012). In order to participate in this study, participants had to have daily interaction with their leaders and be willing to complete three consecutive surveys in one month. A total of 200 participants ($M_{age} = 43.26$, $M_{tenure \ with \ leader} = 5.93$ years, 40% male; 72.5% Caucasian, 10% Asian American, 8.5% Hispanic American; no participants failed the attention check item) completed all three waves of surveys, yielding a response rate of 28.57%.

Each of the three waves were separated by roughly 10 days. At Time 1, participants completed a measure of ethical leadership. At Time 2, participants completed measures of leader humor, trust, and upward social comparison. At Time 3, participants completed a measure of leader liking. Each participant was compensated with $5 for their time per survey and received a $10 bonus if they completed all three surveys.
Measures

**Ethical leadership.** At T1, participants were first asked to rate their leaders’ ethical leadership using a well-established scale developed by Brown, Trevino, and Harrison (2005). A sample item is “My leader makes fair and balanced decisions” (1 = strong disagree to 5 = strongly agree; $\alpha = .94$). We chose this scale to triangulate our previous studies using moral identity as the predictor because a leader’s internalized moral identity might not be visible to followers (e.g., a leader might care strongly about morality internally). However, their ethical leadership behaviors are often visible. Moreover, previous research has found that leaders’ moral identity serves an antecedent to the display of ethical leadership (Mayer et al., 2012). In other words, whereas moral identity can be conceptualized as one’s inner morality, ethical leadership can be conceptualized as the behavioral manifestation of a leader’s moral identity.

**Perceptions of leader’s humor.** At T2, we measured perceptions of leader’s humor by asking participants to complete the same scale as in Study 4 (Thorson & Powell, 1993; $\alpha = .98$).

**Trust.** At T2, we measured trust using McAllister (1995)’s five-item affect-based trust scale. A sample item is “If I shared my problems with my supervisor, I know that he/she would respond constructively and caringly” (1 = strongly disagree to 7 = strongly agree; $\alpha = .95$).

**Leader liking.** At T3, we measured leader liking by asking participants to complete the same scale as in Study 4 (Wayne & Ferris, 1990; $\alpha = .95$).

**Control variable.** At T2, we measured upward social comparison with the leader. We controlled for this factor because to date it is the most commonly used reason to explain the exclusion of moral individuals in social settings (e.g., Parks & Stone, 2010). We adapted a three-item scale from Parks and Stone (2010) in measuring upward social comparison with the leader.
A sample item is “People often ask why I can’t be more like my leader” (1 = strongly disagree to 7 = strongly agree; α = .80).

Results and Discussion

Descriptive statistics for the variables are presented in Table 2.

We conducted OLS regression and first entered ethical leadership and upward social comparison at the same time predicting perceptions of leader humor at T2. Results suggested that perceived leader moral identity, manifested in ethical leadership behavior, is negatively associated with perceptions of leader humor ($\beta = -.23$, $p < .001$, $R^2 = .05$). This result again supports the negative relationship between a moral identity and a sense of humor: followers perceived leaders with strong moral conviction to be less humorous.

We then used the same bootstrapping procedure to test a dual-mediator indirect effect model (Hayes, 2013). Ethical leadership was modeled as an independent variable. Perceptions of leader humor and trust were simultaneously modelled as mediators. Interpersonal liking was modelled as the dependent variable. Upward social comparison was modelled as a control. Results suggested that both indirect effects were significant, but in opposite directions. Again, there was a negative indirect effect of perceptions of leader humor ($\text{coefficient} = -.10$, SE = .05, 95% CI = -.23 to -.02). But at the same time, there was also a positive indirect effect of trust ($\text{coefficient} = .09$, SE = .06, 95% CI = .01 to .24, see Figure 3).

In sum, the results of our two field studies suggest that, regardless of status or the two cultural contexts examined (China and United States), employees with chronically accessible moral identities are less likely to be perceived as humorous, and in turn their likeability is
negatively affected. In addition, Study 5 finds that enhanced trust can be a potential remedy for this effect, by boosting a moral leader’s likeability. Although leaders with strong moral identities are seen as lacking in humor (which negatively affects their likeability), they are at the same time seen as more trustworthy (which positively affects their liability).

**GENERAL DISCUSSION**

In six studies, we found support for our hypotheses that 1) a strong moral identity and a sense of humor are causally and negatively related, as the activation of moral identity leads to reduced humor appreciation and production; 2) individuals with an active moral identity experience a negative indirect effect on their likeability through the mediator of humor, and 3) the same individuals experience a positive indirect effect on their likeability through the mediator of trustworthiness. Overall, this suggests an extension to moral identity theory such that there are opposing indirect effects of moral identity on employee likeability.

In Studies 1a and 2, participants for whom a moral identity was situationally activated were less likely to appreciate and produce humor that involved moral violations. Study 1b conceptually replicates these effects with an individual differences measure of chronic moral identity. Study 3 ruled out the possibility that individuals with strong moral identities compensate for their lack of humor by telling more jokes that do not involve moral violations. These controlled studies converge on the finding that a moral identity (whether temporarily or chronically accessible) is associated with a diminished level of overall humor, but only because it specifically diminishes appreciation and generation of humor that involves breaching moral norms. Individuals with a strong, internalized sense of morality are not any less (or more) likely to engage in humor that stays within the boundaries of common moral norms.
Studies 4 and 5 furthermore extended this investigation to field settings, demonstrating that employees with strong moral identities were perceived to lack a sense of humor, and to the extent they are seen as less humorous their likeability is negatively affected. Finally, Study 5 demonstrated that this effect occurred after controlling for upward social comparisons, and that trust can be an offsetting mediating mechanism that enhances the likability of leaders with strong moral identities. Below, we elaborate on the theoretical contributions and practical implications of our work, highlight limitations of the present studies, and suggest future research directions.

**Theoretical Implications**

Although the notion of a negative relationship between morality and humor can be dated back to ancient philosophy, to our knowledge, the present research is the first to empirically document causal and correlational effects of a moral identity on a person’s sense of humor. Drawing on Benign Violation Theory (McGraw & Warren, 2010), we hypothesized and found that both a temporarily and chronically accessible moral identity lead to reduced appreciation and production of humor that transgresses moral norms. While there is no doubt that moral individuals benefit their organizations and society at large (e.g., financial scandals cost billions), less is known about whether morality may also have negative consequences for the moral individual. The present research suggests that individuals with strong moral identities may be less humorous because they are less likely to engage in the benign moral violations often involved in the display of humor.

Our results also contribute to the broader field of behavioral ethics. To date, scholars have remained largely optimistic about the consequences of being paragons of morality. Indeed, only recently have scholars begun to examine the potential downside of morality, demonstrating for instance that moral individuals can be less creative (Gino & Wiltermuth, 2014). Our results
add a unique insight into this discussion by utilizing BVT to identify another important way in which a moral self-concept can cause unexpected negative consequences. Our findings also complement behavioral ethics research that traditionally examines the antecedents of morality (e.g., Bazerman & Gino, 2012; Tenbrunsel & Smith-Crowe, 2008). Whereas a majority of research focuses on factors that predict moral behaviors, we underscore the importance of examining the potentially negative consequences (reduced levels of some forms of humor) of a temporarily (Studies 1a and 2) and chronically (Studies 1b, 3, 4, and 5) accessible moral identity.

Beyond behavioral ethics, we also contribute to research on humor. Despite the importance and prevalence of humor in social life, it has continued to be an understudied topic. Not only does our research directly speak to the literature on humor, but we also examine both the antecedents and consequences of humor. It appears that a morally-relaxed attitude may facilitate the appreciation and expression of humor. Given the benefits of humor (Mesmer-Magnus et al., 2012), future research should continue to explore individual traits or situational factors that may lead to increased levels of humor production and appreciation.

Our work also contributes to research on the relationship between morality and perceptions of moral individuals. Although intuition suggests that highly moral individuals should be well-liked, research has revealed that moral individuals can be disliked by their peers under some circumstances (Minson & Monin, 2012; Monin et al., 2008; Parks & Stone, 2010; Trevino & Victor, 1992). Scholars have typically used upward social comparisons to explain this puzzling finding. This explanation suggests that when comparing ourselves to morally superior individuals, we feel a sense of moral inferiority, which eventually leads to resentment (for a review, see Monin, 2007). Because individuals in general care a great deal about their self-image (Aquino & Reed, 2002), these feelings of moral inferiority and resentment lead to social
rejection of “do-gooders.” We offer a complementary explanation for the relationship between morality and interpersonal liking that adds to work on social comparisons (Study 5): individuals with chronically accessible moral identity are less humorous, and to the extent they are seen as less humorous their likeability is negatively affected.

**Practical Implications**

Our findings call attention to potential risks that individuals and/or with strong moral identities should be aware of. To be clear, we are not suggesting that they should compromise their values or violate moral standards to be funny and gain likeability in the workplace. Rather, we encourage employees to strive to embody the values they possess while simultaneously cultivating social exchanges that are pleasant and enjoyable. Because the activation of a moral identity appears to suppress humor production, individuals with strong moral identities should seek to deliberately offset this suppression during critical interactions. Although they do not appear to make such an effort spontaneously (Study 3), managers with strong moral identities can make a deliberate effort to tell a few inoffensive linguistic puns while discussing business ethics with subordinates. Developing the humor style (Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003; Saroglou & Scariot, 2002) that works for them may allow ethical leaders to maintain positive ratings of humor, and thus enjoy the benefits of both trustworthiness and likeability.

An additional practical implication is that employees who are high in moral identity should seek other means besides humor to bolster their likeability in the workplace. Given the offsetting effects of their moral identity on humor and trustworthiness, it makes sense for those high in moral identity to find other ways to strengthen their relationships with their colleagues. Establishing relationships of trust with colleagues, helping others, and seeking out positive social interactions more generally may be a few effective approaches to compensate for their reduced
humor. Obtaining likeability gains from these additional actions may help employees with a strong moral identity offset social losses due to decrements in some forms of humor. In contrast, we suggest that those who are low in moral identity and seen as funny but untrustworthy will likely find it hard to enhance their likeability, given how trust is painstaking earned. All in all, we suggest that the net positive effects to being a moral person at work still likely outweigh the negative effects incurred, given how one can compensate for a lack of humor with the aforementioned methods (but perhaps not vice versa).

Limitations and Future Directions

Although we replicate the negative relationship between morality and humor both in experimental and field studies using different manipulations and measures (Anderson et al., 1999), a number of limitations warrant acknowledgment. First, the measurement of humor production in Study 2 is somewhat constrained and artificial. For example, humor generation often has a reactive component (e.g., making a funny comment immediately after a co-worker spills his coffee) which was not captured with this behavioral measure. Likewise, although we capture humor production in general in Study 2, we did not assess specific types of humor such as satire or sarcasm. Finally, other ratings of the focal person’s humor might be prone to rater biases (e.g., the rater him/herself lacks a sense of humor). However, we note that our package of studies utilized different measures of humor – appreciation of humor, humor production in the absence of others, and perceptions of others’ humor – to capture the construct space of humor as fully as feasible. Nevertheless, we recommend that future research employ even more measures of humor to replicate and extend our findings.

Second, although we hypothesized and found a negative, linear relationship between morality and humor in this study, there are reasons to expect that the relationship between a
moral identity and humor might in fact be curvilinear (McGraw et al., 2014), such that people find both highly moral and highly immoral others to be humorless. We re-ran our analyses for both Studies 4 and 5 to test for this potential curvilinear relationship. We added a squared term of moral identity and ethical leadership (in tandem with the linear terms) and examined their effects on perceptions of humor. Results indicated that the curvilinear effect of morality identity (Study 4) and ethical leadership (Study 5) on perceptions of humor was not significant ($\beta = 1.20, p = .22; \beta = -.03, p = .95$, respectively). Nevertheless, we wish to note that restriction of range may undermine our ability to detect curvilinear effects between morality and humor, and we urge future research to examine this link with a wider range of samples and measures.

Third and relatedly, we did not observe a significant main effect of moral identity (Study 4) or ethical leadership (Study 5) on interpersonal liking. Yet, prior research on impression formation, as well as lay beliefs, suggest that immoral people are often disliked by others (e.g., Brambilla & Leach, 2014; Wojciszke, Bazomska, & Jaworski, 1998). Given these prior findings and the observed negative indirect effect of morality to reduced likeability (via humor), we tested whether moral identity (Study 4) or ethical leadership (Study 5) has a curvilinear relationship with interpersonal liking, with very low morality leading to a steep drop in likeability. None of the quadratic terms were significant ($\beta = .50, p = .62$ in Study 4; $\beta = -.37, p = .48$ in Study 5). Although it is difficult to interpret null findings, we speculate there is a restriction of range issue embedded in our field studies. Self-ratings of moral identity (Study 4) and other-ratings of ethical leadership (Study 5) are rather high ($Ms = 4.15$ and $4.03$ out of $5.00$ respectively), although such means are consistent with the literature on moral identity (Hertz & Krettenauer, 2016) and ethical leadership (Ng & Feldman, 2015). Although one can argue that self-ratings of moral identity are prone to social desirability bias, other-ratings of ethical leadership should not
be, or at least to a lesser degree. We speculate that one reason for the latter to display such a high mean is that most leaders who rise to management need to maintain some standards of morality. Field studies sample relatively successful professionals who have achieved the support and trust of many of their colleagues; individuals who gained reputations for poor ethics may have left the organization or leadership positions. In addition, we speculate that displays of humor are much more frequent than displays of immorality. Hence, the relative strength of influence on one’s likeability should be more heavily affected by displays of humor (found in this research) compared to displays of immorality (which are rare).

While field and survey studies of morality might find it difficult to avoid these restriction of range issues, we encourage future experimental work to manipulate moral identity from a third-party perspective to test for curvilinear patterns. For instance, target persons in vignettes might be manipulated to be very high, somewhat high, medium, somewhat low, or very low in moral identity and behavior, and their anticipated humor level, trustworthiness, and likability evaluated by participants. Social perceivers may perceive individuals somewhat low in moral identity (e.g., petty thieves) to be the most generative of humor and therefore (after controlling for the suppressing effects of their low trustworthiness) also more likeable and charming. In contrast, targets very low in moral identity (e.g., rapists and murderers) should not be seen as adept at humor, socially liked, or trustworthy. Another approach may be for future field studies to focus on participants from industries or jobs where immorality might be more common at work (e.g., jobs that involve deception such as used car dealerships) to allow more variance and hence better test for curvilinear effects.

Fourth, various individual differences not captured in the current studies might affect how employees react to their coworkers’ or leaders’ morality identity and thus lack of humor.
For example, followers who themselves have a strong moral identity are likely less concerned about reduced humor on the part of their ethical leaders, increasing the importance of the humor path and reducing that of the trust path. In contrast, followers who themselves enjoy humor might react especially negatively towards ethical leaders who lack a sense of humor. Similarly, contextual factors might moderate these relationships – ethical employees and leaders might suffer the most in industries where humor is especially encouraged (e.g., tourism, entertainment, etc.). In some cultures, leaders may use humor to establish interpersonal trust, as in the *simpatico* ethos in some Latin American countries (Sanchez-Burks, Nisbett, & Ybarra, 2000). We encourage future research exploring for whom and in what contexts the effects we report are especially prominent vs. restricted.

Fifth, although we demonstrated via controlled experiments that the mechanism that underlies the negative association between moral identity and humor is moral people’s unwillingness to engage in benign, mild moral violations in order to generate humor (Studies 1a, 1b, and 2), we were unable to collect such fine-grained data in our field studies (Studies 4 and 5). This is due to the challenges in teasing out the types of jokes in the broader social setting. Future research might ask employees to separately rate their coworkers’ tendency to tell “clean” and “dirty” jokes, or employ experience sampling methods (Hofmann, Wisneski, Brandt, & Skitka, 2014) in the field to assess the frequency of everyday workplace humor that does vs. does not involve moral violations.

Sixth, although we found consistent results for reduced humor as the mechanism that explained the negative association between moral identity and likeability and controlled for upward social comparison, we did not rule out other plausible alternative explanations. For example, perhaps those with strong moral identities are also more likely to report unethical
behavior committed by coworkers, which then causes others to dislike them. We thus encourage future work to control for these alternative explanations. Moreover, our work finds opposing, mediating paths from morality to liking via humorlessness and trustworthiness. Future research should explore potential sub-mediators which explain why morality may simultaneously increase and decrease liking for different reasons. Because trust is generally viewed as a rational-cognitive process and affective judgments about trust appear to dissipate rather quickly (Schoorman, Mayer, & Davis, 2007), it is likely that moral internalization increases rationally-driven perceptions of benevolence (Solomon, 1960), while also leading to more affective, visceral responses of disliking (as a function of humorlessness or upward comparisons). This may explain the tendency to sometimes “know that we should like someone” who is highly ethical even when we actually do not.

Seventh, individual and demographic differences in humor perceptions – what is regarded as funny, contrary to social standards, benign, or deeply offensive – should be examined in greater depth. Although speculative, individuals high and low in moral identity may vary in the appreciation, rather than recognition, of humor containing moral violations. Drawing as before on Benign Violations Theory, we suggest that as opposed to perceiving jokes that internalizers view as moral violations as non-violations, individuals low in moral identity internalization may actively enjoy that immorality so long as the transgression is not too extreme (i.e., benign). For the low internalization individual, the fact that a joke is in breach of common moral standards may be precisely what makes it so funny and makes them want to share it with others. Recognition and appreciation of “dirty” humor may vary based on the person’s endorsement of the five Moral Foundations (Graham, Haidt, & Nosek, 2009), Big 5 Personality Traits (Goldberg,
1993), as well as demographic differences such as gender, age, and socioeconomic status, among other potentially relevant dimensions.

Eighth, although our focus here is on the distinction between humor that does vs. does not involve moral violations, future work should more carefully parse different categories of humor and relate them to moral identity internalization, likeability, and perceived trustworthiness. For instance, Martin et al.’s (2003) theory of humor styles distinguishes between affiliative humor, in which the person engages in spontaneous banter to amuse others, self-enhancing humor that adopts a lighthearted outlook on life as a coping mechanism, self-defeating humor in which the person tells jokes at their own expense in order to amuse others, and aggressive humor which involves making jokes at the expense of others. Self-reported individual differences, other-reports from work peers, or coded writings and utterances could be used to capture these humor styles and assess the extent to which they conform to the general pattern shown here. A strong moral identity may suppress aggressive humor more than the others, and aggressive humor should also fail to project likability. At the same time, affiliative and self-enhancing humor may more effectively establish trust than telling jokes at the expense of oneself or colleagues. Leaders seeking to establish both their social likeability and ethical reputation would do well to tailor their humor style to these key social goals.

Finally, a full consideration of the relationship between humor and morality must also account for instances in which jokes are leveraged to raise ethical issues, such as racism or political corruption. Consider late night talk shows with stand-up comedy routines pointing out the hypocrisies and self-contradictions of those in positions of power. In such cases, in which sarcasm or mockery are used to disparage unethical conduct, moral concerns and humor appear synergistic rather than antagonistic. Another example is the activist group Adbusters, which uses
clever parodies of famous advertising campaigns to point out what they see as the moral shortcomings of certain industries. One viral line of parody ads shows the cartoon character Joe Chemo, who bears an uncanny resemblance to former cigarette spokesperson Joe Camel, in a hospital suffering from lung cancer. Such “culture jamming” (Kozinets & Handelman, 2004) is designed to subvert the marketing of harmful products by linking the company’s message to critical counter-arguments via vivid imagery often involving humor. Thus, while a strong moral identity may generally suppress the use of humor, it is also possible to deploy jokes subversively to undermine unethical practices and institutions. Future research should examine in greater depth ways in which humor is weaponized in mass persuasion contexts, to both moral and immoral ends.

Conclusion

As noted in the opening quote by Mark Twain that “there is no humor in heaven,” people have long considered the idea that there is a tension between morality and humor. Both experimental and field data indicate that such tension not only means that morality may impinge upon humor, but that this can occur at the expense of likeability. Thus, morality can have a downside that was previously overlooked. Although our society will be better off if people are moral, there can be social cost in becoming paragons of morality in organizations.
References


tests and a forecasting tournament to address the controversy. *Journal of Experimental Social Psychology, 79,* 188-199.


Tenbrunsel, A. E., & Smith-Crowe, K. (2008). 13 ethical decision making: where We’ve been and where We’re going. *Academy of Management Annals, 2*(1), 545-607.


Table 1. Means, standard deviations, and correlations of the focal variables (Study 4)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means (SD)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Moral identity</td>
<td>4.15 (.88)</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Perceptions of humor</td>
<td>3.71 (1.12)</td>
<td>-.27*</td>
<td>.96</td>
<td></td>
</tr>
<tr>
<td>3. Interpersonal liking</td>
<td>4.09 (.71)</td>
<td>-.13</td>
<td>.34**</td>
<td>.92</td>
</tr>
</tbody>
</table>

Alpha coefficients are presented on the diagonal

*p < .05

**p < .01
Table 2. Means, standard deviations, and correlations of the focal variables (Study 5)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means (SD)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ethical leadership (T1)</td>
<td>4.03 (.81)</td>
<td>.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Perceptions of humor (T2)</td>
<td>4.44 (1.66)</td>
<td>-.21**</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Leader trust (T2)</td>
<td>5.20 (1.44)</td>
<td>.27**</td>
<td>.35**</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Leader interpersonal liking (T3)</td>
<td>5.22 (1.61)</td>
<td>.01</td>
<td>.28**</td>
<td>.26**</td>
<td>.95</td>
<td></td>
</tr>
<tr>
<td>5. Social comparison (T2)</td>
<td>2.97 (1.27)</td>
<td>.16*</td>
<td>.07</td>
<td>.10</td>
<td>-.01</td>
<td>.80</td>
</tr>
</tbody>
</table>

Alpha coefficients are presented on the diagonal

*p < .05

**p < .01
Figure 1. Humor appreciation by condition (Study 1a)
Notes: **p < .01; *p < .05

*The exact p values were .02 for the effect of moral identity on perceptions of humor, and .009 for the effect of perceptions of humor on interpersonal liking. Please note that the indirect effect analysis does not produce an exact p-value and only the 95% CI.*
Figure 3. Indirect effects in Study 5

Indirect effect = -.10 (95% CI = -.24 to -.02)

Notes: ***p < .001; **p < .01; *p < .05
The exact p values were .001 for the effect of ethical leadership on perceptions of humor, .004 for the effect of perceptions of humor on interpersonal liking, and .02 for the effect of perceptions of humor on interpersonal liking. Please note that the indirect effect analysis does not produce an exact p-value and only the 95% CI.