London Business School

LBS Research Online

K O'Connor, D Effron and B J Lucas
Moral cleansing as hypocrisy: When private acts of charity make you feel better than you deserve
Article

This version is available in the LBS Research Online repository: https://lbsresearch.london.edu/id/eprint/1398/

O'Connor, K, Effron, D and Lucas, B J

(2020)

Moral cleansing as hypocrisy: When private acts of charity make you feel better than you deserve.

Journal of Personality and Social Psychology, 119 (3). pp. 540-559. ISSN 0022-3514

DOI: https://doi.org/10.1037/pspa0000195

American Psychological Association
https://psycnet.apa.org/doiLanding?doi=10.1037/psp...

Users may download and/or print one copy of any article(s) in LBS Research Online for purposes of research and/or private study. Further distribution of the material, or use for any commercial gain, is not permitted.

Running Head: MORAL CLEANSING AS HYPOCRISY

Moral Cleansing as Hypocrisy:

When Private Acts of Charity Make You Feel Better Than You Deserve

Kieran O'Connor
University of Virginia

Daniel A. Effron
London Business School

Brian J. Lucas
Cornell University

Corresponding Author: Kieran O'Connor, McIntire School of Commerce, 125 Ruppel Drive, Charlottesville, VA 22903 University of Virginia, Charlottesville, VA 22904. Phone: 434-924-0887, Email: oconnor@virginia.edu.

Abstract

What counts as hypocrisy? Current theorizing emphasizes that people see hypocrisy when an individual sends them "false signals" about his or her morality (Jordan, Sommers, Bloom, & Rand, 2017); indeed, the canonical hypocrite acts more virtuously in public than in private. An alternative theory posits that people see hypocrisy when an individual enjoys "undeserved moral benefits," such as feeling more virtuous than his or her behavior merits, even when the individual has not sent false signals to others (Effron, O'Connor, Leroy, & Lucas, 2018). This theory predicts that acting *less* virtuously in public than in private can seem hypocritical by indicating that individuals have used good deeds to feel less guilty about their public sins than they should. Seven experiments (*N* = 3,468 representing 64 nationalities) supported this prediction. Participants read about a worker in a "sin industry" who secretly performed good deeds. When the individual's public work (e.g., selling tobacco) was inconsistent with, versus unrelated to, the good deeds (e.g., anonymous donations to an anti-smoking cause vs. an anti-obesity cause), participants perceived him as more hypocritical, which in turn predicted less praise for his good deeds. Participants also inferred that the individual was using the inconsistent good deeds to cleanse his conscience for his public work, and such moral cleansing appeared hypocritical when it successfully alleviated his guilt. These results broaden and deepen understanding about how lay people conceptualize hypocrisy. Hypocrisy does not require appearing more virtuous than you are; it suffices to feel more virtuous than you deserve. (250 words)

KEYWORDS: Hypocrisy, moral judgment, behavioral integrity, guilt, moral cleansing

Moral Cleansing as Hypocrisy:

When Private Acts of Charity Make You Feel Better Than You Deserve

In 2015, hackers released the names of people with paid accounts at Ashley Madison, a website designed to facilitate extramarital affairs. Among the names was Josh Duggar, a family-values activist and the executive director of the Family Research Council, a lobbying group that aims "to champion marriage and family as the foundation of civilization, the seedbed of virtue, and the wellspring of society" (Feinberg, 2015). Given that Duggar appeared to be cultivating a public image that was more virtuous than his private behavior, it is no surprise that many accused him of hypocrisy (Puente, 2015). Now consider the founding CEO of Ashely Madison, Noel Biderman. Though Biderman publicly promotes marital infidelity, he has admitted to practicing marital fidelity in his private life (Walker, 2015). Would people condemn Biderman as a hypocrite if they knew he was faithful to his wife? Perhaps not, because his failure to practice what he preaches does not involve cultivating the false appearance of virtue. However, we will argue that people still see hypocrisy in Biderman's behavior. By investigating how publicly promoting vice while privately practicing virtue affects lay people's judgments of hypocrisy—and the consequences of these judgments for how much they praise the virtue—the present research advances theory about "what counts" as hypocrisy.

Understanding how people think about hypocrisy is important because individuals and organizations suffer severe consequences when they are perceived as hypocritical. Individuals incur greater moral condemnation, punitive sentiment, and other forms of censure for the same wrongdoing when audiences perceive it as hypocritical versus merely wrong (Effron, Markus, Jackman, Muramoto, & Muluk, 2018; Kreps, Laurin, & Merritt, 2017; Laurent, Clark, Walker, & Wiseman, 2014; Powell & Smith, 2013). Scandals involving hypocrisy will damage a

politician's reputation more severely than equivalent scandals without hypocrisy (Bhatti, Hansen, & Olsen, 2013). In a laboratory study, job applicants who had committed a minor transgression were less likely to be selected for a job and more likely to be offered a lower wage when the transgression seemed hypocritical than when it did not (Effron, Lucas, & O'Connor, 2015). When leaders and teams in organizations fail to enact the values they preach, employee performance can suffer and employment turnover can increase (Greenbaum, Mawritz, & Piccolo, 2015; Palanski & Yammarino, 2011; Simons, Friedman, Liu, & McLean Parks, 2007). When organizations themselves seem hypocritical, their reputations and stock prices can take a hit (Janney & Gove, 2011; Stoll, 2002; Wagner, Lutz, & Weitz, 2009).

Examples like these illustrate that hypocrisy often involves inconsistency between a person's public image and private behavior. Several theories suggest that inconsistency only counts as hypocrisy when the person appears *more virtuous* in public than in private (Graham, Meindl, Koleva, Iyer, & Johnson, 2015; Jordan, Sommers, Bloom, & Rand, 2017; Monin & Merritt, 2012; Stone & Fernandez, 2008). In this view, people hate hypocrites for trying to trick them with a "false signal" of virtue (Jordan et al., 2017, p. 356). However, by focusing on inconsistencies between public virtues and private transgressions, the literature's dominant perspectives on hypocrisy offer an incomplete and overly narrow understanding of how laypeople think about hypocrisy. Challenging these theories, we propose that a public image that is *less* virtuous than private behavior also counts as hypocrisy in laypeople's minds. We argue that what bothers people in such cases is not that a person appears more virtuous than she is, but that she feels more virtuous than she deserves.

To illustrate how inconsistency between a public image and private behavior could seem hypocritical, consider two tobacco executives who privately and anonymously donate to different

causes: The first executive funds the fight against smoking and the second funds the fight against obesity. We posit that the first executive is more likely to be judged as a hypocrite, and will thus receive less praise for her donation, than the second executive. Our reasoning is that, compared to the anti-obesity donation, the anti-smoking donation is more easily seen as an attempt to alleviate guilt for promoting tobacco use – a case of “moral cleansing” (Tetlock, Kristel, Elson, Green, & Lerner, 2000). As we develop below, laypeople may consider such moral cleansing as hypocritical because they perceive it to allow an individual to feel more virtuous than she deserves.

The following sections review theories of hypocrisy and their limitations, develop our hypotheses, and present seven experiments and a meta-analysis of their results.

Theories of How Laypeople Judge Hypocrisy

Hypocrisy and inconsistency. Current theories offer limited consensus on “what counts” as hypocrisy to laypeople. One perspective, grounded in cognitive dissonance theory (Aronson, 1968; Festinger, 1957), is that people’s aversion to hypocrisy simply reflects their general dislike of inconsistency (Tedeschi, Schlenker, & Bonoma, 1971). A similar perspective has been influential in the organizational-behavior literature (see Simons et al., 2015, for a review), where a foundational paper conceptualized hypocrisy as “inconsistency between talk and action” (Simons, 2002, p. 20; see also Greenbaum et al., 2015). In this view, hypocrisy “does *not* consider the morality of principles, but rather focuses on the extent to which stated principles are seen as aligning with actions” (Simons, 2002, p. 19).

More recent research, however, has revealed that a simple aversion to inconsistency cannot adequately account for lay judgments of hypocrisy (Alicke, Gordon, & Rose, 2013; Effron, O’Connor, Leroy, & Lucas, 2018; Hale & Pillow, 2015). For example, “saying one thing

then doing another" seems more hypocritical than "doing one thing than saying another" (Barden, Rucker, & Petty, 2005; Barden, Rucker, Petty, & Rios, 2014). That is, the same degree of inconsistency between practicing and preaching seems less hypocritical when the practicing precedes the preaching, because people attribute the inconsistency to "turning over a new leaf." Relatedly, advising others to "do as I say, not as I've done" seems less hypocritical when the advisor has suffered for what she has done, because suffering made the advisor's preaching seem more genuine (Effron & Miller, 2015). Even normative perspectives do not label all inconsistencies as hypocrisy. For example, some studies operationalized hypocrisy as holding the self to lower moral standards than others (Lammers, 2012; Lammers, Stapel, & Galinsky, 2010; Valdesolo & DeSteno, 2007, 2008), but labeled behavior as *hyper*-critical when holding the self to *higher* moral standards than others (Lammers et al., 2010). Going further, in some cases inconsistency between words and deeds can even have positive interpersonal consequences for the actor (Effron & Monin, 2010), such as favorable impression management (Lönnqvist, Irlenbusch, & Walkowitz, 2014).

Based on these and other findings, a recent review concluded that to condemn actions as hypocritical, observers not only need to detect inconsistency, but also must make a morally discrediting attribution about why the inconsistency occurred (Effron, O'Connor, et al., 2018). In other words, the evidence indicates that hypocrisy is an *interpretation* of inconsistent behaviors' meaning, not merely a perception that behaviors are inconsistent. However, current theories disagree about what kind of discrediting attributions hypocrisy involves. We review these theories next and highlight a key tension in the literature that we aim to address with the present research.

Hypocrisy as false signaling. One theory holds that people who practice the same behavior they condemn in others are hypocritical "because they use their condemnation to mislead other people about their moral behavior" (Jordan et al., 2017, p. 387). In other words, preaching virtue sends others a "false signal" that one will act virtuously oneself (Jordan et al., 2017). In support of this perspective, online participants rated a target person more negatively for downloading illegal music if she had previously condemned illegal downloading than if she had not – unless she also admitted to this behavior at the time she condemned it. By admitting to the less-virtuous behavior she condemned, the target clarified that she did not intend her condemnation to signal her virtue. Directly condemning a behavior is not the only way of falsely signaling one's virtue. Working for an organization that promotes a virtuous cause (e.g., safe driving), and then acting inconsistently with that cause (e.g., driving recklessly) seems hypocritical because people interpret the work as a false signal that one values the cause (Effron et al., 2015).

The false-signaling theory of hypocrisy comports with a normative argument made by other scholars: that inconsistency *should* only be considered hypocrisy if it involves the "false appearance of virtue" (Monin & Merritt, 2012, p. 171). Researchers' decisions about how to operationalize hypocritical behavior reflect this sort of normative definition (Graham et al., 2015). For example, participants in one series of studies chose to flip a coin as a fair way of determining who would complete a desirable task (versus a tedious task), but then lied about the coin's outcome to ensure they would get to do this task (Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Batson, Thompson, & Chen, 2002; Batson, Thompson, Seuferling, Whitney, & Strongman, 1999). Batson and colleagues labelled this behavior hypocritical because

it suggests participants wanted to appear moral without being moral (Batson et al., 1999; see also Batson, 2002).

The idea that hypocrisy requires falsely signaling virtue fits well with the etymology of the term itself, from the Greek *hypokrisia*, meaning “to play a part on the stage” (Alicke et al., 2013; Crisp & Cowton, 1994). Akin to Goffman’s (1959) metaphor of the self as a theater, hypocrites seem to be those who present themselves publicly at the “front stage” in ways that appear more virtuous than they are privately at the “back stage.” In other words, this perspective paints hypocrisy as a “moral masquerade” (Batson et al., 1997, p. 1336) occurring “through an act of moral self-presentation” to others (Stone, Wiegand, Cooper, & Aronson, 1997, p. 54).

To summarize, the false-signaling theory posits that inconsistency becomes hypocritical when people’s *public* behavior disingenuously signals their *virtue* to an audience of *other people*. This perspective fits with the term’s etymology, with normative arguments about what *should* count as hypocrisy, and with existing data on what lay people think *does* count as hypocrisy. However, the false-signaling perspective does not share our prediction that people see hypocrisy in the inconsistency between *public vice* and *private virtue*, because such inconsistency does not signal a virtuous image to others. In the next section, we review another theory that does share our prediction and offers a broader perspective on hypocrisy than the false-signaling perspective.

Hypocrisy as claiming undeserved moral benefits. An alternative theoretical perspective is that people will interpret an actor’s inconsistency as hypocrisy if they believe the actor has “claimed an undeserved moral benefit,” defined as a social or psychological reward for virtuous behavior (Effron, O'Connor et al., 2018, p. 65). Because people are motivated to uphold norms of equity (Walster, Walster, & Berscheid, 1978), to believe in a just world (Lerner, 1980; Lerner & Miller, 1978; Ross & Miller, 2002), and to punish free-riders (Fehr & Gächter, 2002;

Jordan, Hoffman, Bloom, & Rand, 2016), it makes sense that they would morally condemn someone for taking a benefit to which he or she is not entitled.

Appearing more virtuous to others than one really is represents one example of an undeserved moral benefit. Appearing virtuous can satisfy people's psychological need to be respected, trusted, and positively regarded, and can also lead to tangible rewards (Axelrod & Hamilton, 1981; Frank, 1988). In this respect, the undeserved-benefits perspective encompasses the false-signaling perspective: Falsely signaling virtue is one way of claiming an undeserved moral benefit.

At the same time, the undeserved-benefits perspective is broader than the false-signaling perspective because appearing virtuous is not the only moral benefit that a person could claim. Other moral benefits include the right to judge or influence others' moral behavior (Effron & Miller, 2015), and – most germane to the present investigation – *feeling* virtuous (Effron, O'Connor, et al., 2018). Because morality is a core part of many people's self-concept (Aquino & Reed, 2002; Steele, 1988), people are willing to make substantial sacrifices to feel virtuous (Andreoni, 1988, 1990; Batson et al., 1989, 1988). For example, they will make anonymous donations to feel generous (Hugh-Jones & Reinstein, 2012; Karlan & List, 2007), or forgo lucrative cheating opportunities to avoid feeling like a cheater (Abeler, Nosenzo, & Raymond, 2019; Bryan, Adams, & Monin, 2013; Mazar, Amir, & Ariely, 2008; Shalvi, Dana, Handgraaf, & De Dreu, 2011). In short, feeling virtuous is a valuable resource that normally must be earned through costly or effortful behavior.

Just as appearing more virtuous to others than one deserves is perceived as hypocritical, we propose that *feeling* more virtuous than one deserves is also perceived as hypocritical. Past research suggests that observers may readily perceive that others' feelings of virtue are unearned.

For instance, cynicism about the motives and self-interest of others (Kruger & Gilovich, 1999; Newman & Cain, 2014; Ratner & Miller, 2001) may lead observers to suspect that a target is motivated to wring the most virtuous feeling out of the smallest good deed, particularly when that target's transgressions are salient (Gawronski & Bodenhausen, 2006), as is the case in our empirical context. Further, people are less likely to perceive others' behavior, compared to their own behavior, as earning the right to feel virtuous (Effron, 2014; Kruger & Gilovich, 2004). Anticipating that an actor is likely to judge her own good deeds against relatively lax standards, observers may be on the lookout for actors who try to get away with feeling virtuous despite not earning these feelings. Observers would condemn these actors as hypocrites, according to the undeserved-benefits perspective.

One situation in which observers may suspect that an actor is trying to feel undeservedly virtuous is when the good deeds she does in private directly contradict her public vices. In such situations, compared to those in which the good deeds are unrelated to the vices, observers should be more likely to construe the good deeds as an attempt to ease the actor's conscience by compensating for the public vices. Observers may construe these goods deeds as "molehills of virtue" but anticipate that the actor interprets them as "mountains of morality" (Effron, 2014). Observers may thus condemn the actor as a hypocrite because they believe she feels more virtuous than she deserves.

For example, consider again the founder of the Ashley Madison website, who publicly encourages marital infidelity but privately stays faithful to his wife. Audiences could impute that he believes his own fidelity compensates for the infidelity he has facilitated for his website's 54 million subscribers, and that he should feel guiltier than he does (May, 2018). Or consider again the tobacco executive who secretly and anonymously donates to fund smoking-cessation

programs. Audiences could infer that her donation alleviates her guilt for selling cigarettes. Although her work harms millions of smokers, with relatively minimal effort she has allowed herself to sleep well at night. More broadly, people may interpret private good deeds as hypocrisy if they seem like an attempt to purchase a clean conscience on the cheap – in other words, to claim a moral benefit that one does not deserve.

Prior research has documented numerous examples of moral behavior motivated by a desire to reduce guilt (Carlsmith & Gross, 1969; Tangney, Stuewig, & Mashek, 2007; Xu, Bègue, & Shankland, 2011). For example, recalling or imagining bad deeds may motivate people to bolster their self-image by making more-virtuous choices (Jordan, Mullen, & Murnighan, 2011; Sachdeva, Iliev, & Medin, 2009; Zhong et al., 2010). Such *moral cleansing* – which occurs when "people counterbalance their bad deeds by behaving in ways that symbolically reaffirm the values that had been undermined" (Gollwitzer & Melzer, 2012, p. 1356) – may help individuals feel more virtuous following ethical lapses (Jordan et al., 2011; Lee & Schwarz, 2011; Zhong & Liljenquist, 2006; Zhong, Liljenquist, & Cain, 2009). Advancing this prior research, we propose that moral cleansing can also make a person vulnerable to accusations of hypocrisy.

Hypotheses

To summarize, both the false-signaling and the undeserved-benefit theories can account for prior work on lay judgments of hypocrisy. In these situations, because of its inconsistency with private vices, public behavior sends a false signal of virtue to an audience of other people. The signal represents a claim to a moral benefit (the appearance of virtue), and the signal's falsity makes the benefit undeserved. To distinguish between these two theories, we examine how people judge actors whose public behavior (e.g., promoting tobacco use) is inconsistent with

private virtues (e.g., secret, anonymous donations to anti-tobacco causes). Here, the public behavior does not signal virtue to an audience of other people (if anything, it signals the opposite). However, the public behavior could imply that the actors are claiming an undeserved moral benefit: feeling more virtuous than their behavior merits. Thus, if people condemn inconsistency as hypocrisy when private virtue follows public vice, it would support the undeserved-benefits theory and suggest that the false-signaling theory cannot account for the full range of lay beliefs about hypocrisy.

The main goal of our research was to test whether and why people would view inconsistency as hypocritical even when it involved public vice and private virtue. Following the undeserved-benefits theory, we hypothesized that inconsistency between private virtues and public behavior increases ascriptions of hypocrisy. The undeserved-benefits theory also specifies the mechanism explaining why people would view inconsistency between public vices and private virtues as hypocritical: They view the private virtues as moral cleansing that has the potential to make the hypocrite feel less guilty about the vices than he or she deserves.

A second goal was to refine the undeserved-benefits theory by exploring whether hypocrisy, in lay people's minds, requires *obtaining* an undeserved moral benefit, or merely *attempting to obtain* such a benefit. In other words, we explore the roles of intentions (Knobe, 2003; Malle & Knobe, 1997) versus outcomes (Alicke, 2008; Cushman, 2008; Mazzocco, Alicke, & Davis, 2004) of guilt-reduction in hypocrisy judgments. The undeserved-benefits theory is clear that *actually* feeling more virtuous than you deserve is hypocritical, as in the case of the tobacco executive whose work harms millions of smokers but whose conscience feels clean after donating to an anti-tobacco cause. The theory is not clear about whether merely *trying* to feel more virtuous than you deserve is hypocritical, as in the case of a tobacco executive who

unsuccessfully attempts to alleviate her guilt with an anti-tobacco donation. On one hand, to the extent that attempting to alleviate guilt seems like a claim to an undeserved moral benefit, it would seem hypocritical. On the other hand, guilty feelings could be a positive signal of moral character (Barasch, Levine, Berman, & Small, 2014) and only seem hypocritical if successfully alleviated. Without formalizing hypotheses, we explored whether successful moral cleansing is necessary to evoke hypocrisy or whether attempted moral cleansing is sufficient.

Our studies also measured a potential downstream consequence of hypocrisy judgments: moral praise. As noted, previous research shows that private transgressions receive greater moral condemnation when they are perceived as hypocritical (Effron, Markus, et al., 2018). By analogy, private virtues should receive less moral praise when they are perceived as hypocritical. Several of our studies tested whether hypocrisy ascriptions would mediate a negative indirect effect of inconsistency on praise.

Testing these hypotheses goes beyond normative debates about what *should* count as hypocrisy (Crisp & Cowton, 1994; Graham et al., 2015; Monin & Merritt, 2012; Simons, 2002; Szabados & Soifer, 2004) and contributes a new perspective to research on how people *actually* judge what counts as hypocrisy (Alicke et al., 2013; Effron et al., 2015; Effron & Miller, 2015; Effron, O'Connor, et al., 2018; Jordan et al., 2017; Kreps et al., 2017). By testing whether inconsistency between public vices and private virtues seems hypocritical to laypeople, we aim to demonstrate that people perceive hypocrisy in a broader array of circumstances than previously appreciated, which helps disentangle different theoretical perspectives. The undeserved-benefits and the false-signaling theories agree that appearing more virtuous than you are can trigger hypocrisy ascriptions, but only the undeserved-benefits perspective asserts that false appearances are not necessary; feeling more virtuous than you deserve is sufficient.

The Present Research

We present seven experiments testing our hypotheses. Study 1 – based on the earlier Ashley Madison example – examined whether participants would judge the founder of a website promoting extra-marital affairs as more hypocritical when they knew (vs. were unaware) that he was faithful to his own spouse. Study 2 tested the same prediction in a new scenario with a more tightly-controlled manipulation. That is, the target person in Study 2 always enacted private good deeds, but we manipulated whether the good deeds were inconsistent with vs. unrelated to his public behavior in his work role. To ensure generalizability and robustness, Study 3 recruited participants from more than 60 different nationalities to consider 20 different contexts where hypocrisy might occur. Study 4 provided a first test of the moral-cleansing mechanism in a measurement-of-mediation design. Two supplemental studies reported in the Online Supplement replicated our previous findings and tested robustness with minor modifications to the procedure. Study 5 manipulated the hypothesized moral-cleansing mechanism directly, and also explored whether *successful* moral-cleansing would be necessary for people to perceive hypocrisy in private virtue, or whether *attempted* (but unsuccessful) moral cleansing would suffice. Finally, to more precisely estimate our effects' size and reliability, we meta-analyzed our studies. Several studies included a measure of praise for the virtuous behavior, because lowered praise could be a downstream consequences of hypocrisy judgments.

Open Practices

We report all conditions, measures, and data exclusions, and we have posted verbatim study materials in the Online Supplement. Stopping rules for data collection were determined in

advance, and Studies 3 and 5 were pre-registered. Study data[1] and analysis code can be found at https://osf.io/uce7n/?view_only=10d4262c8b144f8b93981c6860da16a0.

Study 1: Monogamous CEO Encourages Others to Cheat

Method

Participants. Four hundred-one participants signed up for a "Current Events and News Survey" (M_{age} = 37.25 years, *SD* = 12.25; 213 men, 188 women, 2 no response) on Amazon Mechanical Turk (MTurk; Buhrmester, Kwang, & Gosling, 2011; Horton, Rand, & Zeckhauser, 2011; Paolacci, Chandler, & Ipeirotis, 2010). Each participant received $.55. We promoted data quality by allowing participants to begin the study only if they correctly answered a reading comprehension question. By *a priori* decision, we retained only those participants who spent between 4 and 25 minutes completing the study (*n* = 337) and responded correctly to attention-checks described below. Our analyses are based on the remaining 310 participants. The direction and significance of results were the same when we retained all participants for analysis.

Materials. We wrote and realistically formatted a brief article based on actual reports about Noel Biderman, the founder and CEO of Ashely Madison, a website that connects people seeking extra-marital affairs (the Online Supplement contains the article's full text). At the time of the study, the website had been in the news for being hacked. Participants read different versions of the article depending on randomly assigned condition.

Manipulation. Participants in the *inconsistent-virtue* condition read factual information that Biderman promoted adultery in a published book entitled "Cheaters Prosper" and other public statements ("Life is short. Have an affair"), but that investigative reporting discovered he

[1] The research ethics boards that approved Study 3 withheld permission to post even anonymized data publicly, because the consent forms did not inform participants that their data would be used in this way. However, data from that study are available upon request.

had privately remained faithful to his wife.[2] Participants in the *no-virtue* condition only read that he promoted adultery. Thus, only in the inconsistent-virtue condition did he enact private virtue that directly contradicted his less-than-virtuous public image. Note that these conditions provide a particularly strong test of our hypotheses. That is, we expected participants to rate Biderman as more hypocritical in the inconsistent-virtue condition even though this condition provides more positive information about him (i.e., his marital fidelity).

Hypocrisy measure. Participants reported how much they thought the founder, Biderman, was a hypocrite on an established five-item hypocrisy measure (Effron et al., 2015): how hypocritical, two-faced, phony, genuine (reverse-coded), and insincere they found him (α = .90; 1 = *Disagree Strongly*; 7 = *Agree Strongly*). (Because some studies use a single-item measure of hypocrisy – e.g., Effron & Monin, 2010 – we reran the analyses in all studies using only the single item about whether the target person was a hypocrite. All results remained significant and in the same direction).

Praise measure. Participants indicated how positively they viewed the founder and his behavior using a 13-item scale adapted from previous research about hypocrisy (α = .90; Effron et al., 2015; Effron & Monin, 2010).[3] Nine of these items measured the founder's character on 7-point bipolar scales (starred items were reversed-coded): cruel-kind, nice-awful*, cold-warm, honest-dishonest*, unfair-fair, moral-immoral*, arrogant-humble, good-bad*, likable-dislikable*; three of the items assessed condemnation of the founder's leadership behavior at the company ("Noel Biderman should be punished harshly"; "As CEO of Ashley Madison when the

---

[2] The article emphasized actual accounts of his monogamy with several additional fictional quotes, such as "There is no way I would ever cheat."

[3] This previous research used the measure to assess moral condemnation of transgressions, where higher numbers indicated more negative evaluations. To better capture our interest in praise for virtuous behavior, we reverse-coded the measure so higher numbers indicate more positive evaluations.

hacks occurred, Noel Biderman should be held responsible"; "Noel Biderman should be sued for any impending damages or liability"; 1 = *Disagree Strongly*, 7 = *Agree Strongly*, all reverse-coded); for the last item, participants reported how much they agreed or disagreed on a 7-point scale with the statement, "If Noel Biderman were looking for another job, organizations definitely should not hire him" (reverse-coded). Results do not vary if we analyze judgments of the target's character (9-items) and his behavior (4-items) separately.

To avoid demand characteristics, we administered the praise measure before the hypocrisy measure. That is, we did not want to signal to participants that they should consider hypocrisy when judging the founder's praiseworthiness.

Attention checks. Last, two multiple-choice attention checks assessed whether participants comprehended the focal domain manipulations of the organization's work context (e.g., investing, infidelity, cooking, television streaming; n = 397) and the target's private behavior (e.g., No, the article did not report any information about his romantic life; Yes, he is married and a regular user of Ashley Madison; Yes, he is married and has never cheated in the last 12 years; n = 377).

Exploratory items. In this and in subsequent studies, participants also rated the company on the same 9 semantic differential items used to assess the target person's character. Because our theorizing makes no specific predictions about these items and they yielded inconsistent results across studies, we report them in the Online Supplement and do not discuss them further.

Results and Discussion

Supporting our prediction, participants judged the founder who publicly promoted marital infidelity as more hypocritical when they knew he privately practiced monogamy (inconsistent-virtue condition: M = 5.23, SD = 1.42) than when they did not know (no-virtue condition: M =

4.62, *SD* = 1.31), *t*(309) = 3.93, $p < .001$, *d* = .45. That this effect occurred even though the inconsistent-virtue condition provided more positive information about the founder than the no-virtue condition speaks to the effect's strength.

We next examined a potential downstream consequence of these hypocrisy judgments: reduced praise. As predicted, enacting virtuous behavior that was inconsistent with his public image made the founder seem more hypocritical, which in turn predicted less praise for him and his behavior (see Figure 1). That is, we found a significant negative indirect effect of condition (coded 1 = inconsistent-virtue; 0 = no-virtue) on praise through hypocrisy, *b* = -.29, *SE* = .09, bias-corrected 95% CI computed with 5,000 bootstrap resamples = [-.44, -.15] (Hayes, 2013).

Finally, we examined whether inconsistent, virtuous behavior reduced the founder's praiseworthiness overall – the total effect of the manipulation on praise without specifying an indirect pathway through hypocrisy judgments. We considered this analysis exploratory. On one hand, the apparent hypocrisy of the inconsistent-virtue should decrease praise overall, as the indirect-effect analysis above suggested. On the other hand, the inconsistent-virtue condition provided more positive information about the founder than the no-virtue condition (i.e., that he practiced monogamy), which should increase praise overall. It is unclear what the net effect of these competing processes will be. The results showed that participants expressed more praise for the founder overall when they knew about his marital fidelity (*M* = 3.10, *SD* = 1.04) than when they did not (*M* = 2.52, *SD* = 0.92), *t*(310) = 5.18, $p < .001$, *d* = .59. Thus, in this scenario, his fidelity's virtuousness outweighed its hypocrisy in participants' overall judgments.[4]

---

[4] Comparing the manipulation's *total* effect on praise (i.e., without specifying the path through hypocrisy) to its *direct* effect on praise (i.e., partialing out the path through hypocrisy) provides a statistical illustration of these competing mechanisms. The total effect was significantly weaker than the direct effect, *b* = .58, *SE* = .11, $p < .001$ and *b* = .87, *SE* = .09, $p < .001$, respectively – a case of statistical suppression (MacKinnon, Krull, & Lockwood, 2000). Thus, learning that the founder practiced marital fidelity would have increased the praise he received even

To summarize, Study 1 provides initial support for our hypotheses that inconsistency between a person's private virtues and public behavior increases ascriptions of hypocrisy, and thereby indirectly reduces how much praise the person receives. As noted, Study 1 provided a particularly conservative test because the founder was perceived as more hypocritical in the condition in which he was known to do more good (i.e., participants knew that he had been faithful only in the inconsistent-virtue condition). Still, a cleaner test of our hypotheses would isolate the effect of inconsistency from the effect of doing good. For example, such a test could compare judgments of two target people who practice virtue despite preaching vice: one whose virtue and vice are directly inconsistent (e.g., publicly promotes smoking, but privately donates to anti-smoking causes) and one whose virtue and vice are unrelated (e.g., publicly promotes smoking but privately donates to an anti-obesity cause; or publicly promotes unhealthy eating but privately donates to an anti-smoking cause). Although both targets practice private virtue in all cases, we predict that participants will judge it as more hypocritical if the charitable behavior is inconsistent with (vs. unrelated to) public preaching. We predict that these ascriptions of hypocrisy will mean greater condemnation of the inconsistent virtue than of the unrelated virtue. Study 2 tested this prediction.

Study 2: Hypocrisy for Combatting Addiction

Participants read excerpts from an ostensible case study describing an individual who spent time and money privately contributing to a charitable cause (e.g., reducing tobacco use or reducing gambling addiction). The materials described an organization where the individual was publicly known to work (e.g., consulting for a tobacco or gambling organization) to manipulate whether the mission was inconsistent with or unrelated to the individual's private volunteering.

---

more if his fidelity had not also made him seem hypocritical (for another suppression effect of hypocrisy, see Effron & Monin, 2010).

Method

Participants. One hundred and eighty-three undergraduate students in the U.S. (M_{age} = 20.81 years, *SD* = 1.10; 106 males, 77 females) were recruited for a case study for course credit. With a limited university participant pool, we used a rule-of-thumb to recruit two hundred participants, which after some expected attrition and no-shows would be sufficient to detect moderate effect sizes across conditions. Applying Study 1's exclusion criteria, we dropped four participants who failed an attention check about the organizational context, seventeen who failed a check about the volunteering behavior, and six who spent fewer than four minutes or more than twenty-five minutes as the predetermined study duration window. The analyses below are based on the remaining 156 responses. The direction and significance of results were unaffected if we retained all participants for analysis.

Materials. For a study listed as a "Case Analysis," undergraduates read two excerpts about an organization formatted to resemble cases they had read in class. The first excerpt described an organization in the "sin industry;" depending on the version, it either helped tobacco companies increase sales or helped casinos increase gambling revenue. The second excerpt described a vice president of accounting at this organization, Joseph Larson, who has been working pro-bono 10 hours per week on non-profit campaigns and making "lump sum donations anonymously"; depending on the version, his efforts helped treat either tobacco or gambling addiction.

We combined these materials to create two randomly assigned conditions: an *inconsistent-virtue* condition in which the individual's volunteering directly contradicts his public behavior (i.e., he publicly promotes but privately fights either tobacco or gambling), and an *unrelated-virtue* condition in which his volunteering does not contradict his public behavior

(i.e., he publicly promotes tobacco and privately fights gambling, or vice versa). This design ensures the specific virtuous behavior remains unconfounded from its inconsistency with his public behavior.

Praise measure. After reading these case details, participants responded to the same 13-item praise scale used in Study 1, with three items revised to fit Study 2's scenario. Specifically, the three items in Study 1 that asked about the founder's behavior here assessed the individual's volunteering behavior on 7-point scales regarding how honorable, moral, and praiseworthy it was. Last, participants responded to the same item, if the target were looking for a job, "organizations definitely should not hire him" (reverse coded). As in Study 1, we averaged the 13 items with appropriate reverse-coding to create a praise measure ($\alpha = .91$).

Hypocrisy measure. Next, participants reported how much they thought the individual was a hypocrite on the same 5-item scale used in Study 1 ($\alpha = .86$).

Harm measure. It is possible that any indirect effect of reduced praise for virtuous behavior in the inconsistent-virtue condition could be driven not by hypocrisy judgments of the individual, but instead by perceptions that he was directly causing harm to his own organization (Effron et al., 2015). To address this potential alternative explanation, participants answered how much they agreed or disagreed that the individual's "actions harmed the organization he works for (The Taylor Group)" (1 = *Disagree Strongly*, 7 = *Agree Strongly*). This general item encompassed a number of types of harm, such as whether his behavior harmed the organization's reputation, its credibility, and its financial standing.

Attention check. Last, two multiple-choice attention checks assessed whether participants comprehended the focal domain manipulations of the organization's work context (e.g., tobacco product sales, gambling revenues, or software development, etc.) and the target's

volunteering behavior (e.g., helping tobacco addiction, helping gambling addiction, working on environmental projects, etc.).

Results

Supporting our prediction, participants judged the individual as more hypocritical when his private charitable behavior directly contradicted his less-virtuous public behavior than when it did not (inconsistent-virtue condition: $M = 3.64$, $SD = 1.16$; unrelated-virtue condition: $M = 3.08$, $SD = 1.06$), $t(154) = 3.13$, $p = .002$, $d = .50$.

Participants also praised the individual less for his charitable behavior when it was inconsistent with his public behavior than when it was not (inconsistent-virtue condition: $M = 5.10$, $SD = 1.00$; unrelated-virtue condition: $M = 5.46$, $SD = 0.71$), $t(154) = 2.57$, $p = .011$, $d = .41$. Thus, doing some good may receive more praise than doing no good overall (Study 1), but the same good deeds receive less praise when they contradict one's public image than when they do not (Study 2). Additional analyses showed that the specific type of volunteering behavior or work context did not significantly moderate these results.

Last, we tested and found a significant indirect effect of our manipulation on praiseworthiness through hypocrisy, (see Figure 2), $b = -.12$, $SE = .04$, bias-corrected 95% CI = [-.22, -.05] (Hayes, 2013). Although this finding is consistent with our prediction that hypocrisy would result in reduced praise, analyses such as these cannot address the causal order between the mediator and the dependent variable (Fiedler, Schott, & Meiser, 2011). Additional details about these analyses and other robustness checks are reported in the Online Supplement.

Discussion

Study 2's results provide additional support for our hypotheses. Inconsistency between private virtues and public behavior increased ascriptions of hypocrisy. Overall, this meant that an

individual received less praise for the exact same charitable behavior (e.g., helping anti-tobacco campaigns) when it was inconsistent with his public behavior (e.g., he worked for a tobacco company) than when it was not (e.g., he worked for a fast-food company). Participants' perceptions of how much the target person's virtuous behavior had harmed his company could not fully account for these results.

Study 3: Replicating and Generalizing

Study 3 tested the robustness and generalizability of our prior findings in several ways. First, whereas Studies 1 and 2 recruited participants from American contexts, Study 3 targeted an internationally diverse sample. Second, Study 3 tested our hypotheses across a sample of 20 different virtue-vice vignette contexts. Finally, although Studies 1 and 2 took pains to present the target person's charitable behaviors as private, it is conceivable that participants inferred that the target was still sending false signals to some public audience. To address this limitation, Study 3's materials emphasized that the virtuous behaviors were completely private and anonymous, explicitly stating that the target never told anyone about the charitable behavior and was the only person who knew about it. Study 3 also included a comprehension check to ensure participants understood this behavior was truly private.

Method

We pre-registered this study at https://aspredicted.org/fb3e3.pdf.

Participants. Residents of London, UK enrolled in a business school's subject pool completed this study as part of a series of unrelated studies. Each study session lasted approximately 1 hour, and participants received £10. Our pre-registered sample target was 300 participants based on the resources available and the number of participants required for other studies in the series. Because the lab overbooked slots to guard against no-shows, we ended with

319 people, each of whom responded to three vignettes for a total of 957 responses. The three vignettes were randomly sampled for each participant from our pool of 20 (see below). Following our pre-registration, we excluded responses to any vignettes on which a participant failed at least one of three attention checks (described below). Our final sample was 317 people who provided 847 responses. Of the 314 people who provided demographics, 226 were women, 105 were men, and 1 was non-binary (45% White, 23% South Asian, 14% East Asian, 6% Black, 12% other race or ethnicity; $M_{\text{age}} = 26.00$ years, $SD = 9.52$). Nationals from 63 different countries participated.

Materials and manipulation. The stimuli were vignettes about organizations that cause harm, and people publicly associated with those organizations who act virtuously in private. We created these vignettes by writing 20 descriptions of harmful organizations in a variety of industries (e.g., a factory farm that abused animals; a for-profit school that charges high tuition for second-rate education) and 20 descriptions of an employee’s private, virtuous behavior – one for each industry (e.g., making anonymous, lump-sum donations to improve animal welfare or to help mentor and tutor high-school students applying to college). Depending on randomly-assigned condition, participants read vignettes in which the private virtuous behaviors were either *inconsistent* with the employees’ public work (e.g., a factory-farm employee donates to help animals) or *unrelated* to the public work (e.g., a factory-farm employee donates to help high-school students). Thus, both conditions used the same harmful organizations and privately virtuous employees, but varied which employees worked for which organizations. Each vignette emphasized that only the employee knew about the charitable behavior. The Online Supplement contains the full text for all materials and attention checks, and more detail about how we programmed the randomization.

Procedure. Each participant saw a random selection of three vignettes, all in the same condition. For each vignette, participants first read about the organization and then answered a multiple-choice, comprehension-check question about its industry (e.g., "Based on the vignette you read, what does the Taylor Group do?" Answer: factory farming). Next participants read about the charitable behavior of an employee at the same organization and answered the same 5-item hypocrisy scale as in Studies 1-2, a comprehension-check about the charitable behavior ("What did Jeff Petersen do in the last story you read?" Answer: donated to protect animal welfare), and a comprehension check about how many other people knew about the employee's donation behavior (Answer: no one else knew).

Results

As hypothesized, collapsing across the three repeated measures, participants thought the target person was more hypocritical in the inconsistent-virtue condition ($M = 3.76$, $SD = 1.40$) than in the unrelated-virtue condition ($M = 3.23$, $SD = 1.41$), $d = .37$. This difference was significant in a mixed regression model with fixed effects for condition (1 = inconsistent-virtue, 0 = unrelated-virtue), the specific charitable behavior, and whether a response corresponded to the first, second, or third vignette they rated (dummy-coded), plus random intercepts for participants to account for the fact that each participant responded to up to three vignettes, $b = .54$, $SE_b = .13$, $z = 4.07$, $p < .001$ by either a pre-registered one-tailed test or a two-tailed test.

As a robustness check, we planned to rerun the analyses, including participants' responses only if they correctly recalled that the target person's donation had been completely anonymous. However, this check proved unnecessary because all participants correctly recalled the relevant information (i.e., that only the target employee knew about the donation behavior).

Discussion

Study 3's results replicate our primary finding from Studies 1 and 2. Participants rated the target individuals' private and anonymous virtuous behavior as more hypocritical when it was inconsistent with the targets' public image at work – even though participants confirmed they understood that the private behavior was completely unknown to anyone but the target people themselves. Attesting to the generalizability of our results, these effects emerged with internationally diverse participants judging virtues and vices related to 20 different contexts – from employee rights to animal welfare to childhood education.

The false-signaling perspective on hypocrisy would not have predicted the results of Studies 1–3, because the target people did not send others a false signal of virtue; only the targets themselves knew about their private good deeds. By contrast, the results fit well with the undeserved-benefits perspective. Participants may have interpreted the private behaviors as indicating that the targets were trying to morally cleanse themselves by assuaging their guilt in order to feel more virtuous. However, our studies thus far have not directly measured whether participants interpreted the private behaviors in this way. Study 4 tests mediation by such a measure.

Study 4: Mediation by Guilt Reduction

Study 4's participants again read about a target person whose private virtues were either inconsistent with or unrelated to his less-than-virtuous public work. This time, they also indicated how much they thought the target had performed these virtues to alleviate his guilt. We predicted that participants would infer stronger guilt-reduction motives when the private virtues were inconsistent with the public work, which in turn would predict participants perceiving more hypocrisy and offering less praise.

Recall that we were agnostic about whether *successful* moral-cleansing would be necessary for private virtues to appear hypocritical, or whether *attempted* moral-cleansing would be sufficient. Study 4 does not attempt to distinguish between these possibilities with its measure of guilt-reduction motives, because such motives could precede both successful and unsuccessful attempts. Study 5 teases apart these possibilities with an experimental manipulation.

Method

Participants. We recruited 450 participants from MTurk in exchange for $.55 ($M_{age}$ = 38.04 years, *SD* = 12.81; 212 males, 238 females; 78% employed full- or part-time, 21% full students not employed). Based on Study 1 and 2's results, we used a rule-of-thumb that 200 people per condition would be sufficient to detect our predicted effects, including some attrition. After applying the same exclusion criteria from Studies 1 and 2, 409 participants remained. The direction and significance of results was the same when we retained all participants for analysis.

Materials. The design was similar to Study 2. For a study called "Attitudes and Perceptions of a Company Profile," participants read two ostensible excerpts from a case study. The first described an organization, the Taylor Group, whose mission was to increase sales of either tobacco or fast-food. The second excerpt described an employee, Joseph Larson, who privately volunteered to combat either tobacco addiction or obesity. To emphasize that the volunteering was private, the materials stated that the individual works on these charitable causes from his "personal computer in his private home and never tells anyone about it," and, similar to Studies 2-3, that whenever he donates, he does so anonymously. Like Studies 2-3, combining these materials produced an *inconsistent-virtue* condition in which his charitable work directly contradicted his public behavior (i.e., he publicly promotes but privately fights tobacco, or publicly promotes fast food but privately fights obesity) and an *unrelated-virtue* condition in

which his charity and his work are not inconsistent (i.e., he publicly promotes tobacco and privately fights obesity, or publicly promotes fast food and privately fights tobacco). As in Studies 2-3, this design avoids confounding the specific virtuous behavior with the manipulation of inconsistency. The Online Supplement reports the material's full text.

Measures. After reading both excerpts, participants answered the 13-item praise scale from Study 2 ($\alpha = .96$), the 5-item hypocrisy scale from Studies 1-3 (main dependent measure; $\alpha = .94$), and the harm item from Study 2 (covariate).

Moral cleansing involves behaviors intended to reduce one's guilt (Gollwitzer & Melzer, 2012; Lee & Schwarz, 2010). Thus, for our moral cleansing measure (and proposed mediator), participants completed a 5-item guilt scale adapted from prior research (Izard, Libero, Putnam, & Haynes, 1993), indicating their agreement that the target was motivated by feeling guilt, regret, remorse, blameworthy, and responsible for wrongdoing (1= *Disagree Strongly*, 7 = *Agree Strongly*; $\alpha = .97$). As noted, this measure does not distinguish between perceptions of *attempted* vs. *successful* guilt reduction because participants who infer that the target person was motivated by guilt may also infer that he successfully reduced his guilt.

Last, attention checks asked whether individuals were able to correctly recall the Taylor Group's mission and target individual's behavior.

Results

As in Studies 1-3, and supporting our prediction, participants judged the individual as more hypocritical when his private volunteering directly contradicted his less-virtuous public behavior than when it did not (inconsistent-virtue: $M = 3.86$, $SD = 1.62$; unrelated-virtue: $M = 3.30$, $SD = 1.60$), $t(407) = 3.51$, $p = .001$, $d = .35$. Going beyond our previous studies, participants thought that guilt was more likely to have motivated the target's private good deeds

in the inconsistent-virtue condition (M = 4.78, SD = 1.57) than in the unrelated-virtue condition (M = 3.48, SD = 1.73), $t(407) = 7.95$, $p < .001$, $d = .79$. As in Study 2, they also found the individual less praiseworthy overall in the inconsistent-virtue condition (M = 5.18, SD = 1.24) than in the unrelated-virtue condition (M = 5.39, SD = 1.17), but this effect was not significant, $t(407) = 1.77$, $p = .077$, $d = .18$. As before, the specific context of volunteering or organization mission did not significantly moderate these results.

Last, we replicated the same indirect effect that we found in Study 2, from the manipulation (inconsistent-virtue coded +1; unrelated-virtue coded –1), to hypocrisy, to praise, b = -.15, SE = .05, bias-corrected 95% CI = [-.24, -.07], computed with 5,000 bootstrap resamples (Hayes, 2013). Going beyond our previous studies, we also found a significant indirect effect from the manipulation (inconsistent-virtue = 1, unrelated-virtue = -1) through guilt-reduction to hypocrisy judgments, b = .33, SE = .05, bias-corrected 95% CI [.24, .43], computed with 5,000 bootstrap resamples. We also found a significant indirect effect such that the target's inconsistency made him appear more motivated by guilt, which in turn predicted higher hypocrisy ratings, which in turn predicted lower praise for his good deeds, b = -.19, SE = .03, bias-corrected 95% CI [-.26, -.13] (see Figure 3). The Online Supplement reports additional details about these indirect effects, as well as structural equation analyses suggesting that the relevant models fit the data better than plausible alternative models.

Replications

The Online Supplement reports two additional studies that replicated the main effect of condition on hypocrisy and the indirect effect from the manipulation through guilt-reduction and hypocrisy to praise.

Discussion

Study 4 replicates Studies 1–3: People found a target person more hypocritical when his virtues were inconsistent with (vs. unrelated to) his vices, even though the materials clearly stated that the target enacted the virtue in private and never told anyone about it. Going beyond Studies 1–3, the results offer mediational evidence consistent with our proposed moral-cleansing mechanism. Specifically, participants were more likely to explain a target person's private virtuous behavior as focused on reducing his guilt when it was inconsistent with (vs. unrelated to) his public vices – and these imputed attributions in turn predicted harsher hypocrisy judgments. These results fit with our claim that audiences will find an actor's private virtues more hypocritical to the extent that audiences interpret these virtues as moral cleansing.

However, as with any measurement-of-mediation design, we cannot draw causal conclusions about the relationship between the mediation and the dependent variable, nor can we rule out the role of other, unmeasured variables (Bullock, Green, & Ha, 2010; Fiedler et al., 2011). Moreover, it is unclear whether participants in this study saw hypocrisy in the target person's *attempt* at moral cleansing, or whether they inferred hypocrisy because they presumed that the moral cleansing had been successful and the target person felt absolved of guilt for his public vices. Study 5 addresses these issues by directly manipulating the moral cleansing mechanism and distinguishing between attempted and successful guilt-reduction.

Study 5: Why Charity Becomes Hypocrisy – Attempted or Successful Guilt Reduction?

Study 5's participants read about the harmful public behavior of a target person whose private charitable contribution was described (a) as a *successful* attempt to alleviate his guilt, (b) as an *unsuccessful* attempt to alleviate his guilt, or (c) without mentioning any such feelings (control condition). In this way, Study 5 directly manipulated moral cleansing. If *attempted*

moral cleansing is sufficient to produce hypocrisy judgments, then participants should perceive more hypocrisy in the two conditions involving guilt motives than in the control condition. If, on the other hand, *successful* moral cleansing is necessary to appear hypocritical in this context, then participants should perceive more hypocrisy in the first condition than in the other two.

The undeserved-benefits perspective posits that moral cleansing seems hypocritical to the extent that it involves feeling (or attempting to feel) more virtuous than one deserves. Study 5 thus included a direct measure of how much participants thought the target person felt overly virtuous and tested this measure as a mediator.

Study 4 suggests that private good deeds seem more like an attempt to reduce guilt when they are inconsistent with (vs. unrelated to) a person's less-virtuous public image. However, the undeserved-benefits perspective does not require that a person display inconsistency to be perceived as hypocritical. Inconsistency can be a cue that a person is claiming such a benefit, but it is not the only cue (cf. Monin & Merritt, 2012). Study 5 tested the prediction that even when the target person's public and private behavior are not directly contradictory, the target will seem more hypocritical when the private behavior involved (successful or perhaps merely attempted) moral cleansing. In other words, we tested whether manipulating the target's guilt reduction could "turn on" hypocrisy judgments in the condition that seemed least hypocritical to our previous studies' participants (i.e., the unrelated-virtue condition). Thus, all participants considered a target person who did something charitable in a domain unrelated to his public image (e.g., a fast-food executive who contributes to anti-smoking causes).

To ensure that participants understood the target person's good deeds were unambiguously private and anonymous, Study 5 participants read that no one else knew about

the charitable behavior, and then completed a comprehension check. Additionally, we tested generalizability to non-American participant populations.

Method

We preregistered the study at https://aspredicted.org/hr8pw.pdf.

Participants. Twelve hundred participants signed up for "A Business Profile Study" on Prolific Academic (ProA), a UK-based online data collection platform whose participants have been shown to be more diverse and provide data of comparable, if not higher, quality compared to competing platforms (Peer, Brandimarte, Samat, & Acquisti, 2017). To recruit English-speaking participants outside North America, we restricted our sample to individuals from the U.K., Australia, New Zealand, and Ireland (M_{age} = 36.46 years, *SD* = 13.25; 747 female (66%), 381 male (34%), 85 (7%) no response; 78.6% employed full or part-time). To ensure the total number of completed responses, the system allowed 1275 participants to begin the survey. Each participant earned $1.21 for completing the study.

Materials. Similar to previous studies, participants read a two-part vignette labeled as a "Business Profile" task. The first part described an ostensibly real pharmaceutical organization, Pantheon Pharma, that manufactures popular opioid pain-relief medications that are highly addictive and linked with life-threatening outcomes. The description was based on the recent case of Purdue Pharma and the Sackler family, who manufactured the opioid OxyContin (Van Zee, 2009). The article described the family who owns the company and marketed these drugs knowing that they were associated with life-threatening addictive risks.

The vignette's second part described an employee at the company who recently made a charitable contribution to an unrelated cause: an art museum (a detail from the real Purdue Pharma case). In this second vignette, we randomly assigned participants to one of three

conditions: *control*, *guilt-persists* (i.e., attempted moral cleansing), and *guilt-relief* (i.e., successful moral cleansing). All participants read about the same charitable behavior of Raymond Sampson, current CEO of Pantheon Pharma. In the control condition, participants read that Sampson is a "long-time art lover and collector. Recently, he decided to make an anonymous donation to an art museum in the northeastern United States. Sampson made the donation anonymously and in a lump-sum payment so that he would be the only person who knew about it."

In both the *guilt-persists* and *guilt-relief* conditions, participants also read about why he donated: "In recent years, he has felt increasing guilt about the company's actions and patient outcomes linked to Pantheon's addictive painkillers." Then participants in the *guilt-persists* condition read:

> After making the donation to the art museums, Sampson feels no relief and just as bad, if not worse, about himself than he did before. Given what his company did, he still feels guilty. He can't sleep at night and thinks of himself as a bad person. 'I feel so guilty,' he thinks to himself.

In the *guilt-relief* condition, participants instead read:

> After making the donation to the art museums, Sampson feels overwhelming relief and much better about himself. Despite what his company did, he doesn't feel guilty anymore. Now he can sleep at night and feels like he can return to thinking of himself as a good person. 'Goodbye guilt,' he thinks to himself.

The Online Supplement contains the full text of these materials.

Measures. Participants answered the 13-item praise scale from Studies 2–4 (α = .94) and the 5-item hypocrisy from Studies 1–4 (α = .86). As a manipulation check, participants answered a 6-item guilt scale similar to Study 4's, but slightly re-worded to assess how much guilt he felt *after* making the donation (α = .92).

As a direct measure of the undeserved moral benefits the target person might experience in this context, Study 5 asked participants whether he felt more virtuous than he deserves to feel using a 7-item measure: “After donating, to what extent do you think that Raymond Sampson feels…better about himself, more pleased, more proud, less regretful, less remorseful, less guilty, and more virtuous…” Each item ended with “…than he deserves” (1= *Disagree Strongly*, 7 = *Agree Strongly*; $\alpha = .96$). We put this measure last so it could not influence participants’ responses to the other measures. Our pre-registered prediction was that if participants perceived more hypocrisy in the guilt-relief condition than in the guilt-persists and control conditions, then this measure would mediate the effect.

Following our pre-registration, we excluded individuals who failed at least one of the multiple-choice attention checks that assessed whether participants were able to recall Pantheon Pharma’s behavior (Answer: “Manufacture and market painkiller drugs”; $n = 140$), where the CEO made his donation (Answer: “To an art museum”; $n = 178$) and how private and anonymous the donation was (“How many people know about the donation that Sampson made?” Answer: “Only Sampson knows”; $n = 213$). As before, we also excluded any participants with duplicate IP addresses ($n = 53$) or with an IP address outside the restrictions specified on the Prolific platform ($n = 16$). The analyses below are based on the remaining participants ($N = 1{,}008$). All results had the same direction and significance-level when we included all complete responses.

Because we tested the three conditions in a context where the donation behavior was unrelated to the work behavior, this study did not collect perceptions of harm to the organization.

Results

Manipulation check. Confirming that participants interpreted the manipulation as we expected, their ratings of the target's guilty feelings differed significantly by condition, $F(2, 1005) = 860.26$, $p < .001$. Participants thought the target felt significantly less guilt in the guilt-relief condition ($M = 2.18$, $SD = 1.05$) compared to both the guilt-persists condition ($M = 5.09$, $SD = 1.09$), $t(1005) = 37.10$, $p < .001$, $d = 2.72$, and the control condition ($M = 2.38$, $SD = 0.91$), $t(1005) = 2.53$, $p = .012$, $d = .24$. The latter two conditions were also significantly different from one another, $t(1005) = 34.35$, $p < .001$, $d = 2.70$.

Hypocrisy. Hypocrisy ratings differed significantly among the three conditions, $F(2, 1005) = 46.98$, $p < .001$ in a one-way ANOVA, and pairwise comparisons suggest that *successful* guilt-reduction was necessary to appear hypocritical, whereas *attempted* guilt-reduction was insufficient. Specifically, people rated the target as more hypocritical when his donation successfully alleviated his guilt about selling opioids (*guilt-relief* condition: $M = 5.33$, $SD = 1.19$) compared to when the donation failed its intended purpose (*guilt-persists* condition: $M = 4.48$, $SD = 1.14$), $t(1005) = 9.54$, $p < .001$, $d = .72$, and also compared to when no information about its purpose or effect was provided (*control* condition: $M = 4.76$, $SD = 1.18$), $t(1005) = 6.28$, $p < .001$, $d = .48$. Interestingly, participants rated the target person as less hypocritical in the guilt-persists condition than in the control condition, $t(1005) = 3.16$, $p = .002$, $d = .24$, perhaps because they gave him moral credit for continuing to feel guilty about selling opioids (Barasch et al., 2014; Stearns & Parrott, 2012).

Feeling better than he deserved. Our direct measure of undeserved moral benefits showed similar results. Participants thought that the target felt better than he deserved to feel when his donation absolved his guilt ($M = 6.12$, $SD = 1.09$), compared to when his guilt persisted ($M = 3.36$ $SD = 1.48$), $t(1005) = 28.53$, $p < .001$, $d = 2.12$, and also compared to the control

condition (M = 5.28, SD = 1.17), $t(1005) = 8.59$, $p < .001$, $d = .73$. Interestingly, participants also thought that he felt worse than he deserved to feel when his guilt persisted, compared to the control condition, $t(1005) = 19.73$, $p < .001$, $d = 1.44$.

Following our pre-registered analysis plan, we also found a significant indirect effect of condition (guilt-relief = +2, guilt-persists = -1, control = -1) on hypocrisy through perceptions that the target felt better than he deserved, $b = .19$, $SE = .02$, bias-corrected 95% CI = [.16, .23]. The Online Supplement reports additional details about this indirect effect.

Praise. The amount of praise participants gave the target for acting virtuously depended on the condition, $F(2, 1005) = 49.32$, $p < .001$. Participants rated the target as significantly less praiseworthy for the same charitable donation when he experienced guilt-relief (M = 2.53, SD = 1.12) than when his guilt persisted (M = 3.37, SD = 1.12), $t(1005) = 9.39$, $p <.001$, $d = .75$, and also compared to the control condition (M = 3.21, SD = 1.25), $t(1005) = 7.50$, $p <.001$, $d = .57$. The difference between the guilt-persists and control conditions was not significant, $t(1005) = 1.79$, $p = .075$, $d = .13$. Consistent with our pre-registered analysis plan, we also found that there was a significant indirect effect from this successful guilt-reduction condition (coded: +2; guilt-persists -1, control -1) on praise, through perceptions that the target feels better than he deserves to feel, and through hypocrisy judgments (see Figure 4), $b = -.12$, $SE = .01$, bias-corrected 95% CI = [-.15, -.10], computed with 5,000 bootstrap resamples (Hayes, 2013). The Online Supplement reports additional robustness tests of the indirect effects on praise.

Discussion

Studies 1-4 have argued that people view a target person as more hypocritical when his private virtues are inconsistent (vs. not inconsistent) with his or her public vices because the inconsistency makes the private virtues seem like moral cleansing. Whereas Study 4's mediation

analysis provided correlational evidence consistent with this mechanism, Study 5 provided causal evidence by directly manipulating moral cleansing (Spencer, Zanna, & Fong, 2005). The same private, virtuous behavior seemed more hypocritical when it alleviated the target person's guilt for his public vices than when it did not. The results also clarify a boundary condition: Using private virtues as a strategy to cleanse one's conscience may only seem hypocritical if the strategy works. In fact, perhaps because feeling guilty implies one has a conscience (Ketelaar & Tung Au, 2003; Tangney & Dearing, 2002; Tangney et al., 2007), continuing to feel guilty after attempted moral cleansing appears to send a positive signal to others (Barasch et al., 2014; Stearns & Parrott, 2012; Tangney & Dearing, 2002).

Why did participants perceive successful guilt-reduction as hypocritical? We predicted that a person who alleviated his guilt about selling harmful opioids by anonymously donating to the arts would be perceived as feeling more virtuous than he deserved – and this is exactly how participants perceived him. At least in this context, using good deeds to alleviate your guilt was tantamount to claiming an undeserved moral benefit.

Meta-Analysis

To more precisely estimate the size of our effects across all individual studies, including the two studies reported in the Online Supplement, we performed an internal meta-analysis. Our goals were to help future researchers estimate statistical power for studying our effect, and also to assess the robustness of the effect of inconsistency on praise, which (in contrast to the effect of inconsistency on hypocrisy) was not statistically significant in all studies.

Overall, our hypotheses received consistent support across seven studies (see Tables 1 and 2). Because Studies 1 and 5 did not include a separate unrelated-virtues condition, we could not include them in the analysis. Using Stata's *metan* function (Harris et al., 2008), we treated

study as a random effect, but treating it as a fixed effect using the inverse variance method produced identical conclusions.[5]

First, we examined the manipulation's effect on hypocrisy. Recall that each of Study 3's participants could respond to three vignettes; we computed Study 3's effect size based on each participant's mean hypocrisy rating across the three. As shown in Figure 5, participants in each study judged the target person as more hypocritical when his private virtues were inconsistent with his public behavior (supporting our focal prediction); across studies, the effect size was modest, $d = .44$ [.35, .54], and equivalent to the average effect size in published psychology research (i.e., $r = .21$, equivalent to $d = .43$; Richard, Bond, & Stokes-Zoota, 2003).

Next, we examined the effect of performing inconsistent (vs. unrelated) virtues on praise. Recall that except in Study 3 where we did not measure praise, our other studies all found significant indirect effects from the inconsistency manipulation to praise. However, the total effect of the manipulation on praise was not significant in all studies. Our theorizing does not require a total effect, but meta-analytically examining its reliability is useful for understanding inconsistency's potential downstream consequences. Across studies, there was a small but significant effect of the inconsistency manipulation on praise (see Figure 6), such that participants praised the same behavior less when it was inconsistent with (vs. unrelated to) a person's public behavior, $d = .18$ [.08, .29]. In addition to providing a better estimate of effect sizes across our studies, the meta-analysis also suggests that the manipulation's effect on hypocrisy and praise is reliable.

General Discussion

[5] Supplementary Study S1 included two hypocrisy measures that showed the same results. We averaged them for the meta-analysis, but the results are consistent if we use either measure independently. Results are also consistent if we include the Study 1 data. The *metan* function estimates random effects using the DerSimonian-Laird method.

Scholars typically characterize hypocrites as individuals who appear more virtuous in public than they behave in private. Yet in seven experiments with vignettes about 24 different behaviors and participants from more than 60 nationalities, we find that laypeople also judge others as hypocritical for virtuous private actions that directly contradict less-virtuous public behaviors. Participants viewed a founder who preached marital infidelity as more hypocritical when he privately practiced fidelity than when he did not (Study 1), and individuals working in a "sin-industry" as more hypocritical when their private volunteering behavior helped the very people their organizations harmed than when it helped people not harmed by their organizations (Studies 2-4, and Supplementary Studies S1-S2). Overall, the same private good deeds received less praise when they directly contradicted public vices than when they did not (Meta-Analysis), an effect that was associated with hypocrisy judgments in mediation tests (Studies 2 and 4, and Supplementary Studies S1-S2). For example, an individual received less praise for volunteering to fight tobacco addiction when he worked for a tobacco company than when he worked for a casino. Even when good deeds are private, anonymous, and secret – i.e., they are not motivated by a desire to manage others' impressions – the moral credit one receives for them depends on whether they contradict one's public behavior.

Why do private good deeds invite hypocrisy judgments when they contradict one's public behavior? It is easier to infer that the actor is trying to use these good deeds to compensate for harmful public behavior when the good deeds are directly inconsistent with the harm. For example, such moral cleansing is a more salient explanation for why a tobacco executive secretly donates to anti-tobacco causes than for why the executive secretly donates to anti-gambling causes. People give others less credit for moral behavior than the others give themselves (Effron, 2014; Kruger & Gilovich, 2004), and they assume others will overclaim credit for positive

outcomes (Kruger & Gilovich, 1999). Thus, observers may readily suspect that moral cleansing makes the actor feel more virtuous than they think she deserves. For example, they might infer that the anti-tobacco donation was an all-too-easy way to purchase a clean conscience. More generally, failing to practice in private the exact behavior you promote in public can suggest to observers that you are enjoying an undeserved moral benefit (Effron, O'Connor, et al., 2018), regardless of whether you are promoting virtue or vice. If you publicly promote virtue while privately practicing vice, the moral benefit is appearing more virtuous than you deserve. If you publicly promote vice while privately practicing virtue, the moral benefit is feeling more virtuous than you deserve.

Three key results support this interpretation of our findings. First, Study 4's participants were more likely to infer that a target person acted virtuously in order to alleviate his guilt when the good deeds were directly inconsistent with his less-virtuous public behavior than when they were not – and these inferences mediated an indirect effect from inconsistency to hypocrisy. This finding supports our prediction that inconsistency with public behavior would make good deeds seem like moral cleansing. Second, directly manipulating moral cleansing had a causal effect on hypocrisy judgments while holding inconsistency constant. That is, Study 5's participants ascribed more hypocrisy to the target person when his good deeds alleviated his guilt about his public behavior than when his good deeds did not. Third, we found direct evidence that people associate hypocrisy with undeserved moral benefits. Telling Study 5's participants that the target person had alleviated his guilt led them to infer that he felt more virtuous than he deserved, which in turn predicted how hypocritical they found him.

Theoretical Contributions

Our findings make several theoretical contributions. First, we demonstrate that people conceptualize hypocrisy more broadly than previously appreciated. Although the canonical hypocrite is someone who privately practices the same vice she publicly preaches against (Crisp & Cowton, 1994; Stone & Fernandez, 2008), we show that audiences will levy a hypocrisy penalty even when the vice is public and the virtue is private. Even in the absence of virtue-preaching, unambiguously prosocial behaviors like volunteering or donating to charity can seem more hypocritical if they directly contradict less-virtuous public behaviors.

Second, our results distinguish between two competing theories of hypocrisy: the false-signaling perspective (Jordan et al., 2017) and the undeserved-benefits perspective (Effron, O'Connor et al., 2018). Both theories predict that people will perceive hypocrisy in someone who privately practices the same vice she publicly preaches against. The inconsistency between her public preaching and her private virtue suggests that she appears more virtuous than she is – a "false signal" that counts as an undeserved moral benefit. Both theories can thus account for most prior work on lay judgments of hypocrisy, because this work mainly measures judgments of target people who act more virtuously in public than in private. However, the false-signaling perspective cannot account for the present findings. Because the target people in our scenarios acted *more* virtuously in private than in public, they did not falsely signal virtuousness to others – if anything, their public behavior undersells their prosocial behaviors – yet participants still perceived this inconsistency as hypocrisy (Studies 1–4). Moreover, manipulating a target person's private feelings affected how hypocritical he seemed to participants (Study 5), and a measure of private feelings mediated the effect of the target's inconsistency on participants' judgments of hypocrisy (Study 4). These effects occurred even though the studies held constant the target people's public behavior and thus the signals they sent to others.

In contrast to the false-signaling perspective, the undeserved-benefits perspective predicts exactly these findings. Private virtues that contradict public vices are also seen as claiming an undeserved moral benefit: feeling virtuous despite one's public vices. The extent to which one enjoys this benefit depends on private feelings (e.g., guilt), even holding public and private behaviors constant. And indeed, participants were more likely to identify the target person as feeling more virtuous than he deserved when his private virtues alleviated his guilt about his public behavior (Study 5).

Although the false-signaling perspective was only recently formalized and tested (Jordan et al., 2017), its assumption that hypocrisy requires presenting oneself as more virtuous than one truly is has guided much prior work (Aronson, Fried, & Stone, 1991; Batson et al., 1997; Monin & Merritt, 2012; Stone & Fernandez, 2008; Stone et al., 1997). Our results require rethinking this assumption. We do not challenge the notion that hypocrisy *can* involve deceitfully virtuous public display, but we argue that lay conceptualizations of hypocrisy are not fundamentally about false signals to others (see also Alicke et al., 2013; Hale & Pillow, 2015). It is worth emphasizing again that the undeserved-benefits perspective encompasses the false-signaling perspective. Appearing virtuous to others is a moral benefit, and so is feeling virtuous. People do not like those who claim either benefit without deserving it.

A potential way to reconcile our findings with the false-signaling perspective is to speculate that participants believed the target's private behavior falsely signaled a virtuous character to themselves. This intrapersonal interpretation is consistent with our claim that the targets were penalized for feeling more virtuous than they deserved. Said differently, the targets seem to have interpreted their private behavior as more of a "true signal" about their own moral character than participants thought they should. However, the idea that private behavior can

represent a false signal to oneself stretches the false-signaling theory of hypocrisy beyond its original formulation. Jordan and colleagues explicitly theorize that "hypocrites are disliked because they use their condemnation to mislead *other people* about their moral behavior" (Jordan et al., 2017, p. 356-357, italics added). Normative arguments about how hypocrisy should be defined similarly emphasize the importance of presenting a deceitfully virtuous image to an external audience (Monin & Merritt, 2012). Some researchers have argued that claiming to be fair in public while acting unfairly in private can involve self-deception (Batson et al., 1997, 1999), but studies on lay judgments of hypocrisy have treated self-deception as a factor that could *mitigate* hypocrisy (see Alicke et al., 2013, Scenario 1, p. 682). Thus, even considering that "feeling more virtuous than you deserve" could be described as false signaling to the self, our findings would not have been predicted by previous theoretical approaches to hypocrisy, which posit a central role of managing *others'* impressions.

In addition to supporting the undeserved-benefits theory of hypocrisy, our findings sharpen the theory. Previously, it was unclear whether *attempting to obtain* an undeserved moral benefit is sufficient to appear hypocritical, or whether *successfully obtaining* it is necessary (Effron, O'Connor, et al., 2018). Our results were more consistent with the latter possibility, at least when the moral benefit is feeling virtuous. In Study 5, learning that a secret charitable donation had successfully alleviated a target person's guilt made him seem more hypocritical, but learning that his donation was motivated by guilt without successfully alleviating it made him seem *less* hypocritical.

These findings point to an additional theoretical contribution: our demonstration that guilt can make people appear hypocritical. On one hand, guilt is a "moral emotion" that can drive pro-social behavior (Andreoni, 1990; de Hooge, Nelissen, Breugelmans, & Zeelenberg, 2011;

Ketelaar & Tung Au, 2003; Nelissen, Dijker, & deVries, 2007; Tangney & Dearing, 2002; Tangney et al., 2007). Laypeople tend to interpret moral emotions as indicative of moral character (Barasch et al., 2014), so virtuous behavior motivated by guilt may receive positive evaluations. On the other hand, virtues motivated by guilt could be seen as attempts to claim undeserved moral benefits, which would elicit negative evaluations. Study 5's results reconcile these competing predictions by suggesting that people will offer more praise for a good deed when it is motivated by guilt, but may offer less praise when the good deed successfully alleviates the guilt. A clean conscience can seem hypocritical if people suspect you feel more virtuous than you deserve.

Finally, our studies provide a novel perspective on moral cleansing and the potential disconnect between how people view their own versus others' good deeds. Previous work has shown that, following wrongdoing, individuals are motivated to reduce negative emotional states through moral cleansing (de Hooge et al., 2011; Gollwitzer & Melzer, 2012; Jordan et al., 2011; Mazar & Zhong, 2010; Sachdeva et al., 2009; Tangney et al., 2007; Tetlock et al., 2000). However, our studies are the first to our knowledge to directly examine how others evaluate behavior they perceive as moral cleansing. To oneself, these actions might resolve discomfort and make individuals feel a sense of enhanced or repaired moral identity. To others, however, these same actions may lead audiences to interpret good deeds that seem inconsistent with prior behavior as more hypocritical. The implication is that successful moral cleansing may wash away guilt, but leave the stain of hypocrisy.

Limitations and Future Directions

A potential alternative explanation is that participants saw hypocrisy in the possibility that the target person's charitable behavior might enable his vice behaviors at the harmful

organization. For example, by alleviating his guilt with an anti-tobacco donation, the tobacco executive might feel licensed to continue peddling tobacco (Effron, 2016; Effron & Conway, 2015; Effron & Monin, 2010). This explanation is similar to our own because it involves guilt reduction, but different because it assumes that guilt reduction is only hypocritical insofar as it enables the target person to continue causing harm. However, a study reported in the Online Supplement (Supplementary Study S1) found no support for this alternative. Participants read about a target who attempts to leave a harmful organization, but must continue to work there due to external constraints (e.g., a bad economy and no available job options) not because his charitable behavior enables him to stay. The results showed that participants still levied a hypocrisy penalty when the same virtuous behavior was inconsistent with (vs. unrelated to) the target's employment. Being forced to stay at the harmful organization did not interact with the inconsistent- vs. unrelated-virtue condition, and the effect sizes of these hypocrisy judgments for the target's charitable behavior were not measurably different when he could not leave the organization as when he could leave (*d*s = .47 vs. .49). The virtuous behavior may have alleviated his guilt about working at the organization, but it was not what caused him to stay. Thus, private good deeds can seem hypocritical even when they do not license harmful public behavior.

Feeling virtuous about good deeds should not always evoke hypocrisy. Our results suggest that private good deeds can seem hypocritical if they make an actor seem like he feels more virtuous than he deserves, and prior research suggests that observers have good reason to expect an actor to give himself more credit for his good deeds than observers are prepared to give (Effron, 2014; Kruger & Gilovich, 2004), particularly when they are aware of the actor's transgressions (Gawronski & Bodenhausen, 2006). Nonetheless, the right behavior can probably

convince observers that an actor has earned a clear conscience after transgressing, and the undeserved-benefits perspective predicts that observers would not perceive hypocrisy in such situations. For example, donating to the arts was apparently not an adequate basis, in the eyes of Study 5's participants, for feeling absolved of guilt after pushing addictive and lethal opioids – but donating one's entire fortune and devoting the remainder of one's life to fight poverty might have been seen as meriting a reduction in guilt. The question of what behaviors will convince observers that a clean conscience does not imply hypocrisy must await future research.

Our research captured a range of contexts, and our experimental methodology allowed us to directly test perceptions of the exact same behaviors under conditions holding all else equal. That said, experimental vignette studies can raise questions about generalizability to real-world situations. To limit concerns about generalizability, our vignettes used real-world organizational examples (Study 1 and 5), realistic materials such as news articles (Study 1), and familiar case formatting relevant to our student population (Study 2). A second limitation is that all our vignettes, inspired by real-world examples of the phenomenon they aimed to capture, equated the target's public behavior with a professional role such as pharmaceutical executive. We predict that our findings would generalize to non-professional contexts, but we leave that question to be addressed by future research. Third, it is important not to overinterpret the data from mediation analysis as causal evidence of mechanism (Bullock et al., 2010; Fiedler et al., 2011; Green et al., 2010). To bolster Study 4's mediational data with causal evidence, Study 5 manipulated our proposed mechanism. Fourth, future research should explore cultural differences in the results, as people make different attributions about moral inconsistency outside the West (Effron, Markus, et al., 2018). The majority of our studies were conducted with participants from the United States, although two studies with internationally diverse samples – including more than 60

nationalities from North and South America, Europe, Africa, Asia, and Oceania – showed broad evidence of cross-cultural generalizability (Studies 3 and 5).

A final limitation is that we were not able to test all incarnations of hypocrisy. On one hand, the fact that private and anonymous good deeds can seem hypocritical implies that false moral signals (Jordan et al., 2017), the intent to deceive others (Crisp & Cowton, 1994), and publicly condemning others (see Crisp & Cowton, 1994; Hale & Pillow, 2015) are not necessary to elicit hypocrisy judgments. Additionally, Study 5 suggests that even inconsistency itself may be unnecessary to elicit hypocrisy judgments (cf. Monin & Merritt, 2012). On the other hand, several instances of hypocrisy fell beyond the scope of our studies. For example, our results do not directly address the question of whether people judge inconsistency as less hypocritical when it results from "weakness of will," as in the case of an anti-drug crusader who fails to kick his private heroin addiction (Szabados & Soifer, 2004). The undeserved-benefits perspective would predict minimal perceptions of hypocrisy if people perceived the crusading addict as wracked with guilt, though this was not a focus of our studies. Similarly, we cannot address the question of whether transgressing one value in order to uphold another value would count as hypocrisy to laypeople (Alicke et al., 2013; Graham et al., 2015), although again the moral benefits perspective predicts the answer depends on the attributions people make about how the actor feels about the transgression. Last, appearing virtuous and feeling virtuous may not be the only moral benefits that could trigger hypocrisy judgments; so could having the right to influence others to be more or less virtuous. For example, a student who tells peers to party but then secretly studies is probably seen as a hypocrite (Alicke et al., 2013) because he claims the benefit of trying to influence others without paying the cost of doing the same thing himself. Future work should test the predictions of the undeserved-benefits theory about when these and other

ambiguous cases of moral inconsistency count or do not count as hypocrisy in the eyes of lay people.

Practical Implications

Our research points to the practical question of how employees in "sin-industry" organizations are judged for their good deeds outside work. For example, how would people think about the volunteering and charitable contributions by any of the 1.7 million people working in the gambling industry (Pierceall, 2014), the 3.8 million in fast food (Samadi, 2013), or the 12 million in tobacco manufacturing (International Labour Organization, 2014)? Our theorizing and results suggest that the answer may depend on how audiences believe the individuals feel about themselves in light of their good deeds. Some audiences may be more likely to cast them as hypocrites to the extent that they believe individuals feel better than they deserve – for example, the individuals feel virtuous despite not doing enough good deeds, doing the "wrong" kinds of good deeds, or doing good deeds that have little positive impact.

Conclusion

Seven studies document a broader interpretation of hypocrisy than has been previously understood: Even do-gooders face moral condemnation as hypocrites if their private charity appears inconsistent with their public-facing identity. Together, these provide new evidence that the shadow of hypocrisy can extend even to those virtuous deeds if they seem to help individuals morally cleanse themselves.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*(4), 1115–1153.

Alicke, M. (2008). Blaming badly. *Journal of Cognition and Culture*, *8*(1–2), 179–186.

Alicke, M., Gordon, E., & Rose, D. (2013). Hypocrisy: What counts? *Philosophical Psychology*, *26*(5), 673–701. https://doi.org/10.1080/09515089.2012.677397

Andreoni, J. (1988). Privately provided public goods in a large economy: the limits of altruism. *Journal of Public Economics*, *35*(1), 57–73.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464–477.

Aquino, K., & Reed, I. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423–1440.

Aronson, E. (1968). Dissonance theory: Progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of Cognitive Consistency: A Sourcebook* (pp. 5–27). Chicago: Rand McNally.

Aronson, E., Fried, C., & Stone, J. (1991). Overcoming denial and increasing the intention to use condoms through the induction of hypocrisy. *American Journal of Public Health*, *81*(12), 1636–1638. https://doi.org/10.2105/AJPH.81.12.1636

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology*, *107*(3), 393–413. https://doi.org/10.1037/a0037207

Barden, J., Rucker, D. D., & Petty, R. E. (2005). “Saying one thing and doing another”: Examining the impact of event order on hypocrisy judgments of others. *Personality and Social Psychology Bulletin*, *31*(11), 1463–1474. https://doi.org/10.1177/0146167205276430

Barden, J., Rucker, D. D., Petty, R. E., & Rios, K. (2014). Order of actions mitigates hypocrisy judgments for ingroup more than outgroup members. *Group Processes & Intergroup Relations*, *17*(5), 590–601. https://doi.org/10.1177/1368430213510192

Batson, C. D. (2002). Justice motivation and moral motivation. In M. Ross & D. T. Miller (Eds.), *The justice motive in everyday life* (pp. 91–106). Cambridge: Cambridge University Press.

Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., & Bayly, M. J. (1989). Negative-state relief and the empathy—altruism hypothesis. *Journal of Personality and Social Psychology*, *56*(6), 922–933. https://doi.org/10.1037/0022-3514.56.6.922

Batson, C. D., Dyck, J. L., Brandt, J. R., Batson, J. G., Powell, A. L., McMaster, M. R., & Griffitt, C. (1988). Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, Vol. 55, pp. 52–77. https://doi.org/10.1037/0022-3514.55.1.52

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, *72*(6), 1335–1348. https://doi.org/10.1037/0022-3514.72.6.1335

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, *83*(2), 330–339. https://doi.org/10.1037/0022-3514.83.2.330

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral

hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, *77*(3), 525–537. https://doi.org/10.1037/0022-3514.77.3.525

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bhatti, Y., Hansen, K. M., & Olsen, A. L. (2013). Political hypocrisy: The effect of political scandals on candidate evaluations. *Acta Politica*, *48*(4), 408–428.

Bryan, C. J., Adams, G. S., & Monin, B. (2013). When cheating would make you a cheater: Implicating the self prevents unethical behavior. *Journal of Experimental Psychology: General*, *142*(4), 1001–1005.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. https://doi.org/10.1037/a0018933

Carlsmith, J. M., & Gross, A. E. (1969). Some effects of guilt on compliance. *Journal of Personality and Social Psychology*, *11*(3), 232–239. https://doi.org/10.1037/h0027039

Crisp, R., & Cowton, C. (1994). Hypocrisy and moral seriousness. *American Philosophical Quarterly*, *31*(4), 343–349. https://doi.org/10.2307/20009796

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380.

de Hooge, I. E., Nelissen, R. M. A., Breugelmans, S. M., & Zeelenberg, M. (2011). What is moral about guilt? Acting "prosocially" at the disadvantage of others. *Journal of*

Personality and Social Psychology, *100*(3), 462–473. https://doi.org/10.1037/a0021459

Effron, D. A. (2014). Making mountains of morality from molehills of virtue: Threat causes people to overestimate their moral credentials. *Personality and Social Psychology Bulletin*, *40*(8), 972–985.

Effron, D. A. (2016). Beyond "being good frees us to be bad:" Moral Self-Licensing and the fabrication of moral credentials. In P. A. M. Van Lange & J.-W. van Prooijen (Eds.), *Cheating, corruption, and concealment: The roots of dishonesty.* (pp. 33–54). Cambridge UK: Cambridge University Press.

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32–35.

Effron, D. A., Lucas, B. J., & O'Connor, K. (2015). Hypocrisy by association: When organizational membership increases condemnation for wrongdoing. *Organizational Behavior and Human Decision Processes*, *130*, 147–159. https://doi.org/10.1016/j.obhdp.2015.05.001

Effron, D. A., Markus, H. R., Jackman, L. M., Muramoto, Y., & Muluk, H. (2018). Hypocrisy and culture: Failing to practice what you preach receives harsher interpersonal reactions in independent (vs. interdependent) cultures. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/J.JESP.2017.12.009

Effron, D. A., & Miller, D. T. (2015). Do as I say, not as I've done: Suffering for a misdeed reduces the hypocrisy of advising others against it. *Organizational Behavior and Human Decision Processes*, *131*, 16–32. https://doi.org/10.1016/j.obhdp.2015.07.004

Effron, D. A., & Monin, B. (2010). Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin*, *36*(12), 1618–1634.

https://doi.org/10.1177/0146167210385922

Effron, D. A., O'Connor, K., Leroy, H., & Lucas, B. J. (2018). From inconsistency to hypocrisy: When does "saying one thing but doing another" invite condemnation? *Research in Organizational Behavior*, *38*, 61–75. https://doi.org/10.1016/J.RIOB.2018.10.003

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137.

Feinberg, A. (2015). Family values activist Josh Duggar had a paid Ashley Madison account. *Gawker*. Retrieved from http://gawker.com/family-values-activist-josh-duggar-had-a-paid-ashley-ma-1725132091

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, Ill: Row, Peterson.

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, *47*(6), 1231–1236.

Frank, R. H. (1988). *Passions within reason: The strategic role of the emotions.* WW Norton & Co.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Anchor.

Gollwitzer, M., & Melzer, A. (2012). Macbeth and the joystick: Evidence for moral cleansing after playing a violent video game. *Journal of Experimental Social Psychology*, *48*(6), 1356–1360. https://doi.org/10.1016/J.JESP.2012.07.001

Graham, J., Meindl, P., Koleva, S., Iyer, R., & Johnson, K. M. (2015). When values and behavior conflict: Moral pluralism and intrapersonal moral hypocrisy. *Social and Personality Psychology Compass*, *9*(3), 158–170. https://doi.org/10.1111/spc3.12158

Green, D. P., Ha, S. E., & Bullock, J. G. (2010). Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose. *The Annals of the American Academy of Political and Social Science*, *628*(1), 200–208.

Greenbaum, R. L., Mawritz, M. B., & Piccolo, R. F. (2015). When leaders fail to "walk the talk": Supervisor undermining and perceptions of leader hypocrisy. *Journal of Management*, *41*(3), 929–956. https://doi.org/10.1177/0149206312442386

Hale, W. J., & Pillow, D. R. (2015). Asymmetries in perceptions of self and others' hypocrisy: Rethinking the meaning and perception of the construct. *European Journal of Social Psychology*, *45*(1), 88–98. https://doi.org/10.1002/ejsp.2064

Harris, R. J., Deeks, J. J., Altman, D. G., Bradburn, M. J., Harbord, R. M., & Sterne, J. A. C. (2008). Metan: fixed-and random-effects meta-analysis. *The Stata Journal*, *8*(1), 3–28.

Hayes, A. (2013). Introduction to mediation, moderation, and conditional process analysis. *New York, NY: Guilford*, 3–4. https://doi.org/978-1-60918-230-4

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. https://doi.org/10.1007/s10683-011-9273-9

Hugh-Jones, D., & Reinstein, D. (2012). Anonymous rituals. *Journal of Economic Behavior & Organization*, *81*(2), 478–489.

International Labour Organization. (2014). *Tobacco sector employment statistical update*. Retrieved from http://www.ilo.org/wcmsp5/groups/public/---ed_dialogue/---sector/documents/publication/wcms_329284.pdf

Izard, C. E., Libero, D. Z., Putnam, P., & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology*,

64(5), 847–860. https://doi.org/10.1037/0022-3514.64.5.847

Janney, J. J., & Gove, S. (2011). Reputation and corporate social responsibility aberrations, trends, and hypocrisy: Reactions to firm choices in the stock option backdating scandal. *Journal of Management Studies*, *48*(7), 1562–1585. https://doi.org/10.1111/j.1467-6486.2010.00984.x

Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, *530*(7591), 473–476.

Jordan, J. J., Sommers, R., Bloom, P., & Rand, D. G. (2017). Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychological Science*, *28*(3), 356–368. https://doi.org/10.1177/0956797616685771

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*(5), 701–713. https://doi.org/10.1177/0146167211400208

Karlan, D., & List, J. A. (2007). Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *American Economic Review*, *97*(5), 1774–1793.

Ketelaar, T., & Tung Au, W. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion*, *17*(3), 429–453. https://doi.org/10.1080/02699930143000662

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*(3), 190–194.

Kreps, T. A., Laurin, K., & Merritt, A. C. (2017). Hypocritical flip-flop, or courageous evolution? When leaders change their moral minds. *Journal of Personality and Social*

Psychology, *113*(5), 730–752. https://doi.org/10.1037/pspi0000103

Kruger, J., & Gilovich, T. (1999). " Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology*, *76*(5), 743–753.

Kruger, J., & Gilovich, T. (2004). Actions, intentions, and self-assessment: The road to self-enhancement is paved with good intentions. *Personality and Social Psychology Bulletin*, *30*(3), 328–339.

Lammers, J. (2012). Abstraction increases hypocrisy. *Journal of Experimental Social Psychology*, *48*(2), 475–480. https://doi.org/10.1016/J.JESP.2011.07.006

Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy: Moralizing in reasoning, immorality in behavior. *Psychological Science*, *21*(5), 737–744. https://doi.org/10.1177/0956797610368810

Laurent, S. M., Clark, B. A. M., Walker, S., & Wiseman, K. D. (2014). Punishing hypocrisy: The roles of hypocrisy and moral emotions in deciding culpability and punishment of criminal and civil moral transgressors. *Cognition and Emotion*, *28*(1), 59–83. https://doi.org/10.1080/02699931.2013.801339

Lee, S. W. S., & Schwarz, N. (2010). Washing away postdecisional dissonance. *Science*, *328*(5979), 709.

Lee, S. W. S., & Schwarz, N. (2011). Wiping the slate clean: Psychological consequences of physical cleansing. *Current Directions in Psychological Science*, *20*(5), 307–311.

Lei, P.-W., & Wu, Q. (2007). Introduction to structural equation modeling: Issues and practical considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33–43. https://doi.org/10.1111/j.1745-3992.2007.00099.x

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin*, *85*(5), 1030–1051.

Lönnqvist, J. E., Irlenbusch, B., & Walkowitz, G. (2014). Moral hypocrisy: Impression management or self-deception? *Journal of Experimental Social Psychology*, *55*, 53–62. https://doi.org/10.1016/j.jesp.2014.06.004

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science : The Official Journal of the Society for Prevention Research*, *1*(4), 173–181. https://doi.org/10.1023/A1026595011371

Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*(2), 101–121.

May, A. (2018, June 22). Husband cheating? Ashely Madison says member signups spiked in these cities last year. *USA Today*.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*(6), 633–644. https://doi.org/10.1509/jmkr.45.6.633

Mazar, N., & Zhong, C.-B. (2010). Do green products make us better people? *Psychological Science*, *21*(4), 494–498. https://doi.org/10.1177/0956797610363538

Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology*, *26*(2–3), 131–146.

Monin, B., & Merritt, A. (2012). Moral hypocrisy, moral inconsistency, and the struggle for moral integrity. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality:*

Exploring the causes of good and evil. (pp. 167–184). https://doi.org/10.1037/13091-009

Nelissen, R. M. A., Dijker, A. J. M., & deVries, N. K. (2007). How to turn a hawk into a dove and vice versa: Interactions between emotions and goals in a give-some dilemma game. *Journal of Experimental Social Psychology*, *43*(2), 280–286. https://doi.org/10.1016/J.JESP.2006.01.009

Newman, G. E., & Cain, D. M. (2014). Tainted Altruism: When Doing Some Good Is Evaluated as Worse Than Doing No Good at All. *Psychological Science*, *25*(3), 648–655. https://doi.org/10.1177/0956797613504785

Palanski, M. E., & Yammarino, F. J. (2011). Impact of behavioral integrity on follower job performance: A three-study examination. *The Leadership Quarterly*, *22*(4), 765–786. https://doi.org/10.1016/J.LEAQUA.2011.05.014

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, *5*(5), 411–419. https://doi.org/10.2139/ssrn.1626226

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163.

Pierceall, K. (2014). The US gambling industry is worth $240 billion. *Business Insider*. Retrieved from http://www.businessinsider.com/the-us-gambling-industry-is-worth-240-billion-2014-9

Powell, C. A. J., & Smith, R. H. (2013). Schadenfreude caused by the exposure of hypocrisy in others. *Self and Identity*, *12*(4), 413–431. https://doi.org/10.1080/15298868.2012.687013

Puente, M. (2015). Duggar's ex-employer denounces new scandal. *USA Today*. Retrieved from https://www.usatoday.com/story/life/2015/08/21/duggars-ex-employer-denounces-his-new-

scandal/32129687/

Ratner, R. K., & Miller, D. T. (2001). The norm of self-interest and its effects on social action. *Journal of Personality and Social Psychology*, *81*(1), 5–16.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Ross, M., & Miller, D. T. (2002). *The justice motive in everyday life*. Cambridge University Press.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, *28*(12), 1629–1646. https://doi.org/10.1177/014616702237645

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*(4), 523–528. https://doi.org/10.1111/j.1467-9280.2009.02326.x

Samadi, N. (2013). *Fast food restuarants in the US*. Retrieved from http://lgdata.s3-website-us-east-1.amazonaws.com/docs/807/817331/72221A_Fast_Food_Restaurants_in_the_US_industry_report.pdf

Shalvi, S., Dana, J., Handgraaf, M. J. J., & De Dreu, C. K. W. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, *115*(2), 181–190.

Simons, T. (2002). Behavioral Integrity: The perceived alignment between managers' words and deeds as a research focus. *Organization Science*, *13*(1), 18–35.

https://doi.org/10.1287/orsc.13.1.18.543

Simons, T., Friedman, R., Liu, L. A., & McLean Parks, J. (2007). Racial differences in sensitivity to behavioral integrity: Attitudinal consequences, in-group effects, and "trickle down" among Black and non-Black employees. *Journal of Applied Psychology*, *92*(3), 650–665. https://doi.org/10.1037/0021-9010.92.3.650

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*(6), 845–851. https://doi.org/10.1037/0022-3514.89.6.845

Stearns, D. C., & Parrott, W. G. (2012). When feeling bad makes you look good: Guilt, shame, and person perception. *Cognition & Emotion*, *26*(3), 407–430.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in experimental social psychology* (Vol. 21, pp. 261–302). Elsevier.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*(4), 411–419. https://doi.org/10.1080/10705519809540115

Stoll, M. L. (2002). The Ethics of Marketing Good Corporate Conduct. *Journal of Business Ethics*, *41*(1), 121–129. https://doi.org/10.1023/A:1021306407656

Stone, J., & Fernandez, N. C. (2008). To practice what we preach: The use of hypocrisy and cognitive dissonance to motivate behavior change. *Social and Personality Psychology Compass*, *2*(2), 1024–1051. https://doi.org/10.1111/j.1751-9004.2008.00088.x

Stone, J., Wiegand, A. W., Cooper, J., & Aronson, E. (1997). When exemplification fails: hypocrisy and the motive for self-integrity. *Journal of Personality and Social Psychology*,

72(1), 54–65.

Szabados, B., & Soifer, E. (2004). *Hypocrisy: Ethical Investigations*. Peterborough, Ontario, Canada: Broadview Press.

Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt*. New York: Guilford Press.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*(1), 345–372. https://doi.org/10.1146/annurev.psych.56.091103.070145

Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, *26*(8), 685–695. https://doi.org/10.1037/h0032110

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853–870. https://doi.org/10.1037/0022-3514.78.5.853

Valdesolo, P., & DeSteno, D. (2007). Moral hypocrisy. *Psychological Science*, *18*(8), 689–690. https://doi.org/10.1111/j.1467-9280.2007.01961.x

Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology*, *44*(5), 1334–1338.

Van Zee, A. (2009). The promotion and marketing of oxycontin: commercial triumph, public health tragedy. *American Journal of Public Health*, *99*(2), 221–227.

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

Walker, T. (2015). Ashley Madison hack: Website founder Noel Biderman is married father of two who says he has never cheated on his wife. *The Independent*. Retrieved from https://www.independent.co.uk/life-style/gadgets-and-tech/news/ashley-madison-hack-website-founder-noel-biderman-is-married-father-of-two-who-says-he-has-never-10466406.html

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*.

Xu, H., Bègue, L., & Shankland, R. (2011). Guilt and guiltlessness: An integrative review. *Social and Personality Psychology Compass*, *5*(7), 440–457. https://doi.org/10.1111/j.1751-9004.2011.00364.x

Zhong, C.-B., Ku, G., Lount, R. B., & Murnighan, J. K. (2010). Compensatory ethics. *Journal of Business Ethics*, *92*(3), 323–339. https://doi.org/10.1007/s10551-009-0161-6

Zhong, C.-B., & Liljenquist, K. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, *313*(5792), 1451–1452.

Zhong, C.-B., Liljenquist, K., & Cain, D. M. (2009). Moral self-regulation: LIcensing and compensation. In D. De Cremer (Ed.), *Psychological Perspectives on Ethical Behavior and Decision Making* (pp. 75–89). Charlotte, N.C.: Information Age.

Table 1. Summary of Predicted Tests.

Effect	Study	Test Statistic (t or z)*	*p-value*	*Cohen's d*
Total effect of Condition on Hypocrisy Judgments	1	3.93	<.001	.45
	2	3.13	.002	.50
	3	4.07	<.001	.37
	4	3.51	.001	.35
	5	6.28	<.001	.48
	S1	4.38	<.001	.46
	S2	4.50	<.001	.41
		b	***95% CI***	
Indirect effect of Condition on Praise through Hypocrisy	1	-.29 (.09)	[-.44, -.15]	
	2	-.12 (.04)	[-.22, -.05]	
	4	-.15 (.05)	[-.24, -.07]	
	5	-.16 (.02)	[-.20, -.12]	
	S1	-.17 (.04)	[-.25, -.09]	
	S2	-.18 (.04)	[-.26, -.10]	
Indirect effect of Condition on Hypocrisy through Guilt	4	.33 (.05)	[.24, .43]	
	S1	.20 (.04)	[.13, .29]	
	S2	.12 (.03)	[.07, .18]	
Serial indirect effect of Condition on Praise through Guilt and Hypocrisy	4	-.19 (.03)	[-.26, -.13]	
	S1	-.10 (.02)	[-.14, -.06]	
	S2	-.07 (.02)	[-.11, -.04]	
Serial indirect effect of Condition on Praise through Undeserved Benefits and Hypocrisy	5	-.12 (.01)	[-.15, -.10]	

Note: Standard errors reported in parentheses. For clarity in this table, Study 5 comparisons depict the comparison of means between the successful moral cleansing (i.e., guilt-relief) condition and control condition. Because Supplementary Study S1 included two measures of praise (hypocrisy of work and of charitable behavior), we report the measure about hypocrisy judgments for the target's charitable behaviors. *The test statistic is z in Study 3 because our analysis was a mixed model, and t in all other studies.

Table 2. Means and Standard Deviations of Focal Measures, Studies 1-4, S1-S2.

		Condition		**Total Effect of Condition**		
Measure	**Study**	**Inconsistent Virtue**	**Unrelated Virtue**[1]	**Test Statistic (*t* or *z*)**[2]	***p-value***	***Cohen's d***
Hypocrisy	1	5.23 (1.42)	4.62 (1.31)	3.93	<.001	0.45
	2	3.64 (1.16)	3.08 (1.06)	3.13	.002	0.50
	3	3.76 (1.40)	3.23 (1.41)	4.07	<.001	0.37
	4	3.86 (1.62)	3.30 (1.60)	3.51	.001	0.35
	S1	3.14 (1.61)	2.43 (1.45)	4.38	<.001	0.46
	S2	3.68 (1.52)	3.07 (1.43)	4.50	<.001	0.41
Perceived Guilt	4	4.78 (1.57)	3.48 (1.73)	7.95	<.001	0.79
	S1	4.72 (1.62)	3.47 (1.61)	7.40	<.001	0.78
	S2	5.02 (1.43)	4.25 (1.71)	5.37	<.001	0.49
Praise	1	3.10 (1.04)	2.52 (0.92)	5.18	<.001	–0.59[2]
	2	5.10 (1.00)	5.46 (0.71)	2.57	.01	0.41
	4	5.18 (1.24)	5.39 (1.17)	1.77	.077	0.18
	S1	5.37 (1.14)	5.58 (1.07)	1.72	.087	0.18
	S2	5.08 (1.18)	5.21 (1.08)	1.28	.202	0.11

Note: Standard Deviations reported in parentheses.

[1] In Study 1, this condition was a no-virtue control condition. The control condition provided no information about the target person's private virtues; participants only judged him based on his less-virtuous public persona. In all other studies, the control was an "unrelated virtue" condition in which the target person performed private virtues that were unrelated to his less-virtuous public work.

[2] The test statistic is z in Study 3 because our analysis was a mixed model, and t in all other studies.

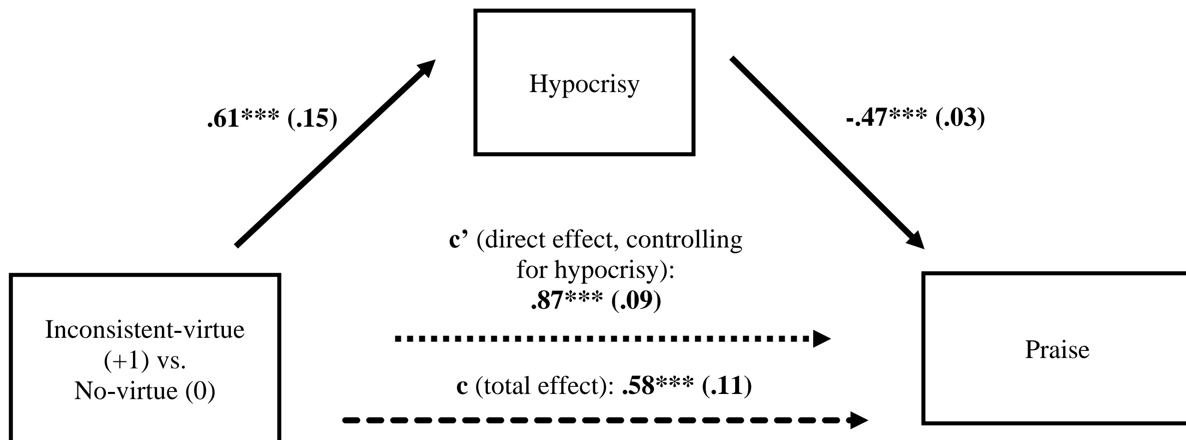
[3] The direction of Study 1's effect is different than the other studies' effect because Study 1 used a different control condition. See note 1 to this table.

Table 3. Means and Standard Deviations of Focal Measures, Study 5.

	Condition		
Measure	**Guilt-Relief**	**Guilt-Persists**	**Control**
Hypocrisy	5.33^{a} (1.19)	4.48^{b} (1.14)	4.76^{c} (1.18)
Guilt (Manipulation Check)	2.18^{a} (1.05)	5.09^{b} (1.09)	2.38^{c} (0.91)
Feels Better Than He Deserves	6.12^{a} (1.09)	3.36^{b} (1.48)	5.28^{c} (1.17)
Praise	2.53^{a} (1.12)	3.37^{b} (1.12)	3.21^{b} (1.25)

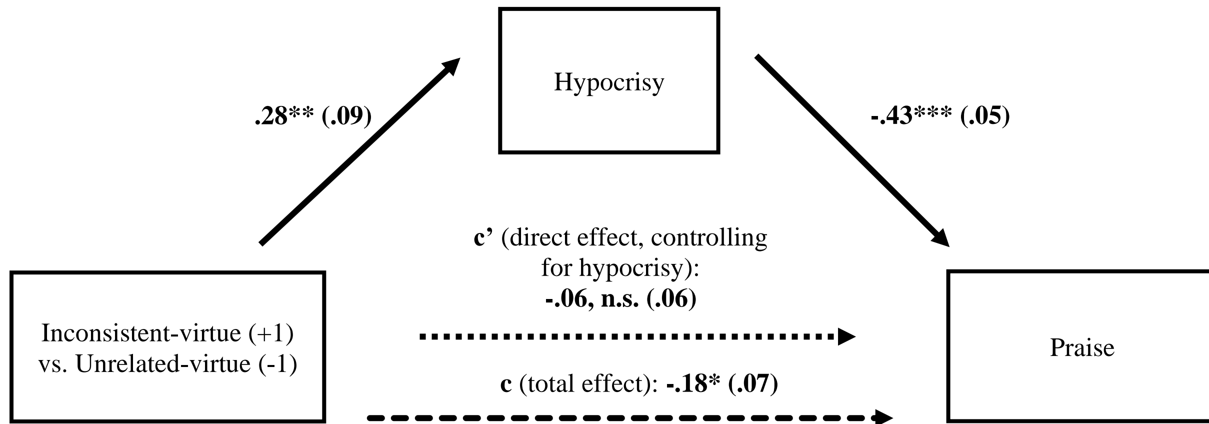
Note: Standard Deviations reported in parentheses. Study 5 means with different superscripts indicate significantly different means.

Figure 1. Hypocrisy judgments mediate the relationship between condition and praise, Study 1.



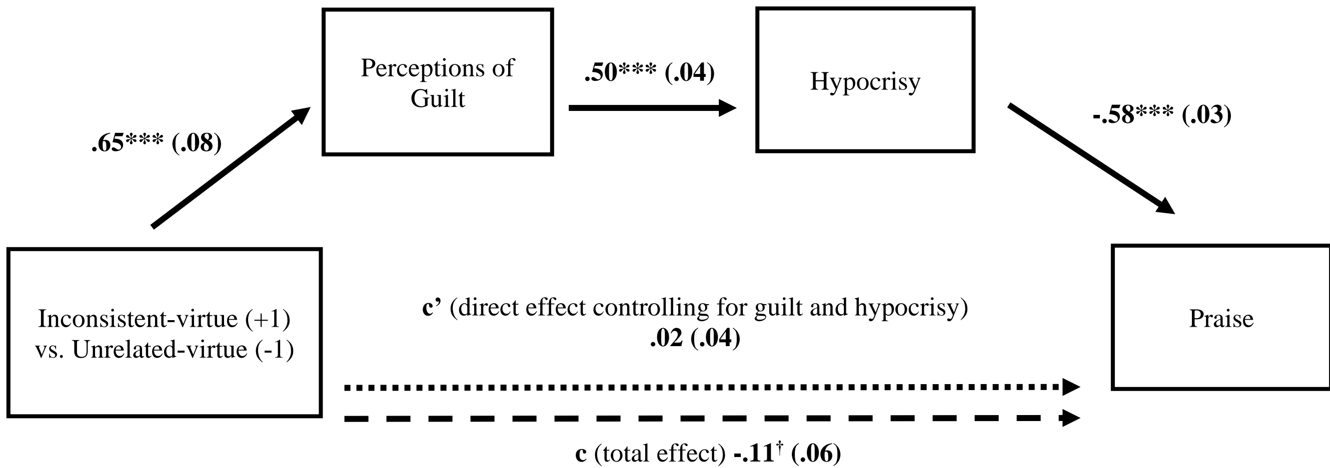
Note: Values are unstandardized path coefficients. **p* < .05, ***p* < .01, ****p* < .001. The indirect effect of condition on praise was significantly mediated by hypocrisy as indicated by the bias-corrected bootstrapped 95% CI around the indirect effect that did not include 0, *b* = -.29, *SE* = .09, [-.44, -.15], using 5,000 resamples.

Figure 2. Perceptions of hypocrisy mediate the relationship between condition (inconsistent- vs unrelated-virtue) and praise, Study 2.



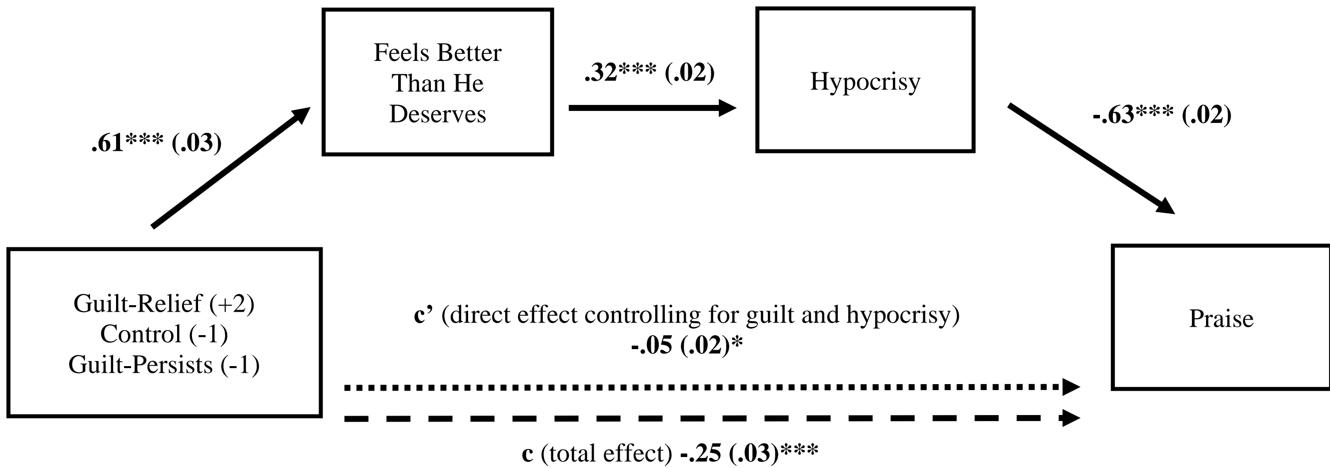
Note: Values are unstandardized path coefficients. **p* < .05, ***p* < .01, ****p* < .001. The indirect effect of condition on praise was significantly mediated by hypocrisy as indicated by the bias-corrected bootstrapped 95% CI around the indirect effect that did not include 0, *b* = -.12, *SE* = .04, [-.22, -.05], using 5,000 resamples. Robustness checks reported in the Online Supplement show that this effect remains significant even when controlling for alternative indirect effects through perceptions of harm to the organization.

Figure 3. Serial mediation results show the significant indirect path from condition to perceptions of guilt, hypocrisy, and praise, in Study 4.



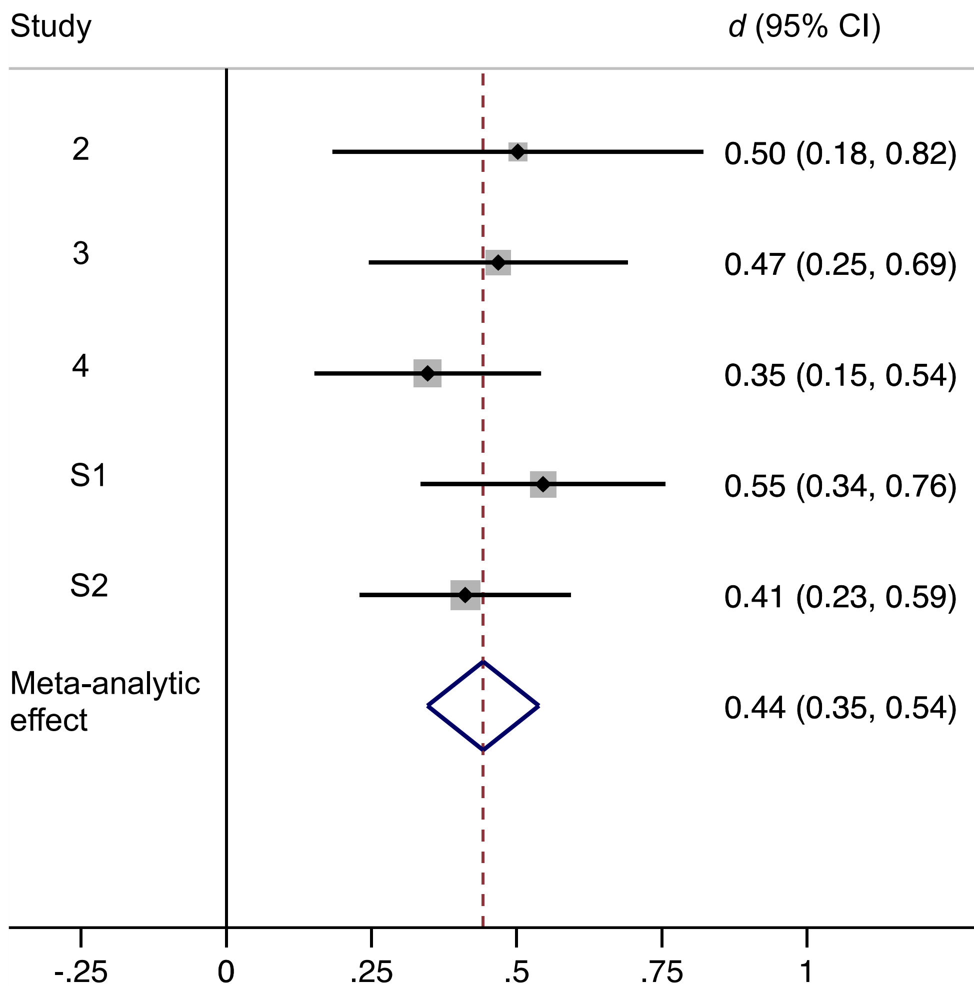
Note: Values are unstandardized path coefficients. Standard errors reported in parenthesis. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. For clarity, alternative non-significant indirect effects among guilt, hypocrisy, and praise are not depicted in this figure. The bold lines show the significant serial indirect effect, $b = -.19$, $SE = .03$, bias-corrected 95% CI [-.26, -.13].

Figure 4. Serial mediation results show the significant indirect path from condition to perceptions of undeserved guilt-reduction, hypocrisy, and praise, in Study 5.



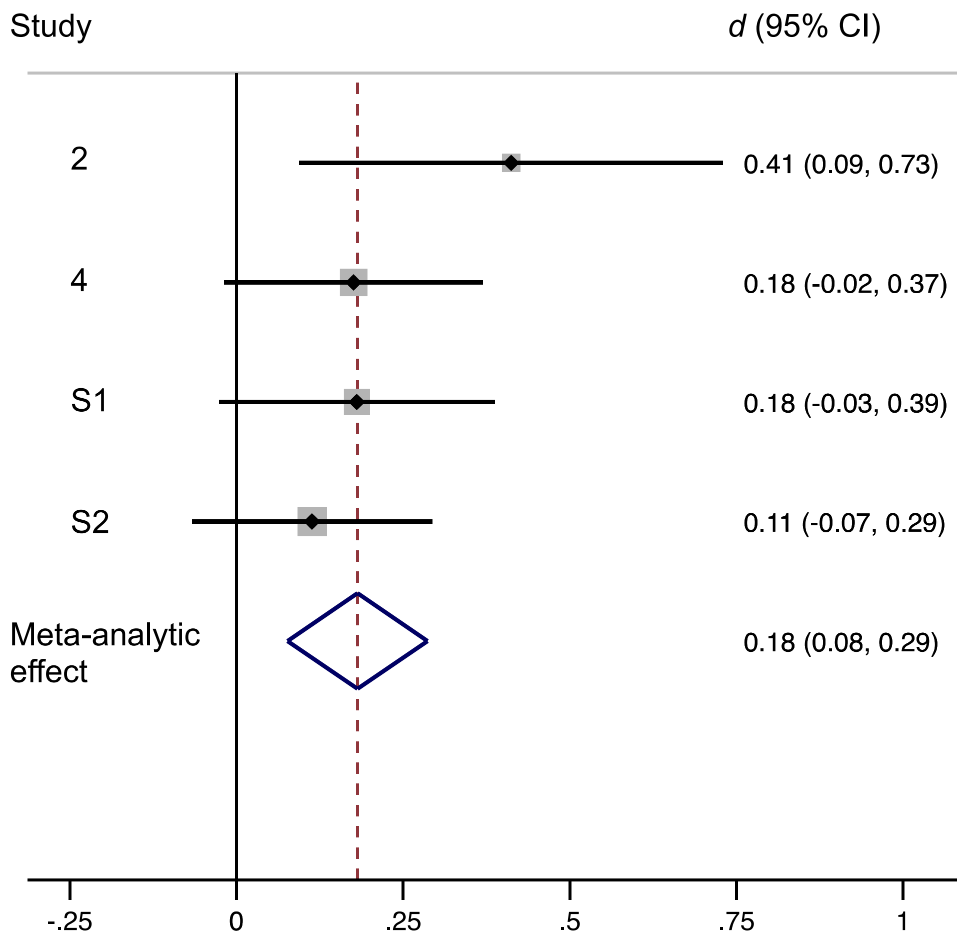
Note: Values are unstandardized path coefficients. Standard errors reported in parenthesis. *p < .05, **p < .01, ***p < .001. For clarity, like Figure 3, this figure does not depict alternative non-significant alternative indirect effects among perceptions that he feels better than he deserves, hypocrisy, and praise. The bold lines show the significant serial indirect effect, b = -.12, SE = .01, bias-corrected 95% CI [-.15, -.10].

Figure 5. Meta-Analysis of the Effect of Inconsistency on Hypocrisy



Note. Studies S1 and S2 are the two supplemental studies reported in the Online Supplement. The larger the effect size, the greater the tendency to perceive more hypocrisy in the inconsistent-virtue condition than in the unrelated-virtue condition. The size of the gray square is proportional to the sample size. CI = confidence interval. The analysis omits Studies 1 and 5 because they did not include a separate unrelated-virtue condition.

Figure 6. Meta-Analysis of Effect of Inconsistency on Praise



Note. Studies S1 and S2 are the two supplemental studies reported in the Online Supplement. The larger the effect size, the greater the tendency to express less praise in the inconsistent-virtue condition than in the unrelated-virtue condition. The size of the gray square is proportional to the sample size. CI = confidence interval. The analysis omits Study 3 because it did not assess praise, and Studies 1 and 5 because they did not include a separate unrelated-virtue condition.