Letting People Off the Hook: When Do Good Deeds Excuse Transgressions?

Daniel A. Effron

Benoît Monin

Stanford University

Abstract

Three studies examined when and why an actor's prior good deeds make observers more willing to excuse – or *license* – his or her subsequent, morally dubious behavior. In a pilot study, actors' good deeds made participants more forgiving of the actors' subsequent transgressions. In Study 1, participants only licensed blatant transgressions that were in a different domain than actors' good deeds; blatant transgressions in the same domain appeared hypocritical and suppressed licensing (e.g., fighting adolescent drug use excused sexual harassment, but fighting sexual harassment did not). Study 2 replicated these effects, and showed that good deeds made observers license *ambiguous* transgressions (e.g., behavior that might or might not represent sexual harassment) regardless of whether the good deeds and the transgression were in the same or in a different domain – but only same-domain good deeds did so by changing participants' construal of the transgressions. Discussion integrates two models of why licensing occurs.

Letting People Off the Hook:

When Do Good Deeds Excuse Transgressions?

Consider how observers might react to the allegation that Martin Luther King, Jr. committed adultery (Abernathy, 1989). In light of King's exemplary contributions to the civil rights movement, observers might be at least somewhat inclined to excuse this alleged transgression. By contrast, consider reactions to Eliot Spitzer, who as New York State district attorney distinguished himself by fighting prostitution and sex trafficking but who later, as governor, hired prostitutes himself. Observers accused Spitzer of hypocrisy, and showed little inclination to excuse his transgression despite his undeniable previous contributions to the fight against prostitution (e.g., Hakim & Santos, 2008). As these two examples illustrate, observers are only sometimes willing to excuse – or *license* – transgressions committed by actors who have a history of moral behavior. Licensing occurs when observers reduce their condemnation of morally dubious behavior or the actor who committed it in light of the actor's prior good deeds. In the present paper, we address the questions of when and why such licensing occurs.

When Do Observers License Transgressions?

Of the many differences between King and Spitzer, one seems especially germane to answering the question of when licensing occurs: while Spitzer's good and bad deeds were in the same domain (prostitution), King's good and (alleged) bad deeds were in different domains (civil rights vs. marital fidelity). When an actor's behavior clearly represents a moral violation – what we refer to as a *blatant transgression* – prior good deeds in a different domain might reduce observers' condemnation by seeming to balance out the bad behavior, while prior good deeds in the same domain might make the actor appear hypocritical. Based on this reasoning, we

hypothesize that observers will license blatant transgressions that are preceded by different-domain, but not same-domain, good deeds.

We further propose that the ambiguity of a transgression is another important moderator of licensing. We define *ambiguous transgressions*, in contrast to blatant transgressions, as suspicious behaviors that could, but need not, represent moral violations. Consider, for example, a White employer who promotes several White employees but passes over two Black employees. Without more information, the promotion decision might seem suspicious: it could represent racial discrimination, or it could represent a good-faith effort to promote the most qualified candidates, who happen in this case to be White. Rather than appearing hypocritical, good deeds in the same domain as this sort of ambiguous transgression (e.g., a history of fighting racial discrimination) might render the "good faith" interpretation more plausible than the "racism" interpretation. Good deeds in a different domain (e.g., a history of fighting prostitution) would seem less relevant for disambiguating the suspicious behavior, but might nonetheless seem to balance it out. Based on this analysis, we hypothesize that ambiguous transgressions (unlike blatant ones) are licensed by both same-domain and by different-domain good deeds.

In the following sections, we develop our predictions about *when* observers will license transgressions by reviewing and integrating research on *why* observers would or would not be willing to license in the first place.

<div align="center">*Why* Do Observers License Transgressions?</div>

Licensing could occur for at least two reasons (see Merritt, Effron, & Monin, 2010; Miller & Effron, in press). We address each in turn.

*Licensing Via Balance: The Moral Credits Model*

Observers may behave as if actors had a moral bank account.  Good deeds would represent *moral credits* to the account, while bad deeds would represent moral debits.  In this model of licensing, actors are permitted to make a moral debit by committing a transgression so long as they have previously accumulated sufficient moral credits to balance out the wrongdoing (Nisan, 1991; see also Hollander, 1958).  In this way, moral credits "purchase" a license to transgress without incurring observer condemnation.  For example, when an actor has obtained moral credits, observers should be less likely to say that it is immoral or inexcusable for him to commit sexual harassment.  Researchers have proposed similar models to explain why people are more likely to license themselves to transgress after they have behaved morally (Jordan, Mullen, & Murnighan, 2009; Sachdeva, Iliev, & Medin, 2009; Zhong, Liljenquist, & Cain, 2009).

*Licensing Via Construal: The Moral Credentials Model*

A second reason why licensing could occur is that an actor's good deeds might change the way observers construe her subsequent, ambiguous behavior.  In this model, good deeds do not grant one a license to transgress per se, but rather license morally dubious behavior by making it seem as if it were not a transgression at all.  Monin and Miller (2001) proposed this model in a series of studies examining actors' willingness to license themselves.  For example, giving male participants the chance to disagree with sexist statements made these participants more likely to describe a stereotypically male job as better suited for men than for women.  Monin and Miller argued that disagreeing with the statements gave participants *moral credentials* as non-sexists, thus increasing participants' confidence that favoring a man for the job would not appear sexist (and instead be construed as, e.g., arising from concern that other employees' sexism would undermine a woman's performance; see also Effron, Cameron, & Monin, 2009).  A moral *credits* interpretation, by contrast, would claim that these participants

felt more comfortable expressing a preference that *would* appear sexist.  If moral credits are like a deposit that allows one to "purchase" the right to transgress, moral credentials provide a lens through which subsequent behavior is construed more favorably.

Some evidence, consistent with the moral credentials model, suggests that an actor's good deeds can change the way observers construe his or her subsequent, ambiguous transgressions.  In one study (Krumm & Corning, 2008), heterosexual participants read vignettes in which targets engaged in several ambiguously discriminatory behaviors toward gays and lesbians (e.g., denying a bank loan to a gay couple), either after establishing moral credentials (e.g., marching in support of gay rights) or not.  Participants perceived behaviors performed by credentialed targets as less discriminatory than different behaviors that were performed by non-credentialed targets.

*When Does Licensing Occur Via Balance vs. Construal?*

Our theorizing about these two models of licensing continues a long tradition of research on person perception and impression formation, which has often considered how observers respond to inconsistent information about others (see Hastorf, Schneider, & Polefka, 1970; Jones, 1990).  The moral credits (balance) model relates to work on how observers average positive and negative characteristics of an actor to create a general impression (e.g., Anderson, 1965), whereas the moral credentials (construal) model fits with theories about how prior information about an actor shapes observers' interpretation of subsequent information about her (e.g., Asch, 1946).  The present model aims to contribute to this literature by applying these familiar processes to the phenomenon of moral licensing and specifying the conditions under which each process will operate.  In so doing, we hope to shed light on when and why licensing will occur.  We view these processes as complementary; as we develop below, we propose that

both the credits model and credentials model accurately describe why licensing occurs, but in different situations. To understand when each model accurately explains licensing, however, it is first necessary to consider why observers may be reluctant to excuse transgressions.

<center>Why Would Observers Be *Reluctant* to License? The Role of Hypocrisy</center>

Moral behavior (especially when it involves efforts to influence others) makes an implicit claim about one's values. Observers seem to have interpreted Spitzer's fight to reduce prostitution, for example, as a claim that he personally opposed prostitution. Transgressions that are in the same domain as prior moral behavior can thus create the appearance of making, and then contradicting, a claim about one's values – in other words, "saying one thing, but doing another," the defining feature of hypocrisy (Barden, Rucker, & Petty, 2005). Hypocrisy elicits negative attributions about the prior moral behavior, which as a result seems less sincere or more selfish and calculated. People also dislike hypocrites because of a general distaste for inconsistent behavior (Cialdini, 1988; Stone, Wiegand, Cooper, & Aronson, 1997; Tedeschi, 1971). It thus seems likely that blatant transgressions that are in the same domain as prior moral behavior could elicit perceptions of hypocrisy, thus making observers reluctant to license. Transgressions in a different domain should likely not appear hypocritical, because they are not directly inconsistent with prior good deeds.

In one study (Powell & Smith, 2009), participants read about a target person who had been punished for academic dishonesty. When the target had previously condemned and helped punish other academically dishonest students, compared to when he had not, participants rated him as more hypocritical and more deserving of the punishment, and expressed greater *schadenfreude* (i.e., pleasure about his negative outcome). Fitting with the idea that moral behavior makes only same-domain transgressions appear hypocritical, *schadenfreude* and

perceptions of hypocrisy were reduced when the target had committed a transgression (stealing) that was unrelated to his moral claims about academic dishonesty.

Theory and research on hypocrisy thus illustrate the potential risks of committing a transgression in the same domain in which one has previously performed good deeds: such situations may not only make observers more reluctant to license, but may also make them happier to dole out punishment.  On the other hand, the moral credentials model suggests that same-domain good deeds should have the potential to facilitate licensing by improving observers' construal of a subsequent transgression, as in the example in which a history of fighting racial discrimination made it more plausible that a hiring decision was based on legitimate, and not racist, motives.  What determines whether prior good deeds provide license versus appear hypocritical?

Integrating Theories of Moral Credits, Moral Credentials, and Hypocrisy

We propose that the effects of good deeds on observers' reactions to subsequent transgressions can best be understood by conceptualizing the ascription of moral credentials and moral credits as complementary processes that promote licensing, and the ascription of hypocrisy as a process that opposes licensing.  By considering how each process is differentially affected by the transgression's ambiguity and domain, our model predicts when licensing should occur and when it should not.  These predictions are summarized in the top half of Table 1.

First, consider the case of blatant transgressions (first two columns in Table 1).  Because by definition such behaviors are not open to favorable reinterpretation, good deeds cannot promote licensing by changing construal.  Good deeds could, however, balance out such transgressions, thus promoting licensing as described in the moral credits model.  Perceived hypocrisy would counteract licensing only when the transgression is in the same domain as the

good deeds. Thus, when transgressions are blatant, we expect licensing to occur via balance (moral credits), and to occur only when the good deeds are in a different domain.

Now consider ambiguous transgressions (third and fourth columns in Table 1). Unlike blatant transgressions, ambiguous transgressions are open to favorable reinterpretation, so good deeds could promote licensing by changing construal, as described by the moral credentials model. Hypocrisy should not undermine licensing because ambiguous transgressions that are construed favorably would not contradict prior good deeds. Same-domain good deeds (fourth column) are most likely to produce licensing via construal, because they are more relevant than different-domain good deeds for disambiguating the transgression. Yet different-domain good deeds (third column) could still produce licensing via balance. Thus, when transgressions are ambiguous, we predict that both same- and different-domain good deeds will have a licensing effect, but for different reasons.

Overview of Studies

We tested these predictions in three studies in which participants read about target individuals who had performed morally dubious acts. A Pilot Study sought to establish that prior good deeds can license transgressions in the eyes of observers (Table 1, first column). Study 1 attempted to show that blatant transgressions are only licensed by good deeds that are in a different domain (first and second column). Study 2 added a manipulation of the transgression's ambiguity in an effort to replicate the hypothesized effect in Study 1 for blatant transgressions, and to show that both same- and different-domain good deeds license ambiguous transgressions but for different reasons (all four columns). In all studies, we operationalized good deeds as actions taken to protect or promote others' welfare (i.e., volunteering to help the homeless,

sheltering hurricane victims, implementing progressive hiring practices, combating sexual harassment, or reducing adolescent drug use).

PILOT STUDY

The Pilot Study tested the hypothesis that good deeds in a different domain than a blatant transgression would produce a licensing effect (see Table 1, first column), defined as a reduction in observer condemnation. Although prior research had established that observers integrate information about an actor's good and bad deeds when judging his moral character (Birnbaum, 1973; Riskey & Birnbaum, 1974), we wished to demonstrate that an actor's prior good deeds would additionally affect how permissible observers found his subsequent transgression. We thus assessed evaluations of permissibility as part of our measure of condemnation.

One hundred thirty-eight participants read two vignettes about target individuals who had committed different transgressions (i.e., a man who drove away without leaving a note after accidentally hitting a parked car, and a woman who committed marital infidelity with a stranger). Prior to reading about each of these transgressions, approximately half the participants also read about good deeds that the same targets had previously performed (i.e., sheltering hurricane victims or helping the homeless). As dependent measures, participants evaluated each target person (how much they would like him or her as a coworker, as a friend, how much they respected him or her, and how moral they found him or her; these items all reverse-coded) and indicated how permissible they found the transgression (how much the target should be blamed, how wrong the transgression was, how much the target deserved be punished, how excusable the transgression was, and how immoral it was) by responding to nine items (endpoints: 1 = "not at all", 7 = "very much"). These nine items were highly intercorrelated ($\alpha$ = .91), so we averaged them to form a *condemnation* scale.

As expected, across both vignettes participants expressed less condemnation when the targets had performed good deeds than when they had not, ($M$s = 4.97 vs. 5.81, $SD$s = 1.24 vs. 1.08, respectively), $F(1, 136) = 27.67$, $p < .0001$ in a good deeds (between-subjects) by vignette (within-subjects) ANOVA that also revealed an unpredicted main effect of vignette but no significant interaction. This finding establishes that good deeds can license subsequent transgressions that are in a different domain, and is consistent with prior work on person perception (Birnbaum, 1973; Riskey & Birnbaum, 1974). We next tested our model's predictions about when and why such licensing would and would not occur.

## STUDY 1

Study 1 tested our prediction that observers would license blatant transgressions based on different-domain good deeds, but not based on same-domain good deeds (Table 1, first and second columns). Because we predicted a different pattern of results for ambiguous and blatant transgressions, we strove in Study 1 to create vignettes in which the target had clearly done something wrong. To enhance our stimuli's realism, we had participants read (fake) newspaper articles that described the target's good and bad deeds in detail.

We expected that the apparent hypocrisy of committing a blatant transgression in the same domain as prior good deeds would suppress licensing. Thus, we predicted that a target's good deeds would reduce condemnation when they were in a different domain than the subsequent transgression (as in the Pilot Study), but not when they were in the same domain. As a test of mechanism, we measured perceptions of the target's hypocrisy, hypothesizing that this measure would suppress licensing in the same-domain condition, but not in the different-domain condition.

## Method

*Participants*

Eighty-three members of a university subject pool were paid $16/hour to complete Study 1 in a packet of unrelated surveys or at the end of an unrelated laboratory experiment. Seven participants were excluded because they had already completed a pilot version of the study. The remaining 76 participants (36 females, 31 males, 9 of unknown gender; 38% White, 29% Asian, 7% Black, 7% Hispanic/Latino, 3% multiracial, remainder other race or unknown) were on average 20.72 years old ($SD = 4.42$).

*Materials*

We created two articles about a high school principal, formatted to appear to have been downloaded from a local newspaper's website. The first article contained information about the principal's good deeds in one of two domains: reducing either drug use or sexual harassment among his students.

> James has tackled the issue of [drug use / *sexual harassment*] head-on by [keeping dealers out of the school, implementing a zero-tolerance policy for drug possession, and increasing the school's extracurricular budget to give students a healthy alternative to drug use / *educating students about how to identify, avoid, and respond to sexual harassment, and implementing a zero-tolerance policy for committing such harassment*].

The article further reported that the principal had begun his "personal crusade" against drugs or harassment before public awareness of the problem had emerged, and had worked "tirelessly for years" on the problem; that his efforts had resulted in a 20% drop in drug use or harassment; and that he had continued to fight drugs or harassment despite the risk that doing so would cause his boss, who wished to avoid public acknowledgement of the school's problems, to fire him.

The second article, ostensibly published eight days after the first article in the same newspaper, reported that the same principal had been either "arrested on charges of drug possession" or "accused of sexual harassment." These two articles enabled us to vary whether or

not the transgression was in the same domain as the prior good deeds. The article about the

drug-related transgression contained the following information:

> According to police reports filed on Wednesday, an officer pulled over James' car for speeding. James appeared intoxicated or high, and the officer asked him to step out of the car. At the officer's request, James emptied his pockets, revealing a bag containing white powder. Police later confirmed that the bag contained approximately 3 grams of cocaine.

The article about the harassment-related transgression contained the following information

instead:

> Carolyn Scherer, a 33-year-old waitress at a local restaurant, came forward on Wednesday with allegations that James, who frequents the restaurant where Scherer works, had made unwelcome sexually explicit remarks to her, offered her money to perform sexual acts with him, and touched her on the buttocks.

Both articles made it clear that the principal had committed the transgression: for the

drug-related transgression, "Police later confirmed that the bag contained approximately 3 grams

of cocaine;" for the harassment-related transgression, "Several customers interviewed for this

article were able to verify Scherer's allegations;" and in both articles, the target indicated that he

would not dispute the charges against him.

*Procedure*

In a survey entitled, "Understanding and Interpreting the News," participants responded

to filler questions (e.g., "What is your primary source of news?"), and were then randomly

assigned to read about one of the two transgressions (drugs vs. harassment). The main

experimental manipulation came just before the transgression article: participants read about the

principal's good deeds in the *same domain* as his transgression (e.g., he fought, then committed,

harassment) or in a *different-domain* (e.g., he fought drug use, then committed harassment); in

the *control* condition, participants did not read about any good deeds. Participants then

responded to filler items about the article (e.g., how well it was written) and the principal's

competence (i.e., his intelligence, strength, and confidence), and then responded to our dependent measures.

## Measures

*Condemnation*

Fifteen items assessed participants' condemnation of the target and the transgression. First, participants evaluated the target on nine 7-point semantic differentials, anchored at -3 and 3 (starred items were reverse-coded): cruel/kind,* nice/awful, cold/warm,* honest/dishonest, unfair/fair,* moral/immoral, arrogant/humble,* good/bad, likeable/dislikeable. Participants also indicated how permissible they found the transgression (referred to as "drug possession" or "behavior towards the waitress") by evaluating it on three semantic differentials (honorable/dishonorable, moral/immoral, excusable/inexcusable; endpoints: -4 and 4; 0 = "neither"), judging how much blame the target deserved (1 = "none at all", 7 = "very much"), and indicating their agreement that the target should "resign from his job," and that other "schools should definitely not hire him" (-3 = "strongly disagree," 0 = "neutral/unsure," 3 = "strongly agree"). These 15 items were highly intercorrelated, so after standardizing them to account for the different response scales, we averaged them into a single condemnation composite, with higher scores indicating greater condemnation ($\alpha$ = .86).

*Hypocrisy*

After evaluating the target but before evaluating the transgression's permissibility, participants indicated whether or not the target "is a hypocrite" (-3 = "strongly disagree," 0 = "neutral/unsure," 3 = "strongly agree).

*Evaluation of Good Deeds*

After evaluating the target but before responding to the hypocrisy item, participants who had read about the target's good deeds (i.e., everyone not in the control condition) rated the actions he had taken "to crusade against (drug use/sexual harassment)" on three bipolar scales, anchored at -4 and 4 (honorable/dishonorable[*], moral/immoral[*], insincere/sincere; starred items reverse-coded) and estimated how much he cared about and was committed to reducing either drug use or sexual harassment (1 = "not at all, " 7 = "very much"). We standardized and averaged these four items so that higher numbers indicated more positive evaluations ($\alpha = .82$).

## Results and Discussion

Preliminary analyses revealed that our results were not qualified by gender, race, or the specific domain (i.e., drugs vs. harassment) of the first article or the second article, nor did any main effects of these variables emerge.

### *Licensing Only by Different-Domain Good Deeds*

As in the Pilot Study, we defined licensing as a decrease in condemnation when the target had performed good deeds compared to when he had not, and we expected to observe this decrease in the different-domain condition, but not in the same-domain condition. A one-way ANOVA established that condemnation was not equivalent across the three conditions, $F(2, 73) = 3.13$, $p < .05$ (see Figure 1). Planned, non-orthogonal contrasts confirmed our specific predictions: participants expressed less condemnation in the different-domain condition ($M = -.21$, $SD = .61$) than in the control condition ($M = .13$, $SD = .59$), $F(1, 73) = 4.03$, $p < .05$, but were just as condemning in the same-domain condition ($M = .13$, $SD = .48$) as in the control condition, $F(1, 73) < .01$, $ns$. These results demonstrate that, as predicted, blatant transgressions were licensed by prior good deeds in a different domain, but not in the same domain.

### *Suppression by Hypocrisy of Same-Domain Licensing*

We hypothesized that when good deeds were in the same domain as the transgression, perceptions of the target's hypocrisy would suppress the licensing effect observed in the different-domain condition. Statistical suppression occurs when controlling for a suppressor variable (in this case, hypocrisy) significantly strengthens the relationship between an independent and dependent variable (whereas mediation would occur if controlling for hypocrisy weakened this relationship; MacKinnon, Krull, & Lockwood, 2000). As described by MacKinnon et al., assessing suppression requires evaluating three regression equations, identical to those in the more familiar analysis of mediation (Baron & Kenny, 1986). A first regression equation – using a contrast that coded the same-domain condition as 1, the control condition as - 1, and the different-domain condition as 0 – confirmed that the target appeared more hypocritical in the same-domain condition than in the control condition, $b = .55$, $t(73) = 4.56$, $p < .001$ (path $a$ in Figure 2). A second equation regressing condemnation against this contrast indicated no same-domain licensing effect overall, $b = -.01$, $t(73) = .13$, $ns$ (path $c$ in Figure 2). Third, if hypocrisy is indeed a suppressor, entering it into this second equation should significantly strengthen this same-domain licensing effect. Results showed that entering hypocrisy – itself a significant predictor, $b = .35$, $t(72) = 3.91$, $p < .001$ (path $b$) – did strengthen the same-domain licensing effect to marginal significance, $b = -.15$, $t(73) = 1.72$, $p = .09$ (path $c'$), and this increase was significant by the Sobel test, $z = 2.97$, $p < .005$. We did not observe a similar suppression effect in the different-domain condition; controlling for hypocrisy *decreased* different-domain licensing from $b = -.18$ to $b = -.13$, although not significantly so, $z = 1.45$, $p = .15$. These results suggest that same-domain good deeds did improve reactions to the principal and his transgression, but that this positive consequence was counteracted by perceptions of

hypocrisy, resulting in no licensing effect overall. In statistical terms, hypocrisy suppressed the licensing effect of same-domain good deeds.

Additional analyses confirmed that licensing by same-domain good deeds was significantly weaker than licensing by different-domain good deeds – an effect fully explained by perceptions of greater hypocrisy in the same-domain condition.[i]

*Evaluation of Good Deeds*

Recall that only participants in the same- and different-domain conditions read about good deeds and were asked to evaluate them. As expected, participants responded less positively to the good deeds after reading about a same-domain transgression ($M = -.40$, $SD = .16$) than after reading about a different-domain transgression ($M = .39$, $SD = .08$), $t(55) = 4.55$, $p < .0001$.

*Summary*

Study 1 again demonstrated a licensing effect: participants responded to the same blatant transgression with less condemnation when they learned that the perpetrator had previously performed good deeds in another domain. Study 1 also added an important qualification to the effect: as predicted by our model, good deeds in the same domain as the transgression failed to produce licensing. Thus, for example, reducing adolescent drug use helped get the target person off the hook for committing sexual harassment, but not for possessing illegal drugs. Further analyses showed that the overall lack of licensing in the same-domain condition resulted from perceptions of hypocrisy, which suppressed an otherwise-modest licensing effect. In fact, the same good deeds appeared less honorable and sincere when they were in the same domain as a subsequent transgression compared to when they were in a different domain. Together, these results show that observers' willingness to license blatant transgressions depends on the similarity between the domains in which an actor transgresses and performs good deeds. The

results also raise the question of whether same-domain licensing ever occurs – an issue that we examine in Study 2.

## STUDY 2

It is important to determine if and when observers will license others based on same-domain good deeds, because actors seem to use same-domain good deeds to license themselves. Prior research shows that acting in unprejudiced ways makes individuals feel comfortable expressing attitudes that could seem prejudiced (Effron et al., 2009; Monin & Miller, 2001). Study 1 found that an actor's good deeds only prompted observer licensing when they were in a different domain than the actor's subsequent transgression, because same-domain good deeds made the actor appear hypocritical. Our model indeed predicts this pattern for blatant transgressions, but it also predicts that same-domain good deeds should license *ambiguous* transgressions – that is, suspicious behaviors that might or might not actually represent transgressions – by disambiguating them favorably for the actor (see Table 1). In Study 2, we tested these predictions by manipulating whether a target's suspicious behavior was ambiguous or represented a blatant transgression. As in Study 1, we also manipulated whether the target had performed good deeds in the same-domain as the transgression or in a different domain; a control condition omitted all information about good deeds. We predicted that different-domain good deeds would license both blatant and ambiguous transgressions, while same-domain good deeds would only license ambiguous ones. Study 2 thus tested all of the predictions derived from our model (shown in Table 1).

Study 2 also tested two psychological processes potentially driving the predicted effects. We hypothesized that same-domain good deeds would not license if they seemed hypocritical (as in Study 1), but would license if they made observers construe the transgression more favorably.

By contrast, we predicted that different-domain good deeds would license by balancing out bad deeds, but without changing how transgressions were construed. To examine these proposed mechanisms, we measured perceptions of hypocrisy and assessed how participants construed the (ambiguous or blatant) transgression.

## Method

### *Participants*

One hundred seventeen participants from a university subject pool received $8 to spend 30 minutes completing a packet of unrelated surveys. Three participants were excluded for having previously completed Study 1 or a related pilot study, and one was excluded for being a minor. Most of the remaining 113 participants were undergraduates ($M$ age = 20.55 years, $SD$ = 5.10; 73 females, 40 males; 31% White, 27% Asian, 12% Black, 10% Latino, 13% multiracial, and 7% other).

### *Materials*

Stimuli were passages from two articles, supposedly excerpted from a local magazine, about a manager at an area technology company. The first article described the manager's good deeds, crediting him with successfully either reducing incidents of sexual harassment or increasing the representation of racial minorities at his company. For example, in the harassment version,

> he lobbied the company's executives to institute anti-sexual-harassment policies, require sensitivity training for all employees, and establish a hotline that enabled employees to report sexual harassment and seek counseling and legal services,

while in the minority representation version, he

> overhauled the company's hiring system by proposing, and then instituting, an affirmative-action policy for racial minorities. Since then, he has designed and implemented several different programs aimed at recruiting minority applicants.

The second article, ostensibly published one month after the first, reported that the manager was being sued for either sexual harassment or racial discrimination. It was initially unclear whether or not the allegations were justified. In the harassment version of this article, the target (Hutchinson) had invited a female employee (Krasne) to dinner to discuss a possible promotion. Krasne later claimed that the target had

> "leered at [her] in a sexually suggestive manner, implying that going to dinner, and perhaps becoming intimate with Mr. Hutchinson, would greatly enhance her chances of promotion." Krasne declined the invitation to dinner; two weeks later, she was informed that she had not received the promotion.

In the discrimination version, the article reported that the target had to fill five positions by promoting his employees.

> Seven of his employees – five of whom were white, two of whom were black, and all of whom had been with the company for equal amounts of time – qualified for the positions. Hutchinson, who is himself white, decided to promote the five white employees.

By manipulating the text of the article's last paragraph, we created an *ambiguous-transgression* version, in which the target provides an alternative explanation for his behavior, and a *blatant-transgression* version, in which the article reveals evidence confirming his culpability and the target admits to the allegations (see Table 2 for the manipulation's text).

*Procedure*

Participants completed an "Impression Formation Study" that prompted them to read the article(s) "in order to form an impression" of the target person. We randomly assigned participants to read either the blatant or the ambiguous version of one of the transgressions. Orthogonally, we manipulated whether participants had previously read about the target's prior good deeds that were in the *same domain* as the alleged transgression (e.g., fought, then accused of, sexual harassment) or in a *different domain* (e.g., fought sexual harassment, then accused of racial discrimination); participants randomly assigned to a control condition did not read about

any prior good deeds. Before completing the primary dependent measures, participants

responded to the three filler items about the target (intelligence, strength, and confidence) used in

Study 1.

<p style="text-align:center"><em>Measures</em></p>

*Condemnation*

Participants first evaluated the target on the same nine semantic differentials used in

Study 1 (i.e., cruel/kind,* nice/awful, cold/warm,* honest/dishonest, unfair/fair,* moral/immoral,

arrogant/humble,* good/bad, likeable/dislikeable; starred items reverse-coded).  Participants also

indicated how permissible they found the transgression (referred to as the target's "decision

about whom to promote" or "behavior towards Rita Krasne" in the discrimination and

harassment vignettes, respectively) using similar items to those used in Study 1, with minor

changes intended to increase the scale's reliability.  These items were four semantic differentials

(perfectly OK/extremely immoral, not at all problematic/extremely problematic, not at all

appropriate/completely appropriate*, and perfectly fine/extremely wrong; endpoints: 1 and 7;

starred item reversed) and two statements (the target "deserves to be fired from his job" and that

"legal action should be taken" against him; -3 = "strongly disagree", 0 = "unsure/neutral", 3 =

"strongly agree").[ii]  As in Study 1, we standardized and then averaged all items to create a single

scale assessing condemnation ($\alpha = .93$).

*Construal*

After evaluating the target himself but before rating the transgression's permissibility,

participants indicated to what extent the transgression represented either racial discrimination or

sexual harassment, depending on what the target had been accused of; 1 = "definitely not (racial

discrimination / sexual harassment)", 7 = "definitely (racial discrimination / sexual harassment)."

*Hypocrisy*

We used the same item to measure hypocrisy that we used in Study 1. Hypocrisy was the last item measured in Study 2, in contrast to Study 1. We changed this placement to rule out the possibility that the suppression effect of hypocrisy observed in Study 1 was obtained because participants were prompted to think about hypocrisy before responding to the other items.

## Results

Gender, race, and the specific domain (i.e., racial discrimination vs. sexual harassment) in the first article or the second article did not moderate our results, but (in contrast to Study 1) the specific domain in the first article was a marginal or significant covariate in some analyses. To reduce error variance, we thus included this covariate in all analyses we report below.[iii]

### Primary Analyses

We predicted that same-domain and different-domain good deeds would have different effects on how much condemnation participants expressed, depending on whether the transgression was blatant or ambiguous. To test this hypothesis, we submitted the condemnation scale to a 3 (good deeds: same-domain vs. different-domain vs. control) X 2 (transgression type: blatant vs. ambiguous) ANCOVA, using the specific domain of the first article (discrimination vs. harassment) as a covariate. Raw means are plotted in Figure 3. Overall, the same-domain and different-domain conditions elicited less condemnation than the control condition ($Ms = -.16$, -.16, and .39; $SDs = .74$, .62, and .64, respectively), $F(2, 108) = 8.83$, $p < .0005$. A main effect of transgression type unsurprisingly revealed that participants condemned blatant transgressions more than ambiguous ones ($Ms = .39$ and. -.43; $SDs = .55$ and .61, respectively), $F(1, 108) = 57.85$, $p < .0001$. These two main effects, however, were qualified by the predicted interaction,

*F*(2, 108) = 3.65, *p* < .05, indicating that the effect of good deeds depended on transgression type.

In the next sections, we examine this interaction by testing our specific hypotheses about when licensing would occur. As in the previous two studies, we defined licensing as a decrease in condemnation when the target had performed good deeds compared to when he had not. First, we compared the same-domain condition to the control condition to test the hypothesis that same-domain good deeds would license ambiguous, but not blatant, transgressions. Next, we compared the different-domain condition to the control condition to test the hypothesis that different-domain good deeds would license both ambiguous and blatant transgressions. To test each of these hypotheses, we conducted separate analyses that regressed condemnation on: a) a contrast code that compared ambiguous and blatant transgressions, b) a contrast code that compared the control condition to the relevant good-deeds condition (i.e., either same-domain or different-domain; see Table 3 for coding), and c) the interaction of these two contrast codes (i.e., their product). Results are summarized in Table 3 and described below.

*Same-Domain Licensing Effect*

We first examined whether the difference between the same-domain and control conditions was moderated by transgression type (Equation 1 in Table 3). As predicted, it was: the interaction term indicated that same-domain good deeds licensed ambiguous transgressions more effectively than they licensed blatant transgressions. But did same-domain licensing occur *only* for ambiguous transgressions, as we expected? To answer this question, we examined simple effects in a separate regression analysis. The contrasts we used and the results are displayed in Table 4. We obtained the predicted pattern, finding evidence that participants

expressed less condemnation in the same-domain condition than in the control condition only when the transgression was ambiguous.

*Different-Domain Licensing Effect*

We next examined whether the difference between the different-domain and control conditions was moderated by transgression type (Equation 2 in Table 3). As predicted, it was not: different-domain good deeds reduced condemnation overall, but this main effect was unqualified by an interaction with transgression type. In other words, participants expressed less condemnation in the different-domain condition than in the control condition regardless of the transgression's ambiguity. This finding supports our hypothesis that different-domain good deeds (unlike same-domain good deeds) license both ambiguous and blatant transgressions.

*Direct Comparison of Same- and Different-Domain Good Deeds*

The results so far support our central hypotheses: different-domain good deeds licensed both blatant and ambiguous transgressions, while same-domain good deeds only licensed ambiguous transgressions. Our analyses suggest that same- and different-domain good deeds had different effects, depending on the transgression's ambiguity, but a direct comparison of the two good deeds conditions is necessary to confirm this. We thus examined whether the difference between the same-domain and different-domain conditions was moderated by transgression type (Equation 3 in Table 3). It was: the significant interaction confirmed that good deeds had different effects on condemnation depending on their domain and the transgression's ambiguity.

*Mechanisms*

Next we evaluated our predictions about mechanism, examining whether or not the target's apparent hypocrisy and participants' construal of his transgression explained the results

that we have described.  First, we tested the prediction that hypocrisy would suppress the same-domain licensing effect only in the blatant condition (as in Study 1), while construal would mediate the same-domain licensing effect only in the ambiguous condition.  Next, we tested the prediction that neither hypocrisy nor construal would explain the different-domain licensing effect.

To examine these predictions, we used the moderated path analysis procedure described by Edwards and Lambert (2007), which enabled us to examine mediation and suppression using similar formulae.  This procedure uses results from two regression equations to calculate path coefficients (i.e., regression slopes) between an independent variable (good deeds), a mechanism variable (hypocrisy or construal), and a dependent variable (condemnation) at each level of a moderator (ambiguity).  By examining these coefficients, we could determine whether the criteria for mediation or suppression were met (Baron & Kenny, 1986; MacKinnon, et al., 2000) for each of the ambiguity conditions.  Then, we could examine whether the mediation or suppression effect (i.e., the *indirect effect* of the independent variable through the mechanism variable) differed significantly based on transgression ambiguity.

The following sections describe these analyses, which we conducted separately for hypocrisy and construal.  For ease of presentation, we display the path coefficients for the analyses of hypocrisy and construal in the same figures.  Specifically, Figures 4 and 5 depict the path coefficients for same-domain and different-domain licensing, respectively, calculated by applying Edwards and Lambert's formulae to the regression results displayed in Table 5. (Numeric subscripts on path coefficients indicate the equation or equations in Table 5 from which the coefficients were calculated; coefficients with subscript *a* refer to ambiguous transgressions, while those with subscript *b* refer to blatant transgressions).  All significance

levels were tested by using a bootstrapping procedure with 1,000 resamples to construct bias-corrected confidence intervals (Edwards & Lambert, 2007; Stine, 1989).

*Mechanisms Behind the Same-Domain Licensing Effect*

*Suppression by hypocrisy when the transgression is blatant.* The results of the moderated path analysis for the same-domain licensing effect are displayed in Figure 4. The top panel shows that in the blatant transgression condition, as in Study 1, hypocrisy suppressed the same-domain licensing effect: participants perceived more hypocrisy in the same-domain condition than in the control condition (path $A_{1b}$); the more hypocritical they perceived the target, the more they condemned him (path $B_{2b}$); and controlling for hypocrisy *increased* the total effect of the same-domain condition from non-significance (path $C_{1,2b}$)[iv] to significance (path $C'_{2b}$). This increase was significant, demonstrating suppression (see indirect effect in Table 6).

In fact, when hypocrisy was controlled, the effect of same-domain good deeds on the condemnation of blatant transgressions was statistically indistinguishable from their effect on ambiguous transgressions (i.e., path $C'_{2b}$ in the top panel of Figure 4 and path $C'_{2a}$ in the bottom panel are not significantly different). In other words, if same-domain good deeds had not made it appear hypocritical to commit a blatant transgression, then they would have licensed both blatant and ambiguous transgressions to the same extent.

*No suppression when the transgression is ambiguous.* The bottom panel of Figure 4 shows that in the ambiguous condition, as predicted, hypocrisy did *not* suppress (or mediate) the same-domain licensing effect. Participants in the ambiguous condition, in contrast to those in the blatant condition, perceived an equivalent amount of hypocrisy in the same-domain and control conditions (path $A_{1a}$), and controlling for hypocrisy left the effect of the same-domain condition on condemnation virtually unchanged (compare paths $C_{1,2a}$ and $C'_{2a}$).

Was the difference in the suppression effect between the blatant and the ambiguous condition significant?  The last column of Table 6 indicates that it was: the 95% confidence interval of this difference did not include 0. Thus, we can conclude that hypocrisy suppressed the same-domain licensing effect when the transgression was blatant, but not when it was ambiguous.

*Mediation by construal when the transgression is ambiguous.*  The bottom panel of Figure 4 shows that, as predicted, construal mediated the same-domain licensing effect when the transgression was open to a favorable interpretation.  Participants interpreted the ambiguous transgression as less representative of harassment or discrimination in the same-domain condition than in the different-domain condition (path $A_{3a}$), and the less they viewed it as a transgression, the less they condemned it (path $B_{4a}$).  Controlling for construal significantly *reduced* the effect of the same-domain condition from significance (path $C_{1,2a}$)[4] to marginal significance (path $C'_{4a}$), and this decrease was significant, indicating mediation (see Table 6).

*No mediation by construal when the transgression is blatant.*  The top panel of Figure 4 shows that, as predicted, construal did *not* mediate same-domain licensing when the transgression was unambiguous.  Consistent with the idea that the transgression was blatant, participants were equally likely to construe it as discrimination or harassment (depending on the specific vignette) in the same-domain condition as in the control condition (path $A_{3b}$).  Controlling for construal left the effect of the same-domain condition virtually identical (compare paths $C_{1,2b}$ to $C'_{4b}$).

Was the difference in the mediation effect between the blatant and ambiguous transgression conditions significant? Again, the last column of Table 6 indicates that it was: the 95% confidence interval for this difference did not include 0.  Thus, we can conclude that

construal mediated the same-domain licensing effect when the transgression was ambiguous, but not when it was blatant.

*Mechanisms Behind the Different-Domain Licensing Effect*

In sharp contrast to the same-domain licensing effect, and consistent with our predictions, the different-domain licensing effect appears not to have been suppressed by hypocrisy or mediated by construal, regardless of transgression type. As can be seen in Figure 5, participants' perceptions of hypocrisy and their construal of the transgression were not affected by learning about different-domain good deeds (the four paths labeled *A*) with one exception: participants perceived marginally *less* hypocrisy in the different-domain condition than in the control condition when the transgression was unambiguous (path $A_{5b}$). As can be seen by comparing the paths labeled *C'* to the paths labeled *C* in Figure 5, controlling for construal and hypocrisy did not significantly change the effect of the different-domain condition on condemnation (see two leftmost columns in Table 6 for significance tests of these indirect effects). Finally, none of the indirect effects differed significantly between the ambiguous and the blatant transgression conditions (see last column in Table 6). Thus, we found no evidence that hypocrisy or construal played a mechanistic role in producing the different-domain licensing effect in either the ambiguous or the blatant transgression conditions.

*Summary of Mechanisms*

The moderated path analyses confirmed our predictions about mechanisms (see Table 6). When good deeds were in the same domain as a transgression, hypocrisy suppressed licensing for blatant transgressions, but not for ambiguous ones, while construal mediated licensing for ambiguous transgressions, but not for blatant ones. By contrast, when the good deeds were in a

*different* domain than a transgression, neither hypocrisy nor construal mediated nor suppressed the licensing that was observed.

## Discussion

Study 2 examined all four cells in our model of licensing (see Table 1), and confirmed our predictions about when and why licensing would occur. As in Study 1, blatant transgressions were licensed by different-domain good deeds, but not by same-domain good deeds, because same-domain good deeds made the target person appear hypocritical. By contrast, the ambiguous transgressions introduced in Study 2 were licensed by both different- and same-domain good deeds, but for different reasons. Namely, same-domain good deeds, unlike different-domain good deeds, changed the way the ambiguous transgressions were construed. This may explain why prior research found that actors licensed themselves to engage in questionable behavior based on a previous action in the same domain. In those studies, the behavior to be licensed was intentionally ambiguous (Effron et al., 2009; Monin & Miller, 2001).

A potential concern with Study 2 is that ambiguous transgressions may have seemed less severe than blatant transgressions. This concern, however, has difficulty accounting for our results. First, it is unclear why a manipulation of severity would produce mediation by construal only under the conditions specified by our model (see Tables 1 and 6). Second, it is unclear why a manipulation of severity would produce a crossover interaction with the good deeds' domain (Table 3, Equation 3): for blatant transgressions, good deeds tended to license more effectively in a different domain than in the same domain, while for ambiguous transgressions, the reverse tendency was apparent (see Figure 3). This crossover, however, is what one would expect if same-domain good deeds made blatant transgressions appear hypocritical but made ambiguous transgressions appear less likely to represent transgressions, as our model specifies.

GENERAL DISCUSSION

The present studies shed light on the question of when and why a history of moral behavior licenses transgressions. The results of our studies are summarized as effect sizes in the bottom half of Table 1, and closely match the hypotheses summarized in the top half. The Pilot Study showed that observers responded to transgressions with less condemnation when the transgressor had previously performed good deeds. In Studies 1 and 2, good deeds licensed blatant transgressions (e.g., sexual harassment) only when the good deeds and the transgressions were in a different domain (e.g., reducing adolescent drug use); good deeds in the same domain (e.g., reducing sexual harassment) made the transgressor appear hypocritical, thus undermining the licensing effect. Study 2 also examined observer reactions to ambiguous transgressions (e.g., suspicious behavior that might or might not represent sexual harassment), finding that good deeds licensed ambiguous transgressions regardless of their domain. Rather than appearing hypocritical, same-domain good deeds favorably disambiguated the suspicious behavior.

These results help integrate prior research on moral licensing with research on hypocrisy, and provide a resolution for a seeming inconsistency. On the one hand, research has suggested that a history of moral behavior can make it seem more legitimate to engage in morally questionable behavior (Krumm & Corning, 2008). On the other hand, behavior that contradicts prior moral claims can seem hypocritical and worthy of punishment (Barden, et al., 2005; Powell & Smith, 2009). The present studies explicate the conditions under which prior moral behavior will make transgressions seem excusable, versus hypocritical and inexcusable.

## Why Do Observers License Transgressions?

The present studies also begin to distinguish empirically between two similar, but theoretically distinct, perspectives on why licensing occurs. As described in our introduction, at

least two processes could underlie licensing. First, good deeds may provide what we have called moral credits, which balance out any moral debits subsequently incurred from committing bad deeds (e.g., Nisan, 1991). Second, good deeds may provide what we have called moral credentials, which change the way ambiguous behavior is construed, making them seem as if they were not transgressions at all (Monin & Miller, 2001). Different researchers have made different assumptions about which process explains licensing, but few have directly assessed process, and no one to our knowledge has previously attempted to distinguish empirically between the two (Merritt et al., 2010; Miller & Effron, in press).

The present results suggest that both processes can operate, albeit under different conditions. For licensing via balance (moral credits), it is not necessary that good deeds change the way observers construe a transgression. Indeed, all three of our studies found that good deeds licensed even blatant transgressions that were not open to favorable reinterpretation, as long as the good deeds and the transgression were in different domains, and Study 2 showed that this licensing effect occurred without a change in how participants construed the transgression. For example, learning that an employer had fought to reduce sexual harassment did not make participants perceive his racially discriminatory promotion decision as any less discriminatory, but nonetheless made them more permissive of it and less harsh in their evaluations of his character.

Licensing via construal (moral credentials), by contrast, should be most likely to occur when a) a transgression is ambiguous, and can thus be reconstrued more favorably, and b) when prior good deeds are in the same domain as the transgression, and thus seem relevant to interpreting it. Indeed, we found that good deeds in the same domain as ambiguous behavior prompted licensing by changing participants' construal of the ambiguous behavior (Study 2).

For example, learning that an employer had fought racial discrimination made participants perceive an ambiguous promotion decision he later made as less likely to represent racial discrimination, thereby increasing their tolerance of the behavior and decreasing their condemnation of his character.

Support for both theories of licensing should not be interpreted as equivocal or contradictory. Rather, this evidence strongly suggests that they represent two complementary routes to licensing. In fact, we believe that both routes can operate simultaneously. Consider the observation that the largest licensing effect size observed in the present studies was in the same-domain/ambiguous condition of Study 2 ($d = 1.31$; see Table 1). Perhaps same-domain good deeds prompted licensing not only by making observers construe the ambiguous transgression more favorably (moral credentials), but also by balancing any part of the ambiguous behavior that was still construed unfavorably (moral credits). Indeed, construal did not entirely explain the licensing effect in this condition: as can be seen in the bottom panel of Figure 4 (path $C'_{4a}$), when construal was statistically controlled, the effect of same-domain licensing on condemnation was still marginally significant. These observations should be replicated before drawing strong conclusions, but we predict that licensing will be most effective when both moral credits and moral credentials can operate simultaneously.

<center>Effects of Transgressions on Perceptions of Prior Good Deeds</center>

We have focused in this paper on how prior good deeds can affect observers' evaluation and construal of subsequent transgressions, but transgressions can also affect the evaluation and construal of prior good deeds. In Study 1, for example, the same good deeds elicited less positive evaluations and seemed less sincere when subsequent transgressions were in the same domain than when they were in a different domain, which probably contributed to ascriptions of

hypocrisy.  Consistent with past theorizing (Trope, 1986), we expect that observer reactions to a transgression will depend in part on how the transgression shapes perceptions of prior good deeds, especially when the good deeds are ambiguous (e.g., when they could reflect either true moral intent or mere self-presentation).[v]

## Self-Licensing

The present studies have focused on how observers react to actors who perform good deeds and then transgress, but we suspect that similar principles apply to situations in which actors *self-license* – that is, permit themselves to transgress.  We know that actors seek to avoid hypocrisy (Stone, et al., 1997), but that acting morally makes them feel comfortable performing morally questionable behavior (for reviews, see Merritt, et al., 2010; Miller & Effron, in press).  Future research should clarify when actors display licensing versus hypocrisy effects. Extrapolating from the present studies suggests that good deeds should liberate actors to perform morally ambiguous behaviors in any domain, but only liberate blatant transgressions that are in a different domain.

## Conclusion

The present research indicates that when people are called upon to judge a morally dubious behavior, they consider not only the behavior itself, but also the moral track record of the actor who performed it.  Establishing an exemplary track record may thus allow actors to minimize the risk that their subsequent misdeeds will elicit condemnation from observers – unless actors attempt to commit blatant transgressions in the same domain as their exemplary behavior.  In that case, they are likely to find themselves without a license.

REFERENCES

Abernathy, R. D. (1989). *And the walls came tumbling down*. New York: Harper & Row.

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression

   formation. *Journal of Experimental Psychology, 70,* 394-400.

Asch, S. (1946). Forming impressions of personality. *Journal of Abnormal and Social*

   *Psychology, 41,* 258-290.

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,* 1463-1474.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social

   psychological research: Conceptual, strategic, and statistical considerations. *Journal of*

   *Personality and Social Psychology, 51,* 1173-1182.

Birnbaum, M. H. (1973). Morality judgment: Test of an averaging model with differential

   weights. *Journal of Experimental Psychology, 99,* 395-399.

Cialdini, R. B. (1988). *Influence: Science and practice*. UK: Harper Collins.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A

   general analytic framework using moderated path analysis. *Psychological Methods, 12,*

   1-22.

Effron, D. A., Cameron, J. S., & Monin, B. (2009). Endorsing Obama licenses favoring Whites.

   *Journal of Experimental Social Psychology, 45,* 590-593.

Hakim, D., & Santos, F. (2008). G.O.P. Consultant says he reported Spitzer trysts in 2007. *The*

   *New York Times*. Retrieved September 5, 2009 from

http://www.nytimes.com/2008/03/24/nyregion/24spitzer.html?scp=6&sq=eliot+spitzer+h ypocrisy&st=nyt.

Hastorf, A. H., Schneider, D. J., & Polefka, J. (1970). *Person perception*. Reading, MA: Addison-Wesley.

Hollander, E. P. (1958). Conformity, status, and idiosyncrasy credit. *Psychological Review, 65*, 117-127.

Jones, E. E. (1990). *Interpersonal perception*. New York: Freeman.

Jordan, J., Mullen, E., & Murnighan, J. K. (2009). *Moral miscreants and unethical angels: Striving to achieve a moral equilibrium*. Manuscript submitted for publication.

Krumm, A. J., & Corning, A. F. (2008). Who believes us when we try to conceal our prejudices? The effectiveness of moral credentials with in-groups versus out-groups. *The Journal of Social Psychology, 148*, 689-709.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*, 173-181.

Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass, 4*, 344-357.

Miller, D. T., & Effron, D. A. (in press). Psychological license: When it is needed and how it functions. In M. P. Zanna and J. M. Olson (Eds.) *Advances in experimental social psychology* (Vol. 43, pp. 117-158). San Diego, CA: Academic Press/Elsevier.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*, 33-43.

Nisan, M. (1991). The moral balance model: Theory and research extending our understanding

of moral choice and deviation. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of

moral behavior and development* (pp. 213-249). Hillsdale, NJ: Erlbaum.

Powell, C. A. J., & Smith, R. H. (2009). *The inherent joy in seeing hypocrites hoisted with their

own petards*. Unpublished manuscript.

Riskey, D. R., & Birnbaum, M. H. (1974). Compensatory effects in moral judgment: Two rights

don't make up for a wrong. *Journal of Experimental Psychology, 103*, 171-173.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of

moral self-regulation. *Psychological Science, 20*, 523-528.

Stine, R. (1989). An introduction to bootstrapping methods. *Sociological Methods and Research,

18*, 243-291.

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*, 54-54.

Tedeschi, J. T. S., Barry R.;  Bonoma, Thomas V.  (1971). Cognitive dissonance: Private

ratiocination or public spectacle? *American Psychologist, 26*, 685-695.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution.

*Psychological Review, 93*, 239-257.

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, NC: Information Age Publishing.

# AUTHOR NOTE

Daniel A. Effron, Department of Psychology, Stanford University, and Benoît Monin, Graduate School of Business and Department of Psychology, Stanford University. Correspondence regarding this article may be sent to the first author via email to effron@stanford.edu, or via post to Department of Psychology, Stanford University, Jordan Hall Building 420, 450 Serra Mall, Stanford, CA 94303.

FOOTNOTES

[1] These analyses used the following contrast code: same domain = 1, different domain = -1,

control = 0. A first equation revealed more condemnation in the same-domain vs. the

different-domain condition, $b = .16$, $t(73) = 2.08$, $p < .05$; a second equation showed the

same pattern for hypocrisy, $b = .68$, $t(73) = 6.31$, $p < .001$; and a third equation showed

that controlling for hypocrisy, $b = .21$, $t(72) = 2.68$, $p < .01$, eliminated the difference in

condemnation between the two domain conditions, $b = .01$, $t(72) = .15$, $ns$ – a change that

was significant by the Sobel test, $z = 2.47$, $p = .01$

[ii] The item asking how much blame the target deserved was not measured in Study 2 because it

failed to correlate highly with the condemnation scale in Study 1.

[iii] Specifically, we contrast-coded the first article's topic (discrimination = -1, harassment = 1,

control = 0) to control for the difference between the discrimination and the harassment

topics without eliminating the difference between them and the control condition (in which

participants did not see the first article).

[iv] Paths $C_{1,2b}$ and $C_{1,2a}$ can both be calculated from either equations 1 and 2 or equations 3 and 4

in Table 5.

[v] Trope's two-stage model of attribution (1986) is also consistent with our distinction between

moral credits and moral credentials. According to Trope, observers can infer an actor's

disposition directly from her prior behavior, or indirectly from the way her prior behavior

shapes their construal of her subsequent behavior. In our model, observers can judge an

actor's moral character and the permissibility of her transgression based directly on her prior

good deeds (balance/credits), or indirectly based on the way her good deeds shapes their

construal of her transgression (credentials).  Consistent with our findings, Trope predicts that

observers will be most influenced by the indirect path when the subsequent behavior is

ambiguous.

TABLES

*Table 1:* Summary of hypotheses and effect sizes

| | Blatant transgressions | | Ambiguous transgressions | |
|---|---|---|---|---|
| | Different-domain good deeds | Same-domain good deeds | Different-domain good deeds | Same-domain good deeds |
| **Predictions** | | | | |
| Mechanism | Balance (moral credits) | Hypocrisy | Balance (moral credits) | Construal (moral credentials) |
| Effect | Licensing | No licensing | Licensing | Licensing |
| **Observed effect sizes** | | | | |
| Pilot Study | **0.89** *** | -- | -- | -- |
| Study 1 | **0.59** * | 0.00 | -- | -- |
| Study 2 | **0.85** ** | 0.38 | **0.85** ** | **1.31** *** |

*Notes:* Positive effect sizes indicate licensing.  Effect sizes are Cohen's *d* for the difference between the (same or different-domain) good deeds condition and the control condition, calculated using the error term of the omnibus analysis reported in the study indicated.  * = $p < .05$, ** = $p < .01$, *** = $p < .001$, testing the hypothesis that this difference is 0.

*Table 2*: Study 2: Text of transgression type manipulation for each specific domain of the second article

| Type of transgression | Domain of second article | |
| --- | --- | --- |
| | Racial discrimination | Sexual harassment |
| Blatant | An email obtained by the press, in which Hutchinson wrote to one of his colleagues that [the two Black employees'] race would make them "unsuitable for management," corroborates the employees' accusations. When questioned about the email, Hutchinson admitted that he was reluctant to give promotions to African-Americans. | An email obtained by the press, in which Hutchinson wrote to one of his colleagues that Krasne had "refused to go on a date with me," corroborates Krasne's accusations. When questioned about the email, Hutchinson admitted that he was hoping to use his power over the promotion decision to persuade Krasne to have sex with him. |
| Ambiguous | When questioned, Hutchinson denied the charges, claiming that the two employees who did not receive the promotion had simply not performed as well as the ones he promoted. | When questioned, Hutchinson denied the charges, claiming that he had suggested the dinner as an informal, but strictly professional, interview. He called Krasne's allegations, "an unfortunate misunder-standing," and noted that Krasne was not the only applicant he had invited to dinner. Hutchinson said he had decided not to promote Krasne simply because the employee he did promote had been a stronger applicant. |

*Table 3*: Study 2: Contrast coding and regression results for analysis of same-domain licensing, different-domain licensing, and direct comparison of domains

| Analysis | Equation | Coding of good deeds | | | Transgression | | Good deeds | | Good deeds X transgression | | Covariate | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Same Domain | Control | Different domain | b | t | b | t | b | t | b | t |
| Same-domain licensing | 1 | 1 | -1 | 0 | -0.39 | 7.53 *** | -0.24 | 3.36 *** | **-0.13** | **2.01 *** | 0.09 | 1.42 |
| Different-domain licensing | 2 | 0 | -1 | 1 | -0.39 | 7.51 *** | **-0.21** | **3.19 *** | 0.03 | 0.48 | 0.05 | 0.87 |
| Direct comparison of domains | 3 | 1 | 0 | -1 | -0.42 | 7.80 *** | <.01 | <.01 | **-0.15** | **2.26 *** | 0.09 | 1.44 |

*Notes*: *b*s are unstandardized coefficients. Boxed statistics are referenced in text. Degrees of freedom for *t*s are 110. Transgression contrast coded -1 for blatant transgression condition and 1 for ambiguous transgression condition. Covariate is specific domain of first article (discrimination = -1, harassment = 1, control = 0). * = *p* < .05, ** = *p* < .01, *** = *p* < .001.

*Table 4*: Study 2: Contrast codes and regression results for simple effects of same-domain licensing in blatant and ambiguous conditions.

| | Coding of contrast | | | | | | Regression results | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Blatant transgressions: good deeds condition | | | Ambiguous transgressions: good deeds condition | | | | |
| Contrast name | Same-domain | Control | Different-domain | Same-domain | Control | Different-domain | b | t |
| Same-domain / blatant | 1 | -1 | 0 | 0 | 0 | 0 | -0.10 | 1.20 |
| Same-domain / ambiguous | 0 | 0 | 0 | 1 | -1 | 0 | -0.37 | 3.77 *** |
| Transgression type | -1 | -1 | -1 | 1 | 1 | 1 | -0.39 | 7.53 *** |
| Covariate | -- | -- | -- | -- | -- | -- | 0.09 | 1.42 |

*Notes*: *b*s are unstandardized regression coefficients. Degrees of freedom for *t*s = 110. *** = $p <$ .001. Covariate is specific domain of first article (discrimination = -1, harassment = 1, control = 0). Statistics in box are referenced in text.

*Table 5*: Study 2: Regression results for moderated path analysis

| Good deeds | Mechanism variable | Equation | Dependent variable | Good deeds | Trans-gression | Good deeds X trans-gression | Mechanism | Trans-gression X mechanism | Covariate |
|---|---|---|---|---|---|---|---|---|---|
| Same-domain (vs. control) | Hypocrisy | 1 | Hypocrisy | 0.22 * | -0.31 *** | -0.29 | -- | -- | -0.06 |
| | | 2 | Condemnation | -0.29 *** | -0.27 *** | -0.04 | 0.36 *** | 0.07 | 0.11 * |
| | Construal | 3 | Construal | -0.25 ** | -0.55 *** | -0.23 * | -- | -- | -0.03 |
| | | 4 | Condemnation | -0.11 * | -0.15 ** | -0.02 | 0.44 *** | 0.06 | 0.10 * |
| Different-domain (vs. control) | Hypocrisy | 5 | Hypocrisy | -0.21 † | -0.29 ** | 0.05 | -- | -- | -0.09 |
| | | 6 | Condemnation | -0.15 ** | -0.30 *** | 0.04 | 0.31 *** | 0.13 * | 0.08 |
| | Construal | 7 | Construal | -0.10 | -0.57 *** | 0.05 | -- | -- | -0.07 |
| | | 8 | Condemnation | -0.17 ** | -0.14 ** | 0.02 | 0.44 *** | 0.1 * | 0.08 † |

*Notes*: Values represent regression coefficients. Hypocrisy and construal were standardized. Equations 1-4 were used to calculate the path coefficients for *same*-domain licensing via hypocrisy (Equations 1 and 2) and construal (Equations 3 and 4) shown in Figure 4. Equations 5-8 were used to calculate the path coefficients for *different*-domain licensing via hypocrisy (Equations 5 and 6) and construal (Equations 7 and 8) shown in Figure 5. Contrasts were coded as in Table 3. † = p < .10, * = p < .05, ** = p < .01, *** = p < .001.

*Table 6*: Study 2: Indirect effects for same- and different-domain licensing via hypocrisy and construal

| Good deeds and mechanism | Transgression type | | Difference | |
| | Blatant | Ambiguous | | |
| | *b* | *b* | *b* | 95% CI |
| --- | --- | --- | --- | --- |
| Same-domain | | | | |
| Hypocrisy | **0.15** ** *(Suppression)* | -0.03 | **-0.18** * | (-.37, -.03) |
| Construal | -0.01 | **-0.24** ** *(Mediation)* | **-0.23** * | (-.43, -.06) |
| Different-domain | | | | |
| Hypocrisy | -0.05 | -0.07 | -0.02 | (-.16, .11) |
| Construal | -0.05 | -0.03 | 0.02 | (-.16, .22) |

*Notes: b*s are regression coefficients. Hypocrisy and construal are standardized. * = $p < .05$; ** = $p < .01$. CI = confidence interval calculated using bias-corrected bootstrapping method (Edwards & Lambert, 2007; Stine, 1989). Difference is significant at .05 level if CI does not include 0.

FIGURES

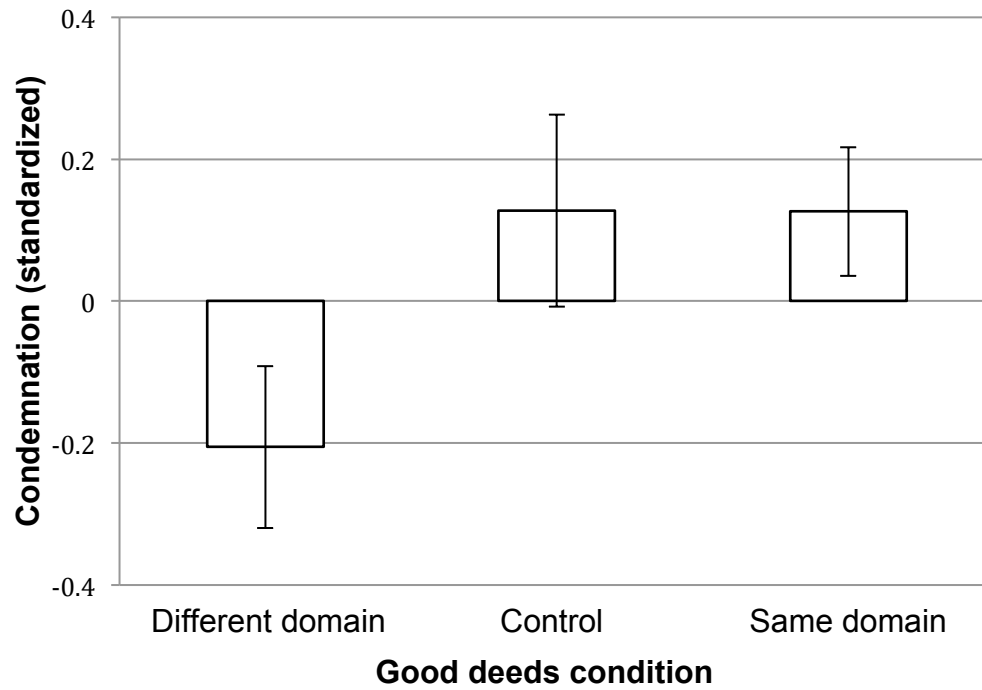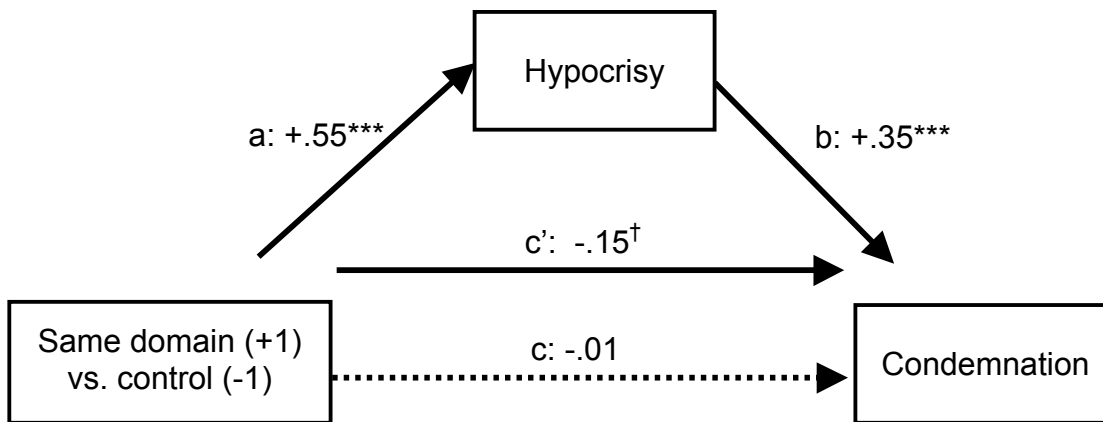*Figure 1.* Study 1: Mean condemnation (± *SE*) by good deeds condition

*Figure 2:* Study 1: Suppression analysis of same-domain licensing effect



*Notes:* Numbers represent regression coefficients. Hypocrisy is standardized. Path c represents the total effect of the independent variable on condemnation; path c' represents the direct effect of the independent variable on condemnation controlling for hypocrisy (see Baron & Kenny, 1986). † = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$

*Figure 3.* Study 2: Mean condemnation (± *SE*), by good deeds condition and transgression type
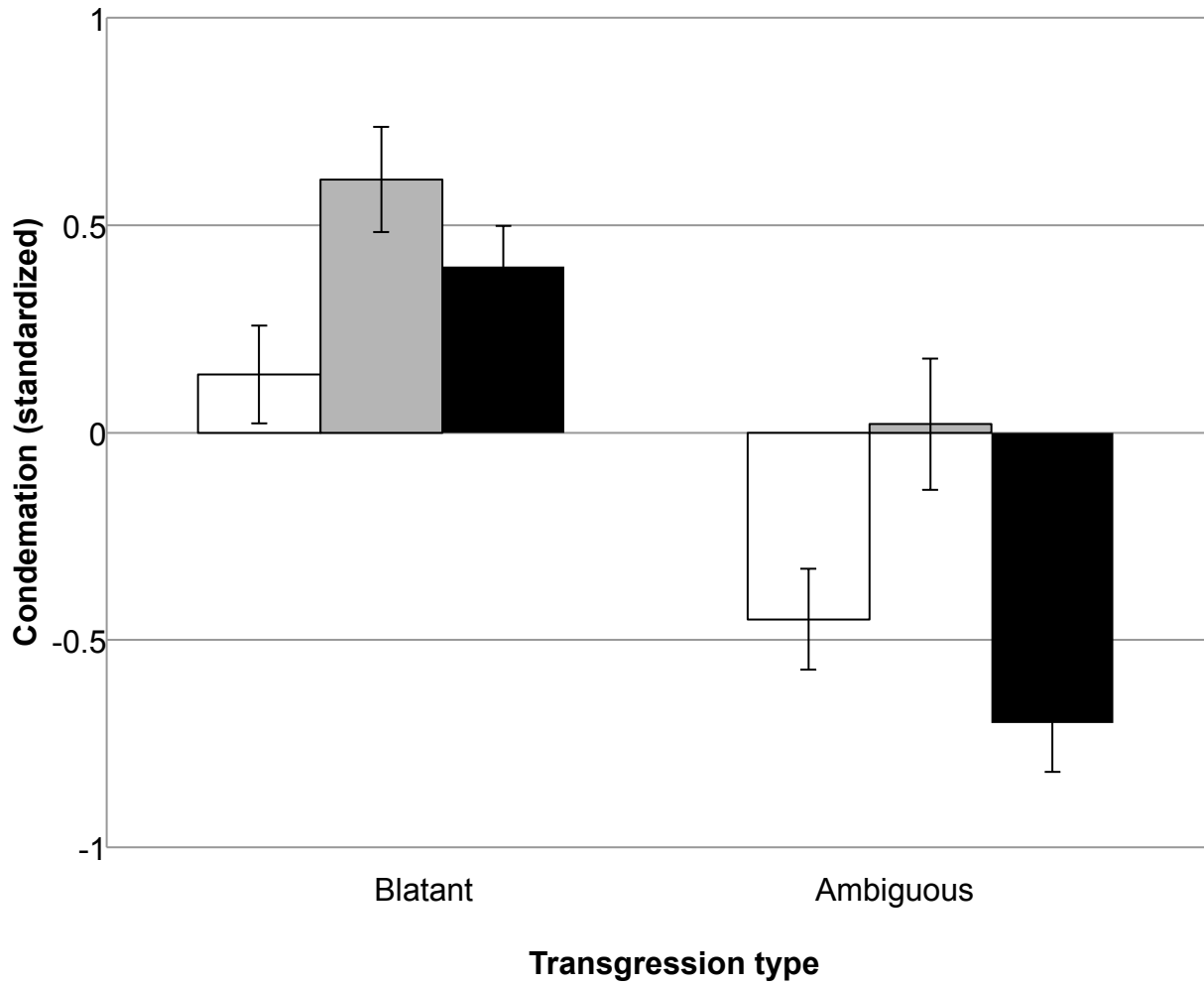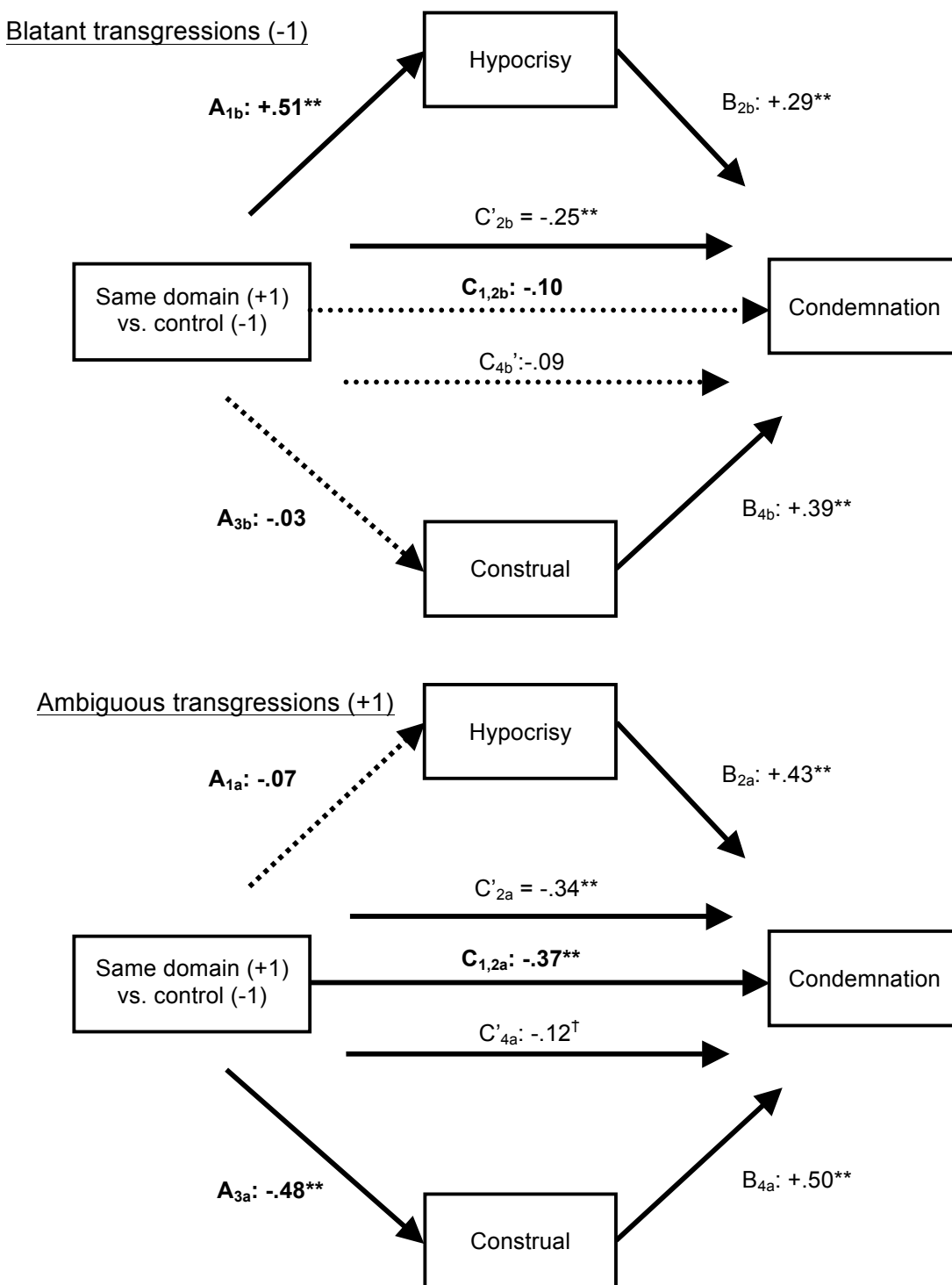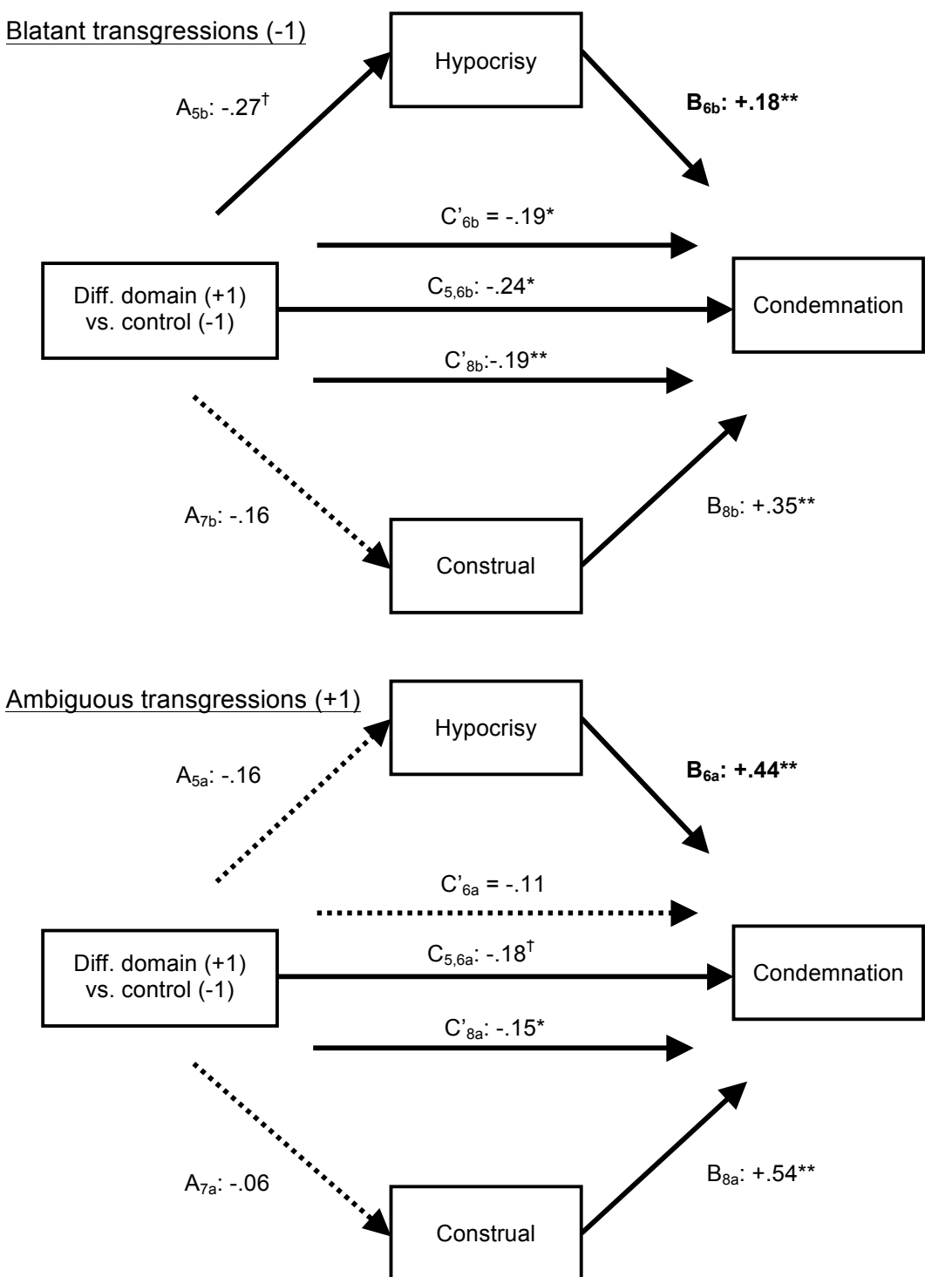
*Figure 4:* Study 2: Moderated path analysis results for *same-domain* licensing for blatant transgressions (top panel) and ambiguous transgressions (bottom panel).

Blatant transgressions (-1)

Hypocrisy

$A_{1b}$: **+.51\*\***

$B_{2b}$: +.29\*\*

$C'_{2b}$ = -.25\*\*

Same domain (+1) vs. control (-1)

$C_{1,2b}$: **-.10**

Condemnation

$C_{4b}'$:-.09

$A_{3b}$: **-.03**

Construal

$B_{4b}$: +.39\*\*

Ambiguous transgressions (+1)

Hypocrisy

$A_{1a}$: **-.07**

$B_{2a}$: +.43\*\*

$C'_{2a}$ = -.34\*\*

Same domain (+1) vs. control (-1)

$C_{1,2a}$: **-.37\*\***

Condemnation

$C'_{4a}$: -.12$^{\dagger}$

$A_{3a}$: **-.48\*\***

Construal

$B_{4a}$: +.50\*\*

*Notes*: Values are path coefficients calculated using the percentile bias-corrected bootstrap method (Edwards & Lambert, 2007) separately for hypocrisy and construal. Paths *C* indicate the total effect of the independent variable on condemnation; Paths *C'* indicate the direct effect controlling for either hypocrisy or construal. Hypocrisy and construal are standardized. Solid lines indicate significant paths, † = *p* < .10, * = *p* < .05, ** = *p* < .01, *** = *p* < .001. Bold-faced coefficients differ significantly between the top panel and the bottom panel, *p* < .05. Numeric subscripts on path coefficients indicate the equation or equations in Table 5 from which the coefficients were calculated; those with subscript *a* refer to ambiguous transgressions; those with subscript *b* refer to blatant transgressions. Paths $C_{1,2}$ may be calculated from either equations 1 and 2 or equations 3 and 4.

*Figure 5:* Study 2: Moderated path analysis results for *different-domain* licensing for blatant transgressions (top panel) and ambiguous transgressions (bottom panel).



**Notes**: Values are path coefficients calculated using the percentile bias-corrected bootstrap method (Edwards & Lambert, 2007) separately for hypocrisy and construal. Paths *C* indicate the total effect of the independent variable on condemnation; Paths *C'* indicate the direct effect controlling for either hypocrisy or construal. Hypocrisy and construal are standardized. Solid lines indicate significant paths, † = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$. Bold-faced coefficients differ significantly between the top panel and the bottom panel, $p < .05$. Numeric subscripts on path coefficients indicate the equation or equations in Table 5 from which the coefficients were calculated; those with subscript *a* refer to ambiguous transgressions; those with subscript *b* refer to blatant transgressions. Paths $C_{5,6}$ may be calculated from either equations 5 and 6 or equations 7 and 8.