

# Understanding the Magnitude of Hypocrisy in Moral Contradictions: The Role of Surprise at Violating Strong Attitudes

Personality and Social Psychology Bulletin  
2024, Vol. 50(12) 1661–1674  
© 2023 by the Society for Personality and Social Psychology, Inc  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/01461672231177773  
journals.sagepub.com/home/pspb



Jacob D. Teeny<sup>1</sup> , Jaroth V. Lanzalotta<sup>2</sup>,  
and Richard E. Petty<sup>2</sup> 

## Abstract

Although two people could both enact similar forms of hypocrisy, one person might be judged as *more* hypocritical than the other. The present research advances a novel, theoretical explanation for a paradigmatic instance of this: the increased hypocrisy ascribed to contradicting a morally (vs. nonmorally) based attitude. In contrast to prior explanations, the present research shows that people infer targets holding morally (vs. nonmorally) based attitudes are more difficult to change. Consequently, when people are hypocritical on these stances, it elicits greater surprise, which amplifies the perceived hypocrisy. Through both statistical mediation and experimental moderation, we provide evidence for this process and show how our explanation generalizes to understanding heightened hypocrisy in other contexts, too (i.e., violating nonmoral attitudes held with certainty vs. uncertainty). Altogether, we provide an integrative, theoretical lens for predicting when moral and nonmoral acts of hypocrisy will be perceived as particularly hypocritical.

## Keywords

hypocrisy, morality, attitude strength, difficulty, inferences

Received June 21 2022; revised accepted May 2 2023

It is well understood that the more people are judged as *hypocrites*, the more negative the social sanctions they face (Bhatti et al., 2013; Kougiannou & Wallis, 2020; Laurent et al., 2014). In the present research, we focus on a prominent and robust context previously shown to amplify this socially punitive judgment: when a person hypocritically contradicts a moral (vs. nonmoral) stance, what we label “the moral-augmentation effect” (Kreps et al., 2017; see also Alicke et al., 2013; Effron et al., 2018; Kreps & Monin, 2014). Within this context, we identify a novel explanation for why hypocrisy on moral versus nonmoral attitudes is perceived as especially hypocritical, and we use this understanding to offer a theoretical umbrella for predicting when other factors might heighten judgments of hypocrisy, too.

## Hypocrisy Judgments

### *The Role of Morality*

Hypocrisy has been defined differently by different scholars (Barden et al., 2005; Jordan et al., 2017; Laurent & Clark, 2019; for a review, see Hale & Pillow, 2015); however, at its core, it involves a “misalignment” or inconsistency between

a person’s initial stance on a topic and a subsequent attitude or behavior (Alicke et al., 2013; Effron et al., 2018). This could entail “saying one thing but doing another,” “doing one thing but saying another,” or any combination of these contradictions. In the present work—and in line with prior research (e.g., Effron et al., 2018)—rather than examining what factors (or “ingredients”) transform a contradiction into “hypocrisy,” we focus on understanding what factors vary the *degree* of hypocrisy ascribed to someone once the elements of hypocrisy are present. Of most relevance, Kreps and colleagues (2017) found that “leaders” who contradicted a position they based on a sense of morality were seen as more hypocritical than when the position was based on practicality (see also, Alicke et al., 2013; Effron et al., 2018; Kreps & Monin, 2014). This phenomenon (aka, the moral-augmentation effect) was shown across 14 studies, a

<sup>1</sup>Northwestern University, Evanston, IL, USA

<sup>2</sup>The Ohio State University, Columbus, USA

### Corresponding Author:

Jacob D. Teeny, Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, USA.  
Email: jacob.teeny@kellogg.northwestern.edu

variety of topics (e.g., capital punishment and sexualization in ads), and for both proattitudinal and counterattitudinal positions on the issue.

To explain this effect, Kreps and colleagues (2017) speculated that moral- (vs. practical-) based hypocrisy conveys ‘moral deceit’: People should be committed to their moral attitudes, and so hypocrisy in this instance is akin to violating a sacred promise. This morality-based explanation is consistent with other research which has explained magnified hypocrisy judgments through a moral lens, such as undeservingly claiming a virtuous status (Effron et al., 2018; O’Connor et al., 2020). Although we do not deny that moral accounts could be relevant for the moral-augmentation effect we are examining, we propose a simpler, nonmoral explanation for it—one in fact hinted at by Kreps and colleagues (2017). That is, within the idea of ‘moral deceit’ is the notion that perceivers simply *expect* morally based attitudes not to be contradicted. This potential lay belief, as we explain shortly, offers a different explanation for the phenomenon that does not rely on moral reasoning strategies. In turn, we can use this explanation to predict when attributions of hypocrisy will be higher or lower in other contexts, too.

### An Attitude Strength Perspective

In better understanding the antecedents and consequences of social judgments, decades of research have developed an extensive literature on *attitude strength* (Petty & Krosnick, 1995; see Luttrell & Sawicki, 2020, for a recent review). This work identifies how dimensions of one’s own attitude (e.g., its certainty and knowledge) predict the influence of that attitude on one’s judgments and behavior. For example, the more *certain* a person is in an attitude, the more likely that attitude is to persist over time and context (Rucker et al., 2014). Although researchers have debated whether an attitude’s moral basis represents another one of these strength dimensions (Philipp-Muller et al., 2020) or something more unique (e.g., Skitka et al., 2005), considerable research shows that it can produce outcomes in line with those of other strength features. For example, like certainty, the more one’s attitude is based on a sense of morality, the more likely that attitude is to persist over time (e.g., Aramovich et al., 2012; Luttrell et al., 2016; Skitka & Morgan, 2014). Notably, dimensions of an attitude’s strength are orthogonal to the topic for which the attitude applies. For example, a person could hold a morally based attitude toward a normatively nonmoral topic (e.g., dislike ice cream because it is produced through animal labor). Or a person could hold a nonmorally based attitude toward a normatively moral topic (e.g., oppose capital punishment because of its cost-ineffectiveness). With regard to hypocrisy, we show how the perceived strength of the target’s attitude (irrespective of how the topic might be normatively perceived) can predict the magnitude of hypocrisy ascribed to the person.

Interestingly, although attitude strength research has extensively examined these dimensions as they pertain to an individual’s own attitude, there is little research on how people perceive the strength of *others’* attitudes (Teeny & Petty, 2022). With regard to moral attitudes specifically, there was one perception we believed to be particularly relevant: inferences about the attitude’s difficulty to change. In general, people generate inferences based on their own personal experiences (cf. Ross & Nisbett, 1991; Thibaut & Kelley, 1959). So, if people typically find their own morally (vs. nonmorally) based attitudes are resistant to changing, they would likely infer that others’ moral attitudes are similarly resistant. This possibility coheres with speculation by Kreps and colleagues (2017) who suggested that people might prescriptively expect others’ moral attitudes not to change. Because this notion plays a critical role in our later theorizing, we ran a brief pilot study to test this belief ( $n = 194$ ; see the web appendix). In it, we provided participants with three very different scenarios, where across them, we found that targets described to hold an attitude based on morality (vs. practicality) were inferred to be more difficult to change ( $p = .003$ ,  $d = 0.44$ ). This finding then provided us with the base rationale to present a new theoretical explanation for the moral-augmentation effect in hypocrisy.

### The Role of Surprise

To reiterate, the prevailing account for the moral-augmentation effect is that hypocrisy on moral (vs. nonmoral) attitudes conveys a unique and therefore particularly offensive form of deceit (Jordan et al., 2017; Kreps et al., 2017; O’Connor et al., 2020). However, based on the observation that people expect moral attitudes to be more persistent, we offer an alternative explanation: Hypocritical contradictions of morally (vs. nonmorally) based attitudes are particularly *surprising*, and these greater feelings of surprise amplify hypocrisy judgments.<sup>1</sup>

Surprise is a well-studied construct in both social and cognitive psychology for its potential to affect people’s judgments and decision-making (Mellers et al., 1997; Reisenzein et al., 2019). Surprise emerges from people’s inability to “make sense” of novel information, where the more difficult it is to integrate new information with an existing representation (or expectation), the more one is surprised (Foster & Keane, 2015; Maguire et al., 2011). Notably, surprise is considered a temporally primary psychological response, where multimethod approaches reveal that surprise interrupts ongoing mental processes and thereby precedes related, value-laden judgments (Noordewier et al., 2016). Because of this, surprise has been shown to have a variety of consequences for subsequent evaluations, where arguably its most noted influence has been in terms of amplifying evaluative extremity (aka, *decision affect theory*; Mellers et al., 1997). The more surprise one experiences (e.g., getting an unexpected “fail” on

an exam), the more extreme (e.g., disappointed) the consequent reaction tends to be (e.g., Durso et al., 2021; Shepperd & McNulty, 2002). With hypocrisy, this suggests that greater surprise at a target's hypocritical contradiction should heighten hypocrisy judgments. Said differently, less surprising hypocrisy should result in lower judgments of it. Thus, because morally (vs. nonmorally) based attitudes are inferred to be more difficult to change, it should be harder for perceivers to mentally integrate a morally based, hypocritical contradiction. The resultant, greater surprise should then elicit greater perceptions of hypocrisy.

According to this logic, other nonmoral factors that also prompt greater surprise at a hypocritical contradiction should also lead to greater judgments of hypocrisy. For example, previous research suggests that people perceive others' certain (vs. uncertain) attitudes as more difficult to change (Petrocelli et al., 2010).<sup>2</sup> Consequently, should a target hypocritically contradict a certain (vs. uncertain) attitude, it should lead to more surprise and therefore greater judgments of hypocrisy. Indeed, past research has found that hypocrisy on broadly construed "strong" (vs. "weak") attitudes is rated as more hypocritical (Laurent & Clark, 2019); however, no empirical explanation of this effect was provided—nor have hypocrisy judgments been linked explicitly to attitude certainty. Thus, in addition to testing our novel process for the moral-augmentation effect, we also test whether it can apply to other, nonmoral factors to both provide convergent validity for the explanation as well as expand its generalizability. In sum, we offer a new way of conceiving the moral-augmentation effect in hypocrisy, which in turn offers a novel theoretical lens for both organizing other findings in the hypocrisy literature as well as predicting new contexts in which this social judgment should be magnified.

## Research Overview

In Studies 1a and 1b, we provide initial evidence for the proposed process behind the moral-augmentation effect through sequential mediation analyses on two separate topics. In Study 2, we use a nonmoral topic and a novel, nonmoral magnifier of hypocrisy (attitude certainty) to provide convergent evidence for the role of surprise in enhancing hypocrisy judgments. Study 3 simultaneously manipulates the target's attitude morality as well as certainty to highlight their similar effects and processes within a single experiment. Studies 4 and 5 take a process-by-moderation approach (Spencer et al., 2005) to show how reducing surprise at a hypocritical contradiction—either through a dispositional trait of the target's attitude or a situational context the target experienced—can blunt the moral-augmentation effect.

For the following studies, we report all materials, manipulations, and measures in the web appendix. Two studies were preregistered (Studies 2 and 5), and all data can be found at this OSF link: [https://osf.io/gszpw/?view\\_only=320daf9cfb5](https://osf.io/gszpw/?view_only=320daf9cfb5)

64d8492b270cd77c956ba. No data were collected after reaching our target sample and all sample sizes were determined in advance. All exclusions are reported in the primary text for all studies, and sensitivity analyses are reported to indicate the minimum effect size that our samples had the power to detect.

## Studies 1a and 1b: The Role of Surprise in the Moral-Augmentation Effect

In our first studies, we take two scenarios from past research (Kreps et al., 2017) and manipulate whether the target expresses an attitude based on morality or practicality. We then measure the inferred difficulty to change the target's attitude, the surprise elicited at their hypocrisy, and then the perceived degree of it. In Study 1b, we also collect judgments about the target's overall character following his hypocrisy. Specifically, we assess two specific character judgments from past hypocrisy research: the degree to which the target is "morally deceptive" and engages in "false signaling." We will ultimately control for these measures to determine the explanatory power of our proposed process in the moral-augmentation effect.

### Participants and Design

Study 1a (capital punishment) received 205 MTurk participants, and Study 1b (sexualization in advertising) received 211 MTurk participants through CloudResearch (as do all the following studies; Litman et al., 2017). Each study consisted of a two-cell (attitude basis: moral vs. practical) between-subjects design. As criteria used throughout this research, we excluded participants who failed a simple attention check of recalling the focal topic in a free response box (Study 1a: moral,  $n = 12$ ; practical,  $n = 11$ ) or multiple choice (Study 1b: moral,  $n = 3$ ). Thus, our final samples entailed: Study 1a,  $n = 182$  ( $M_{\text{age}} = 37.0$ ; female = 52%) and Study 1b,  $n = 208$  ( $M_{\text{age}} = 41.2$ ; female = 44%). A sensitivity analysis (two-tailed,  $\alpha = .05$ , power = 80%; G\*Power; Faul et al., 2007) indicates we could detect a main effect difference of  $d = 0.42$  for Study 1a and  $d = 0.39$  for Study 1b, both medium-sized effects (Cohen, 1988).

### Procedures

**Study 1a.** Participants read a vignette about Cody, whose home state would be voting on whether to legalize capital punishment. Participants read one of the two vignettes describing Cody's opposition to capital punishment as being based on moral reasons (e.g., "it's immoral to execute a prisoner") or practical reasons (e.g., "it's too expensive to execute a prisoner"), adapted from Kreps et al. (2017). Participants then reported the inferred difficulty to change the target's attitude as well as perceptions of the target's

moral bases. Following this, participants then read that Cody behaved hypocritically and voted in favor of the death penalty. Participants then rated their surprise at Cody's contradiction and the magnitude of his hypocrisy.

**Study 1b.** In a similar paradigm, participants read a vignette about Mike, whose marketing team would be voting on whether to use sexualized advertisements. Here, Mike expressed opposition to using sexualized content for moral reasons (e.g., "using sexualized content to sell products seems morally wrong") or practical reasons (e.g., "using sexualized content to sell products seems like it won't make as much money"). In line with Study 1a, participants reported the inferred difficulty to change Mike's stance as well as perceptions of his attitude's moral basis. Participants then learned that Mike was hypocritical on his initial position and voted in favor of using sexualized content. In addition to reporting their surprise at and degree of Mike's hypocrisy, participants made character judgments about him in regard to his moral deceptiveness and false signaling.

## Measures

In both studies, participants responded on similar scales (1 = *Not at all*, 11 = *Completely*). For the *inferred difficulty to change* the target's attitude, participants responded to "How difficult do you think it would be to change [target's] position on the [topic]?" The manipulation check, the target's *perceived moral basis*, was captured with "To what extent do you think [target's] position is based on his sense of morality?"

After learning about the target's hypocritical contradiction, participants reported their *surprise* at the contradiction with "How surprised are you by [target's] vote?" In Study 1a, the degree of *hypocrisy* was captured with two items adapted from prior research (Barden et al., 2005; Kreps et al., 2017): "To what extent is [target] a hypocrite?" and "How hypocritical is [target]?" Because these items were highly correlated ( $r = .96, p < .001$ ), we only collected the latter item in Study 1b. Furthermore, in Study 1b, we also collected two items to account for character judgments, the target's *moral deceit* ("To what extent did you feel Mike was 'morally deceitful' in this scenario?") and *false signaling* ("To what extent did you feel like Mike was 'falsely signaling' his opinion on this topic when he was first speaking?").

## Results and Discussion

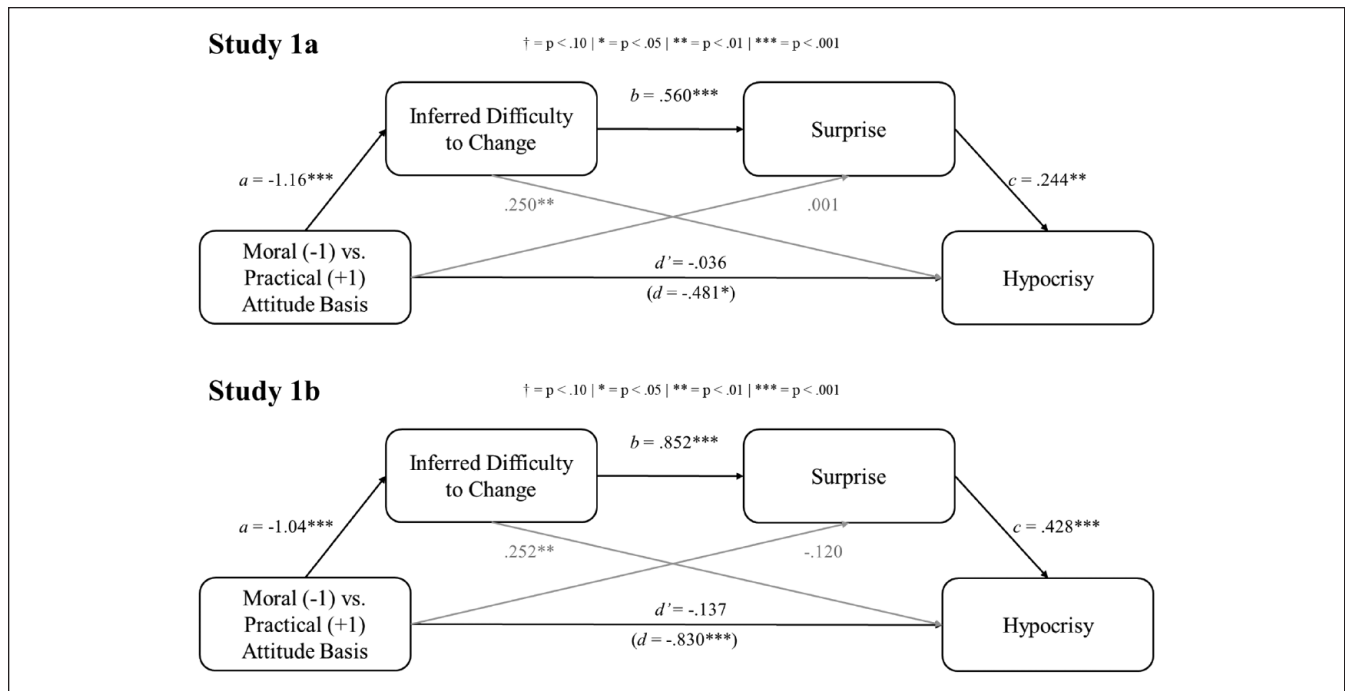
**Study 1a.** The following analyses all come from a two-cell (moral vs. practical) analysis of variance (ANOVA). First, we confirm that our manipulation worked: The target in the moral (vs. practical) condition was perceived as more morally based in his stance,  $M_{\text{moral}} = 9.90, SD = 1.46$  vs.  $M_{\text{practical}} = 4.03, SD = 3.05; F(1, 180) = 267.33, p < .001, 95\%$  confidence interval (CI) = [5.16, 6.57];  $d = 2.46$ .

Replicating prior work (Kreps et al., 2017), hypocrisy on a morally (vs. practicality) based attitude was perceived as more hypocritical,  $M_{\text{moral}} = 6.86, SD = 2.92$  vs.  $M_{\text{practical}} = 5.87, SD = 2.99; F(1, 180) = 5.11, p = .025, 95\%$  CI = [0.126, 1.86];  $d = 0.34$ . In this research, the target described to hold a moral (vs. practical) attitude was indeed inferred as more difficult to change,  $M_{\text{moral}} = 7.86, SD = 2.24$  vs.  $M_{\text{practical}} = 5.55, SD = 2.82; F(1, 180) = 37.15, p < .001, 95\%$  CI = [1.56, 3.06];  $d = 0.91$ , and elicited more surprise when contradicting his attitude,  $M_{\text{moral}} = 8.03, SD = 2.78$  vs.  $M_{\text{practical}} = 6.76, SD = 3.04; F(1, 180) = 8.75, p = .004, 95\%$  CI = [0.426, 2.13];  $d = 0.44$ .

Finally, we examined the plausibility of our theorized relationships, where targets with morally (vs. practicality) based attitudes are perceived as more hypocritical, and these heightened hypocrisy judgments are linked to inferences of greater difficulty to change and thereby increased surprise when change occurs. Using Model 6 of Hayes' (2018) PROCESS macro, this indirect effect was significant ( $B = -.16, SE = .07; 95\%$  CI = [-0.324, -0.033]; Figure 1, top panel).

**Study 1b.** Conducting the same two-cell (moral vs. practical) ANOVA, we again found a successful manipulation: The moral (vs. practical) target's attitude was perceived as more morally based,  $M_{\text{moral}} = 9.75, SD = 1.51$  vs.  $M_{\text{practical}} = 5.56, SD = 3.30; F(1, 206) = 136.57, p < .001, 95\%$  CI = [3.48, 4.90];  $d = 1.63$ . We also found that contradicting a morally (vs. practicality) based attitude was perceived as more hypocritical,  $M_{\text{moral}} = 7.74, SD = 3.01$  vs.  $M_{\text{practical}} = 6.08, SD = 3.05; F(1, 206) = 15.60, p < .001, 95\%$  CI = [0.831, 2.49];  $d = 0.55$ . Replicating the findings of Study 1a, the moral (vs. practical) target was inferred to be more difficult to change,  $M_{\text{moral}} = 8.17, SD = 2.06$  vs.  $M_{\text{practical}} = 6.08, SD = 2.47; F(1, 206) = 43.35, p < .001, 95\%$  CI = [1.46, 2.71];  $d = 0.92$ , and elicited greater surprise at the contradiction,  $M_{\text{moral}} = 8.10, SD = 3.05$  vs.  $M_{\text{practical}} = 6.08, SD = 3.41; F(1, 206) = 20.06, p < .001, 95\%$  CI = [1.13, 2.90];  $d = 0.62$ . Consistent with expectations from past work (Jordan et al., 2017; Kreps et al., 2017), we also found that targets with morally (vs. practicality) based attitudes were judged as more morally deceptive,  $M_{\text{moral}} = 6.23, SD = 3.01$  vs.  $M_{\text{practical}} = 4.68, SD = 2.92; F(1, 206) = 14.15, p < .001, 95\%$  CI = [0.736, 2.36];  $d = 0.52$ , and as engaging in greater false signaling,  $M_{\text{moral}} = 6.85, SD = 2.97$  vs.  $M_{\text{practical}} = 5.60, SD = 3.27; F(1, 206) = 8.31, p = .004, 95\%$  CI = [0.395, 2.10];  $d = 0.40$ .

Using Model 6 of Hayes' (2018) PROCESS macro, we examined the same mediational pattern as in Study 1a (Figure 1, bottom panel). Again, we found a significant indirect effect ( $B = -.38, SE = .10; 95\%$  CI = [-0.592, -0.216]). Moreover, when jointly controlling for perceptions of the target's moral deceptiveness and false signaling, the indirect effect remained intact ( $B = -.07, SE = .03; 95\%$  CI = [-0.166, -0.006]).<sup>3</sup> Altogether, these results support our proposed process for the moral-augmentation effect. In the "General



**Figure 1.** These Figures Represent the Sequential Mediation of Moral Basis to Hypocrisy in Study 1a (Top Panel) and Study 1b (Bottom Panel). Coefficients Displayed are Unstandardized Values.

Discussion” section, though, we further consider the potential role of character judgments in this phenomenon and in hypocrisy findings more broadly.

## Study 2: Surprise and Hypocrisy in a Nonmoral Context

The next study further illustrates that perceived attitude changeability and surprise can enhance judgments of hypocrisy independent of any morality-based attributions. That is, if our proposed mechanism emerges when using a nonmoral predictor of attitude changeability (attitude certainty) and for an instance of hypocrisy with limited to no moral associations (contradicting one’s stance on what dessert to serve), it suggests our explanation for the moral-augmentation effect does not require any latent (i.e., unassessed) moral reasoning strategies. Furthermore, this study demonstrates how our current theorizing can be used to identify novel magnifiers of hypocrisy judgments, namely, expressions of certainty in one’s attitude.

### Participants and Design

This study’s hypotheses, sample size, design, and analyses were all preregistered ([https://aspredicted.org/CQH\\_SJK](https://aspredicted.org/CQH_SJK)). Because this was a novel context, we prepared for a smaller effect size and recruited 300 participants (receiving 301), which provides .80 power (two-tailed,  $\alpha = .05$ ) to detect an effect size approximately 40% smaller ( $d = 0.33$ ) than what was observed

on our focal outcome, surprise, in our prior studies (weighted by sample,  $d = 0.54$ ). Returning to MTurk, we excluded participants based on the preregistered criteria of a topic attention check (uncertain:  $n = 4$ ; certain:  $n = 3$ ) for a final sample of 294 ( $M_{\text{age}} = 38.48$ ,  $SD = 10.88$ , 37.8% female) in a two-cell (attitude certainty: high vs. low) between-subjects design.

### Procedure and Measures

The overall design of this study was similar to Studies 1a/1b; however, the topic for the target’s hypocrisy was very different: a preference between ice cream flavors for a group party. Moreover, instead of varying the moral basis of the attitude, we varied whether the target publicly stated his initial preference for chocolate (vs. vanilla) ice cream with high certainty (e.g., “chocolate is definitely better” and “I’m pretty confident on this”) or low certainty (e.g., “chocolate is maybe better” and “I’m not that confident on this”).

After learning about the target’s ice cream preference for the group party, participants then reported the perceived certainty in the target’s attitude (“How certain do you think Cody is in his opinion?” 1 = *Not at all*, 11 = *Completely*) as well as the perceived difficulty to change his mind and the degree of morality underlying his opinion (on the same items as Study 1b). Next, participants learned of the target’s hypocrisy: He ultimately voted to serve vanilla (instead of chocolate) ice cream at the party. Participants then reported their surprise at this outcome as well as how hypocritical they perceived him to be on the same items as Study 1b.

## Results and Discussion

Conducting a two-cell ANOVA (high vs. low certainty) on the manipulation check, we found that the higher certainty target was indeed rated as more certain,  $M_{\text{high}} = 10.24$ ,  $SD = 1.40$  vs.  $M_{\text{low}} = 4.53$ ,  $SD = 2.92$ ;  $F(1, 292) = 453.41$ ,  $p < .001$ , 95% CI = [5.18, 6.24];  $d = 2.49$ ; however, the targets did not differ in their perceived moral basis,  $M_{\text{high}} = 3.09$ ,  $SD = 2.87$  vs.  $M_{\text{low}} = 2.86$ ,  $SD = 2.76$ ;  $F(1, 292) = 0.49$ ,  $p = .483$ , 95% CI = [-0.415, 0.877];  $d = 0.082$ . The higher certainty target was rated as more difficult to change, though,  $M_{\text{high}} = 9.73$ ,  $SD = 1.82$  vs.  $M_{\text{low}} = 4.34$ ,  $SD = 2.82$ ;  $F(1, 292) = 377.73$ ,  $p < .001$ , 95% CI = [4.85, 5.94];  $d = 2.27$ , and he elicited more surprise when he did,  $M_{\text{high}} = 9.68$ ,  $SD = 2.72$  vs.  $M_{\text{low}} = 5.66$ ,  $SD = 3.35$ ;  $F(1, 292) = 143.85$ ,  $p < .001$ , 95% CI = [3.36, 4.68];  $d = 1.32$ . In line with our theorizing, the target higher in certainty was also rated as more hypocritical,  $M_{\text{high}} = 8.11$ ,  $SD = 3.21$  vs.  $M_{\text{low}} = 5.22$ ,  $SD = 3.37$ ;  $F(1, 292) = 56.60$ ,  $p < .001$ , 95% CI = [2.13, 3.65];  $d = 0.88$ . Notably, these mean ratings of hypocrisy are descriptively similar to what was observed in Studies 1a/1b, suggesting that any quality which enhances the perceived difficulty to change a target's attitude (e.g., moral bases or high certainty) can enhance the surprise at one's hypocrisy, thereby amplifying perceptions of it.

Next, we tested the same sequential mediation as in Studies 1a/1b (Hayes, 2018; model 6), where the effect of condition (high vs. low certainty) on the degree of hypocrisy judgments was sequentially mediated through the perceived difficulty to change the target's attitude and the surprise elicited at changing. Again, we found a significant indirect effect ( $B = .19$ ,  $SE = .04$ ; 95% CI = [0.107, 0.278]). Moreover, if we control for perceptions of the target's morality, the indirect effect does not change ( $B = .19$ ,  $SE = .04$ ; 95% CI = [0.109, 0.281]).

### Study 3: The Joint Effect of Morality and Certainty on Hypocrisy

The prior study highlights how our nonmoral explanation for hypocrisy magnification can occur when moral attitudes are not involved, suggesting our surprise-based process is not facilitated through unassessed moral perceptions. To further support this point, we next manipulate both the target's attitude basis (moral vs. practical) and the target's attitude certainty (high vs. low). If key to the moral-augmentation effect is the inferred strength of the attitude, contradicting an attitude held with certainty (vs. uncertainty) should produce a similar increase in ratings of hypocrisy as contradicting a moral (vs. nonmoral) attitude (as observed descriptively in Studies 1 vs. 2). Thus, the certainty and morality inductions were each expected to result in main effects on hypocrisy judgments; however, an interaction was also feasible if high levels of these variables produced a ceiling effect on perceptions of change (and thus surprise).

## Participants and Design

We received 402 MTurk participants for a 2 (attitude basis: moral vs. practical)  $\times$  2 (attitude certainty: certain vs. uncertain) between-subjects design. In line with our earlier studies, we excluded those who failed a simple attention check of recalling the focal topic (moral-low certainty,  $n = 3$ ; practical-high certainty,  $n = 1$ ; moral-high certainty,  $n = 1$ ) for a final sample of 397 ( $M_{\text{age}} = 38.3$ ; female = 48%). Using the same criterion for the sensitivity analysis as our earlier studies (two-tailed,  $\alpha = .05$ , power = 80%), this sample can detect the predicted main effect differences on surprise as small as  $d = 0.28$ .

## Procedure and Measures

The procedure largely mirrored Study 1a, albeit with an additional manipulation of the target's attitude certainty. That is, participants read one of four possible vignettes describing the target character's attitude basis and certainty in their attitude toward capital punishment: (a) based on morality and high in certainty, (b) based on morality and low in certainty, (c) based on practicality and high in certainty, or (d) based on practicality and low in certainty. The language used to manipulate the target's attitude basis reflected Studies 1a and 1b, while certainty was manipulated using language from Study 2 (see the web appendix for details).

After reading this initial vignette, participants reported their inferences about the difficulty to change the target's attitude as well as perceptions of the target's moral attitude basis and attitude certainty. Participants then learned about the target's hypocritical contradiction and evaluated their surprise at it and the perceived hypocrisy of the target. All items were adapted versions of the items from Study 2.

## Results and Discussion

For the following analyses, all results come from a 2 (attitude basis: moral vs. practical)  $\times$  2 (attitude certainty: high vs. low) ANOVA. First, we note that our manipulations were largely successful: The moral (vs. practical) target was seen as more morally based,  $M_{\text{moral}} = 9.13$  vs.  $M_{\text{practical}} = 4.11$ ;  $F(1, 393) = 377.71$ ,  $p < .001$ , and the high (vs. low) certain target was seen as more certain,  $M_{\text{certain}} = 7.70$  vs.  $M_{\text{uncertain}} = 5.09$ ;  $F(1, 393) = 99.58$ ,  $p < .001$ . These effects were not qualified by interactions ( $ps > .313$ ). At the same time, there was some spillover between manipulations, where the moral target was seen as more certain, and the certain target was seen as more moral (see the web appendix for details). Because these effects were somewhat expected as moral bases and attitude certainty have been linked in prior research (e.g., Luttrell et al., 2016; Skitka et al., 2005), we address this spillover effect as done in prior research by statistically controlling for the alternative manipulation check when testing conditional comparisons (e.g., control for

perceptions of certainty when testing the main effect of morality on judgments of hypocrisy). When we do this, all reported findings remain consistent. To simplify the reporting of the results, we reserve those additional analyses for the web appendix, reporting below the pure tests of condition on our focal outcomes.

Looking at the inferred difficulty to change the targets' attitude, targets with morally (vs. practicality) based attitudes were perceived as more difficult to change,  $M_{\text{moral}} = 6.95$ ,  $SD = 2.73$  vs.  $M_{\text{practical}} = 5.43$ ,  $SD = 2.85$ ;  $F(1, 393) = 34.38$ ,  $p < .001$ , 95% CI =  $[-2.03, -1.01]$ ;  $d = 0.55$ , as were targets expressing more (vs. less) certain attitudes,  $M_{\text{certain}} = 7.22$ ,  $SD = 2.67$  vs.  $M_{\text{uncertain}} = 5.10$ ,  $SD = 2.72$ ;  $F(1, 393) = 67.53$ ,  $p < .001$ , 95% CI =  $[-2.64, -1.62]$ ;  $d = 0.77$ . There was no interaction,  $F(1, 393) = 1.18$ ,  $p = .278$ .

Next, we examined participants' surprise. As observed previously, hypocritical behavior from targets with morally (vs. practicality) based attitudes elicited more surprise,  $M_{\text{moral}} = 8.20$ ,  $SD = 2.74$  vs.  $M_{\text{practical}} = 6.27$ ,  $SD = 3.43$ ;  $F(1, 393) = 40.30$ ,  $p < .001$ , 95% CI =  $[-2.53, -1.33]$ ;  $d = 0.62$ , as did targets expressing more (vs. less) certain attitudes,  $M_{\text{certain}} = 7.93$ ,  $SD = 3.12$  vs.  $M_{\text{uncertain}} = 6.47$ ,  $SD = 3.24$ ;  $F(1, 393) = 23.60$ ,  $p < .001$ , 95% CI =  $[-2.07, -0.878]$ ;  $d = 0.46$ . There was no interaction,  $F(1, 393) = 2.09$ ,  $p = .150$ .

Finally, we examined the degree of hypocrisy ascribed to the target, where hypocritical contradictions for morally (vs. practicality) based attitudes were perceived as more hypocritical,  $M_{\text{moral}} = 7.41$ ,  $SD = 2.96$  vs.  $M_{\text{practical}} = 6.60$ ,  $SD = 3.06$ ;  $F(1, 393) = 7.56$ ,  $p = .006$ , 95% CI =  $[-1.38, -0.230]$ ;  $d = 0.27$ . The same was true for targets with more (vs. less) certain attitudes,  $M_{\text{certain}} = 7.72$ ,  $SD = 2.85$  vs.  $M_{\text{uncertain}} = 6.25$ ,  $SD = 3.02$ ;  $F(1, 393) = 25.46$ ,  $p < .001$ , 95% CI =  $[-2.06, -0.903]$ ;  $d = 0.50$ . These effects were not qualified by an interaction ( $F < 1$ ).

Finally, we tested the viability of our theoretical model, again using Hayes' (2018) PROCESS macro for testing sequential mediation (Model 6). We ran two separate mediation models, first, on the effect of the target's attitude basis (moral vs. practical) on the degree of hypocrisy ascribed to them through the inferred difficulty to change and reactions of surprise at the target's hypocritical behavior. Second, we tested the same indirect effect, replacing the target's attitude basis with the target's attitude certainty (certain vs. uncertain). In both cases, the indirect effects were significant (target's attitude basis:  $B = .35$ ,  $SE = .09$ ; 95% CI =  $[0.186, 0.552]$ ; target's attitude certainty:  $B = .51$ ,  $SE = .12$ ; 95% CI =  $[0.305, 0.757]$ ).<sup>4</sup>

#### Study 4: A Dispositional Trait Reduces Surprise to Reduce Hypocrisy

Across four experiments, we showed how targets with moral (vs. practical) attitudes—just like targets with certain (vs. uncertain) attitudes—are inferred to be more difficult to

change, in turn eliciting more surprise, which was associated with magnified judgments of hypocrisy. In our final studies, we provide evidence for this surprise-based process through experimental moderation (Spencer et al., 2005). That is, if we reduce the expected difficulty to change a target's moral (or certain) attitude, it should attenuate surprise and therefore reduce how hypocritical someone is perceived to be. To that effect, we next describe targets' moral (or certain) attitudes as dispositionally changeable, showing how describing characteristic mutability on these attitudes results in less surprise at a hypocritical contradiction, lowering judgments of their hypocrisy.

#### Participants and Design

We engaged 456 MTurk participants for a 3 (attitude type: moral vs. confident vs. weak)  $\times$  2 (difficulty to change: low vs. control) between-subjects design. In line with our prior studies, we excluded those who failed a simple attention check of recalling the focal topic (which consisted of two items in this study due to the length of the stimuli) or didn't answer the focal outcome measures (moral-control,  $n = 4$ ; moral-changeable,  $n = 4$ ; certain-control,  $n = 4$ ; certain-changeable,  $n = 4$ ; weak-control,  $n = 5$ ; weak-changeable,  $n = 3$ ) for a final sample of 432 ( $M_{\text{age}} = 41.4$ ; female = 46%). Conducting a sensitivity analysis ( $\alpha = .05$ , power = 80%), we could detect an interaction as small as  $d = 0.30$ .

#### Procedure and Measures

Participants were given a brief description about a historical region in Eastern Europe ostensibly known as "Hargov," and then received one of the three sets of focal information. One group of participants read about the people of Hargov (Hargovians') sense of morality (i.e., "their morality represented their core beliefs about what is right and wrong"). Another group read a similar paragraph on Hargovians' belief confidence (i.e., "their confidence represented their certainty and sense of correctness in what they believed"). A final group read about Hargovians' "weak" attitudes (i.e., "their uncertainty represented their sense of doubt in what they believed"). Of these three groups, participants in the "control" condition received no further information than this initial paragraph. However, those in the "low difficulty to change" condition received additional information that described these attitudes as changeable.

More specifically, for changeable participants across the moral, confident, and weak attitudes conditions, they read the following: "When presented with new information or a new perspective, Hargovians would change what they [morally vs. confidently vs. uncertainly] believed quite easily." Notably, in all cases, the core meaning of the attitude type was kept constant (i.e., "their [morality vs. confidence vs. uncertainty] still represented what they fundamentally

thought [was morally right vs. was correct vs. had some doubt about]). Participants then answered the attention checks as well as a manipulation check of how difficult they thought it would be to change a Hargovian's (moral vs. confident vs. uncertain) belief on the same item from Study 3.

Next, participants in all conditions received the exact same information that Hargovians dealt with violent criminals either through banishment or execution. One specific Hargovian, Kasmira, was described to favor banishment over execution, and she held this belief in line with participants' earlier condition (i.e., morally vs. confidently vs. uncertainly). Later in the scenario, Kasmira's community held a vote on what to do with a violent criminal, and Kasmira contradicted her earlier opinion by voting for execution. Participants then reported their surprise at this and how hypocritical they thought she was on the same measures from Study 3.

## Results and Discussion

For the following analyses, we first ran a 3 (attitude type: moral vs. confident vs. weak)  $\times$  2 (difficulty to change: low vs. control) between-subjects ANOVA. However, on every outcome measure, the descriptive and inferential results for the moral and confident conditions (as expected) were statistically equivalent. To simplify the reporting of the results, below we collapse across the moral and confidence conditions to test a 2 (attitude type: "strong" vs. weak)  $\times$  2 (difficulty to change: low vs. control) ANOVA. For the results on the full analysis (which again conceptually replicate everything reported below), see the web appendix.

First, Hargovians were perceived as less difficult to change if they were described as changeable,  $M_{\text{changeable}} = 4.19$  vs.  $M_{\text{control}} = 8.20$ ;  $F(1, 428) = 193.59$ ,  $p < .001$ , 95% CI = [2.93, 3.89];  $d = 1.52$ , or held a weak attitude,  $M_{\text{weak}} = 5.15$  vs.  $M_{\text{strong}} = 6.78$ ;  $F(1, 428) = 40.19$ ,  $p < .001$ , 95% CI = [-2.04, -1.07];  $d = 0.51$ . Importantly, these effects were qualified by the predicted interaction,  $F(1, 428) = 48.49$ ,  $p < .001$ , 95% CI = [-4.38, -2.45]. That is, in the control conditions, there was a greater difference between strong and weak attitudes in their perceived difficulty to change,  $M_{\text{strong}} = 9.26$ ,  $SD = 1.65$  vs.  $M_{\text{weak}} = 6.00$ ,  $SD = 2.68$ ;  $F(1, 428) = 89.83$ ,  $p < .001$ , 95% CI = [-3.94, -2.59];  $d = 1.46$ , compared with those in the changeable conditions,  $M_{\text{strong}} = 4.14$ ,  $SD = 2.51$  vs.  $M_{\text{weak}} = 4.30$ ,  $SD = 3.09$ ;  $F(1, 428) = 0.19$ ,  $p = .662$ , 95% CI = [-0.840, 0.534];  $d = 0.06$ .

A similar pattern of effects unfolded for surprise. Participants were less surprised when Kasmira hypocritically violated a changeable attitude,  $M_{\text{changeable}} = 5.34$  vs.  $M_{\text{control}} = 7.43$ ;  $F(1, 428) = 36.77$ ,  $p < .001$ , 95% CI = [1.17, 2.29];  $d = 0.69$ , or a weak attitude,  $M_{\text{weak}} = 4.82$  vs.  $M_{\text{strong}} = 7.20$ ;  $F(1, 428) = 67.45$ ,  $p < .001$ , 95% CI = [-2.90, -1.78];  $d = 0.80$ . Again, both of these effects were qualified by the predicted interaction,  $F(1, 428) = 12.95$ ,  $p < .001$ , 95% CI = [-3.17, -0.930]; Figure 2, left panel. That is, those in the

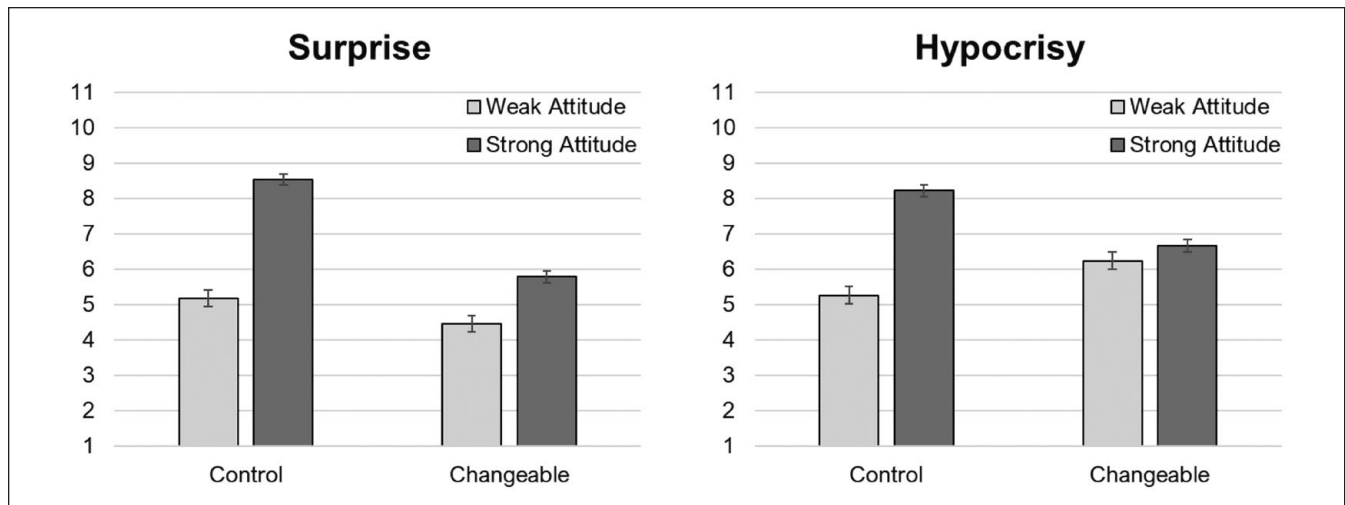
control conditions exhibited a significantly greater difference in surprise between hypocritical violations of a strong and weak attitude,  $M_{\text{strong}} = 8.53$ ,  $SD = 2.30$  vs.  $M_{\text{weak}} = 5.17$ ,  $SD = 2.68$ ;  $F(1, 428) = 70.82$ ,  $p < .001$ , 95% CI = [-4.15, -2.58];  $d = 1.51$ , compared with those in the changeable conditions,  $M_{\text{strong}} = 5.78$ ,  $SD = 3.17$  vs.  $M_{\text{weak}} = 4.46$ ,  $SD = 2.99$ ;  $F(1, 428) = 10.49$ ,  $p = .001$ , 95% CI = [-2.11, -0.516];  $d = 0.43$ .

Finally, the hypocrisy results cohered to predictions. Participants judged Kasmira as less hypocritical when she violated a weak attitude,  $M_{\text{weak}} = 5.74$  vs.  $M_{\text{strong}} = 7.46$ ;  $F(1, 428) = 36.67$ ,  $p < .001$ , 95% CI = [-2.27, -1.12];  $d = 0.58$ . Here, there was no main effect of the changeable condition,  $F(1, 428) = 1.07$ ,  $p = .303$ , but there was an interaction,  $F(1, 428) = 18.70$ ,  $p < .001$ , 95% CI = [-3.67, -1.38]; Figure 2, right panel. Those in the control conditions exhibited a greater difference in judgments of hypocrisy between violations of a strong and weak attitude,  $M_{\text{strong}} = 8.22$ ,  $SD = 2.55$  vs.  $M_{\text{weak}} = 5.26$ ,  $SD = 3.05$ ;  $F(1, 428) = 52.05$ ,  $p < .001$ , 95% CI = [-3.76, -2.15];  $d = 1.05$ , compared with those in the changeable conditions,  $M_{\text{strong}} = 6.66$ ,  $SD = 2.97$  vs.  $M_{\text{weak}} = 6.23$ ,  $SD = 3.03$ ;  $F(1, 428) = 1.08$ ,  $p = .300$ , 95% CI = [-1.25, 0.386];  $d = 0.14$ . Thus, as predicted, enhanced changeability reduced the impact of strong versus weak attitudes on hypocrisy.

Finally, a moderated mediation analysis (Hayes, 2018; Model 7) tested the effect of attitude type (strong vs. weak) on the magnitude of the hypocrisy judgment through feelings of surprise, treating the changeability factor as a moderator for the attitude type's effect on the mediator (surprise). The index of moderated mediation was significant ( $B = -.40$ , boot  $SE = .13$ ; 95% CI = [-0.672, -0.167]), showing that the mediational model was significantly weaker in the changeable condition ( $B = .26$ , boot  $SE = .09$ ; 95% CI = [0.077, 0.440]) compared with the control condition ( $B = .66$ , boot  $SE = .11$ ; 95% CI = [0.448, 0.887]).

## Study 5: A Situational Context Reduces Surprise to Reduce Hypocrisy

In our final study, we take a similar but different methodological approach to Study 4, focusing specifically on moral attitudes. That is, rather than manipulating the attitude's dispositional changeability, we created a situational context that should reduce surprise if someone's moral attitude changed. Although we suspected our induction could lower surprise at a hypocritical behavior for both those with morally as well as practicality-based attitudes (i.e., learning that the target attended an event that provided morally based counterarguments), we expected this effect to be pronounced for those with moral attitudes because "matched" arguments historically result in greater attitude change (Luttrell et al., 2019; Teeny et al., 2021). Thus, we predicted an interaction, where moral (vs. practical) hypocrisy in the control condition should be judged as more hypocritical; however, this difference



**Figure 2.** These Figures Represent the Conditional Results of the 2 (Attitude Type: Weak vs. Strong)  $\times$  2 (Difficulty to Change: Low vs. Control) ANOVA on Feelings of Surprise (Left Panel) and Judgments of Hypocrisy (Right Panel) in Response to the Target's Hypocritical Behavior.

should be attenuated when the surprise at the hypocrisy is led to be similar between moral and practical targets.

### Participants and Design

This study's hypotheses, sample size, design, and analyses were all preregistered ([https://aspredicted.org/RBG\\_S61](https://aspredicted.org/RBG_S61)). For this study, an *a priori* power analysis ( $\alpha = .05$ , power = 80%) for a small interaction ( $d = 0.20$ ) entailed recruiting 787 participants, for which, we recruited 800 MTurk participants (receiving 802) for a 2 (attitude basis: moral vs. practical)  $\times$  2 (surprise: control vs. low) between-subjects design. As before, we excluded those who failed a simple attention check of recalling the focal topic (practical-control,  $n = 10$ ; practical-low,  $n = 7$ ; moral-control,  $n = 10$ ; moral-low,  $n = 9$ ) for a final sample of 766 participants ( $M_{\text{age}} = 39.9$ ; female = 58%).

### Procedure and Measures

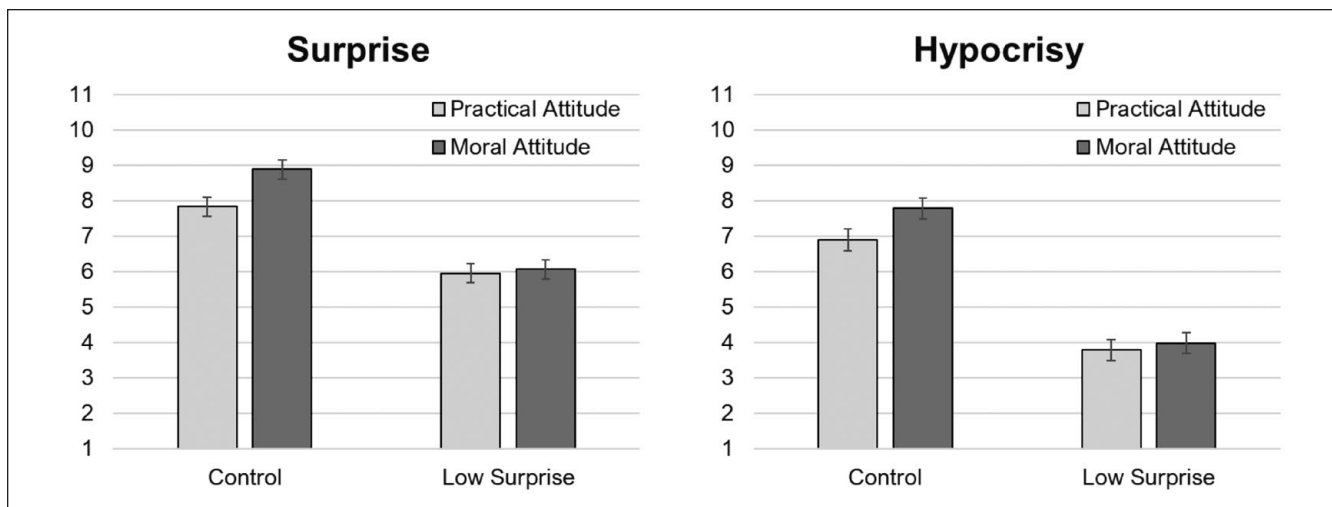
The procedure largely mirrored that of Study 1a, albeit with an additional, surprise-reducing manipulation for half of the participants. First, all participants read either about a target's morally based or practicality-based opposition to capital punishment. For participants in the control condition, they proceeded directly to learning about the target's hypocritical behavior (i.e., voting in support of capital punishment), before reporting their surprise at this and their ratings of how hypocritical the target was perceived to be (on the same items as Study 2). The other half of the participants were provided intervening information about the target.

Specifically, the low-surprise participants were told that a day after Cody expressed his opposition to capital

punishment, he accidentally attended an event with speakers who explained why they morally supported capital punishment. No specific rationale for their position was provided, only that it was morally based (e.g., the speakers supported capital punishment because it was "the most moral and fair thing for society"). This situational context was intended to suggest that Cody might change after the experience, which should reduce participants' surprise if he did. Afterward, participants learned about Cody's hypocritical behavior and reported their surprise at it as well as the degree to which they found it to be hypocritical.

### Results and Discussion

For the following analyses, all results come from a 2 (attitude basis: moral vs. practical)  $\times$  2 (surprise: low vs. control) ANOVA. First, our manipulation of surprise was successful. Although the morally (vs. practicality) based target's hypocrisy elicited more surprise,  $M_{\text{moral}} = 7.54$ ,  $SD = 3.00$  vs.  $M_{\text{practical}} = 6.88$ ,  $SD = 2.81$ ;  $F(1, 762) = 11.23$ ,  $p < .001$ , 95% CI = [0.264, 1.01];  $d = 0.23$ , the low (vs. control) surprise condition resulted in less surprise, too,  $M_{\text{low}} = 5.99$ ,  $SD = 2.81$  vs.  $M_{\text{control}} = 8.43$ ,  $SD = 2.50$ ;  $F(1, 762) = 163.66$ , 95% CI = [-2.80, -2.06];  $p < .001$ ,  $d = 0.92$ . Notably, these main effects were qualified by an interaction,  $F(1, 762) = 8.48$ ,  $p = .004$ , 95% CI = [-1.85, -0.360]; Figure 3, left panel. Decomposing this, we found that morally (vs. practicality) based targets in the control-surprise condition reported greater feelings of surprise,  $M_{\text{moral}} = 9.02$ ,  $SD = 2.34$  vs.  $M_{\text{practical}} = 7.83$ ,  $SD = 2.52$ ;  $F(1, 762) = 19.61$ ,  $p < .001$ , 95% CI = [0.662, 1.72];  $d = 0.49$ . However, this difference was eliminated in the low-surprise condition,  $M_{\text{moral}} = 6.04$ ,  $SD = 2.86$  vs.  $M_{\text{practical}} = 5.95$ ,  $SD = 2.77$ ;  $F(1, 762) = 0.10$ ,  $p = .756$ , 95% CI = [-0.444, 0.611];  $d = 0.03$ .



**Figure 3.** These Figures Represent the Conditional Results of the 2 (Attitude Basis: Moral vs. Practical)  $\times$  2 (Surprise: Low vs. Control) ANOVA on Feelings of Surprise (Left Panel) and Judgments of Hypocrisy (Right Panel) in Response to the Target's Hypocritical Behavior.

For hypocrisy, we replicated the finding that morally (vs. practicality) based hypocrisy was judged as more hypocritical,  $M_{\text{moral}} = 5.96$ ,  $SD = 3.48$  vs.  $M_{\text{practical}} = 5.33$ ,  $SD = 3.23$ ;  $F(1, 762) = 8.66$ ,  $p = .003$ , 95% CI = [0.201, 1.00];  $d = 0.19$ . Here, the low (vs. control) surprise condition also resulted in lower judgments of hypocrisy,  $M_{\text{low}} = 3.85$ ,  $SD = 2.81$  vs.  $M_{\text{control}} = 7.44$ ,  $SD = 2.90$ ;  $F(1, 762) = 306.0$ ,  $p < .001$ , 95% CI = [-3.98, -3.18];  $d = 1.26$ . Again, these two main effects were qualified by the expected interaction,  $F(1, 762) = 5.45$ ,  $p = .020$ , 95% CI = [-1.76, -0.152]; Figure 3, right panel. Specifically, morally (vs. practicality) based targets in the control condition were judged as significantly more hypocritical,  $M_{\text{moral}} = 7.97$ ,  $SD = 2.74$  vs.  $M_{\text{practical}} = 6.89$ ,  $SD = 2.96$ ;  $F(1, 762) = 13.93$ ,  $p < .001$ , 95% CI = [0.512, 1.65];  $d = 0.38$ . However, this difference was eliminated in the low-surprise condition,  $M_{\text{moral}} = 3.91$ ,  $SD = 2.91$  vs.  $M_{\text{practical}} = 3.79$ ,  $SD = 2.71$ ;  $F(1, 762) = 0.19$ ,  $p = .667$ , 95% CI = [-0.444, 0.693];  $d = 0.04$ .

Finally, a moderated mediation model (Hayes, 2018; Model 7) tested the effect of the target's attitude basis (moral vs. practical) on judgments of hypocrisy through surprise, treating the surprise condition as a moderator for the effect of the attitude basis on the mediator (surprise). A significant moderated mediation index emerged (index = -.30, boot  $SE = .11$ ; 95% CI = [-0.511, -0.096]), showing that the mediational analysis was significant in the control condition ( $B = .33$ , boot  $SE = .07$ ; 95% CI = [0.190, 0.468]) but nonsignificant in the low-surprise condition ( $B = .02$ , boot  $SE = .08$ ; 95% CI = [-0.130, 0.184]). Thus, with these final results, we showed how even situational factors that reduce perceivers' surprise at a target's hypocrisy—even in the case of a moral position—can reduce the degree to which the target is perceived as hypocritical.

## General Discussion

The more one is judged a hypocrite, the greater the negative consequences faced (Bhatti et al., 2013; Kougiannou & Wallis, 2020; Laurent et al., 2014). One reliable magnifier of this judgment is when the hypocrisy occurs on a morally (vs. nonmorally) based attitude (e.g., Kreps et al., 2017). In the present work, we extended that finding to a new type of target (i.e., an everyday person/nonleader) and provided a novel, theoretical explanation for it. That is, a common inference of those who hold morally based attitudes is that these attitudes are more difficult to change. This can then result in greater surprise when hypocrisy occurs, amplifying its perception. Offering this new account for the phenomenon, we then showed how this same process could be used to predict when another, nonmoral attribute of a target's attitude would magnify hypocrisy. Specifically, greater perceptions of attitude certainty can similarly enhance hypocrisy judgments through the same surprise-based mechanism. Altogether, then, we believe this research helps to develop an integrative framework for understanding and identifying when both moral and nonmoral factors are likely to heighten the hypocrisy attributed to hypocrites.

### Predicting Magnified Hypocrisy in Other Instances

We believe the processes documented in this research could be useful in helping to understand prior findings in the hypocrisy literature. For example, Alicke and colleagues (2013) found that targets were judged as more hypocritical when contradicting attitudes which the targets had voiced publicly. Violating a public (vs. private) opinion could be more surprising (e.g., because it seemed less likely to change), which

could therefore enhance hypocrisy. Effron and Miller (2015) also found that people were rated as more hypocritical when they preached against deeds for which they had previously benefited (vs. suffered). Advocating against a prior benefit could be more surprising (e.g., because a stance on a personally beneficial outcome should be less likely to change), which could amplify hypocrisy. Importantly, we do not contend that surprise can account for all factors shown to enhance hypocrisy judgments. For example, people from independent (vs. interdependent) cultures tend to exhibit more pronounced hypocrisy effects (Effron et al., 2018); however, it might not be due to differential surprise. Alternatively, interdependent cultures are more accepting of inconsistency (Spencer-Rodgers et al., 2010); thus, surprise at these contradictions might be reduced (Luttrell et al., 2022).

Beyond new interpretations of prior findings, this work suggests novel instances when judgments of hypocrisy could be magnified. Specifically, previous research has shown that when a person's attitude is more extreme, more important, or less ambivalent, the attitude tends to be stronger (e.g., more resistant to change; Luttrell & Sawicki, 2020). However, like morality and certainty, there is little research on how these attitude qualities are perceived in others (Teeny & Petty, 2022). If people make the same inferences as observed for certainty and morality (i.e., that these attitudes are more difficult to change), we would expect contradicting these attitudes to similarly produce greater judgments of hypocrisy. Thus, although the present research focused on morality and certainty, our findings provide a broad range of future directions for when to expect specific instances of hypocrisy will be judged as particularly hypocritical.

Although we have focused on predicting the magnitude of hypocrisy judgments via the effect of attitude characteristics on perceivers' surprise, another consideration is how they affect hypocrisy via changes in the target's perceived disposition. That is, observing someone act hypocritically on a moral or certain attitude could affect broader conclusions about "what kind of person" this target is (cf. Robinson et al., 2017). For example, observing a target hypocritically violate their moral/certain attitude could lead to the perception they are dispositionally dishonest. In turn, this character judgment could increase the hypocrisy ascribed to them in addition to any effects of surprise so far noted. In Study 2, we demonstrated the explanatory power of surprise alongside two such judgments (i.e., moral deceptiveness and false signaling); in Study 5, a specifically situational, surprise-reducing factor still lowered hypocrisy judgments. However, more work is needed to better test and understand the influence of character judgments in these (and other) contexts on the degree of hypocrisy people attribute to others.

### *What Underlies Hypocrisy Itself?*

In the present research, we focused on understanding why similar forms of hypocrisy on morally (vs. nonmorally)

based positions could amplify judgments of it. In providing evidence for our surprise-based process, though, we also observed that hypocrisy judgments can emerge even when contradictions occurred for nonmorally based attitudes.<sup>5</sup> In contrast to some other accounts (e.g., Monin & Merritt, 2012), this suggests that violating a moral attitude specifically is not a prerequisite for the perception of hypocrisy—leading us to the question, what is?

The debate surrounding the "ingredients" of hypocrisy is still ongoing (Laurent & Clark, 2019). For example, some have contended that hypocrisy requires establishing a "standard" of behavior for others that the hypocrite then violates (Lammers, 2012; Valdesolo & DeSteno, 2008). At the same time, separate work has shown this standard might not be necessary (Alicke et al., 2013). As already noted, some research has claimed that moral violations are required for judgments of hypocrisy, whereas the present research suggests they are not (see also, Laurent & Clark, 2019). Interestingly, the only element that hypocrisy scholars universally seem to agree upon is that the target must demonstrate "inconsistency" between an attitude, behavior, or some combination of the two (e.g., what is often referred to as "word-deed misalignment"; Effron et al., 2018; Huppert et al., 2023; Laurent & Clark, 2019). In which case, it is the nature and purpose of the contradiction that then becomes vital in determining the extent or frequency with which attributions of "hypocrisy" are applied. In the present work, our argument is only that when the elements for hypocrisy exist (whatever those might be), surprise can magnify hypocrisy judgments. For example, if a politician wore a goofy hat one day (a surprising behavior), it would unlikely be seen as hypocritical. However, if the politician flip-flopped positions on an issue (a relatively unsurprising behavior because politicians are known to do that), it would likely seem hypocritical. Our point is that if this latter politician's position change *was* made to be more surprising (e.g., because absolute certainty in the position had been expressed), it would likely be judged as even more hypocritical. In sum, we believe there are many interesting questions about precisely what makes a person come to be seen as a hypocrite, and from the current work, we now understand that when those elements of hypocrisy *are* present, surprise will magnify judgments of it.

### *Limitations and Conclusions*

Although we aimed to be rigorous in our theoretical and empirical approach, we nonetheless acknowledge some limitations. First, all our participants were from the United States and online-based samples. Thus, if people from different cultures or backgrounds are unlikely to infer that those with moral (or certain) attitudes are difficult to change, we would be unlikely to observe the processes documented in this research. Second, although we provided evidence for a surprise-based, nonmoral process in the moral-augmentation effect, we do not definitively test whether this mechanism

“better accounts” for the phenomenon relative to other, morally based explanations put forth (Jordan et al., 2017; Kreps et al., 2017). Thus, we want to emphasize only that the mechanism we documented was supported even after controlling for related, moral judgments and emerged in contexts where contradictions of moral attitudes were not plausibly relevant.

Finally, although we generally showed that morality and certainty exhibited similar effects on hypocrisy, we do not claim they generated “equivalent” effects. There could very well be contexts where one type of attitude would be more likely than the other to elicit greater hypocrisy judgments (or even different, downstream consequences). Thus, all we argue is that one aspect of morality (its perceived difficulty to change) is comparable to but not redundant with attitude certainty when it comes to the processes we outlined; however, we do not rule out the possibility that other factors specific to either type of attitude could be important, too. In closing, we hope the present research has provided a new and valuable perspective for both the literatures on hypocrisy as well as the moral bases of attitudes, as we believe both areas of research are ever more important for understanding the world’s social-cultural climate ahead.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: All funding for this research came directly from research funds distributed by the authors’ respective institutions.

### ORCID iDs

Jacob D. Teeny  <https://orcid.org/0000-0001-6112-534X>

Richard E. Petty  <https://orcid.org/0000-0002-2870-8575>

### Supplemental Material

Supplemental material is available online with this article.

### Notes

1. Kreps et al. (2017) post hoc assessed participants’ surprise at targets’ contradictions in a few studies and found that moral violations were indeed more surprising than practical ones; however, they did not treat surprise as a mediator for the enhanced hypocrisy judgments (as we do) but as a predictor (along with hypocrisy) of support for the target.
2. To further confirm this assumption, we conducted another pilot test ( $n = 195$ ) using the same vignettes as the moral attitudes pilot. Here, we described the targets as holding their stance with certainty versus uncertainty. As expected, certain (vs. uncertain) attitudes were perceived as more difficult to change. See the web appendix.

3. Mediation analyses testing the effects of these character judgments as alternative mediators (i.e., the moral-practical condition on [moral deceptiveness or false signaling] on hypocrisy judgments) showed both patterns were significant. This provides supportive evidence for past theorizing and their role in the moral-augmentation effect. However, if we additionally control for our proposed mediators (i.e., difficulty to change and surprise) either separately or jointly, these alternative mediational tests become nonsignificant. See the web appendix for details.
4. Controlling for the nonfocal perception in the reported mediational models (i.e., controlling for certainty when testing the effect of target’s attitude basis; controlling for perceived moral bases when testing the effect of the target’s attitude certainty) does not change the significance of these mediational analyses. See the web appendix.
5. In addition to the results of Study 2 (which used a nonmoral manipulation of hypocrisy magnification [attitude certainty] and a nonmoral topic [a preference between ice cream flavors]), we can also examine the hypocrisy judgments in the “nonmoral” conditions of Studies 3 to 5 (i.e., when the target expressed an entirely practical stance on the topic). In line with Study 2, all these analyses revealed differences in hypocrisy magnitude as a function of attitude certainty, even when the basis for the opinion was purely nonmoral (see the web appendix for details).

### References

- Alicke, M., Gordon, E., & Rose, D. (2013). Hypocrisy: What counts? *Philosophical Psychology*, *26*(5), 673–701. <https://doi.org/10.1080/09515089.2012.677397>
- Aramovich, N. P., Lytle, B. L., & Skitka, L. J. (2012). Opposing torture: Moral conviction and resistance to majority influence. *Social Influence*, *7*(1), 21–34. <https://doi.org/10.1080/15534510.2011.640199>
- 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>
- 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.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Academic Press. <https://doi.org/10.4324/9780203771587>
- Durso, G. R. O., Petty, R. E., Briñol, P., Siev, J. J., Hinsenkamp, L., & Sawicki, V. (2021). Dampening affect via expectations: The case of ambivalence. *Journal of Personality and Social Psychology*, *121*(6), 1172–1194. <https://doi.org/10.1037/pspa0000248>
- 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*, *76*, 371–384. <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>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <https://doi.org/10.3758/BF03193146>
- Foster, M. I., & Keane, M. T. (2015). Why some surprises are more surprising than others: Surprise as a metacognitive sense of explanatory difficulty. *Cognitive Psychology*, 81, 74–116. <https://doi.org/10.1016/j.cogpsych.2015.08.004>
- 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>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Huppert, E., Herzog, N., Landy, J., & Levine, E. (2023). On being honest about dishonesty: The social costs of taking nuanced (but realistic) moral stances. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspa0000340>
- 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>
- Kougiannou, N. K., & Wallis, M. O. (2020). Chimneys don't belch out carnations! The (in)tolerance of corporate hypocrisy: A case study of trust and community engagement strategies. *Journal of Business Research*, 114, 348–362. <https://doi.org/10.1016/j.jbusres.2019.08.029>
- 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>
- Kreps, T. A., & Monin, B. (2014). Core values versus common sense: Consequentialist views appear less rooted in morality. *Personality and Social Psychology Bulletin*, 40(11), 1529–1542. <https://doi.org/10.1177/0146167214551154>
- 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>
- Laurent, S. M., & Clark, B. A. (2019). What makes hypocrisy? Folk definitions, attitude/behavior combinations, attitude strength, and private/public distinctions. *Basic and Applied Social Psychology*, 41(2), 104–121. <https://doi.org/10.1080/01973533.2018.1556160>
- Laurent, S. M., Clark, B. A., 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 & Emotion*, 28(1), 59–83. <https://doi.org/10.1080/02699931.2013.801339>
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, 49(2), 433–442.
- Luttrell, A., Petty, R. E., Briñol, P., & Wagner, B. C. (2016). Making it moral: Merely labeling an attitude as moral increases its strength. *Journal of Experimental Social Psychology*, 65, 82–93. <https://doi.org/10.1016/j.jesp.2016.04.003>
- Luttrell, A., Petty, R. E., Chang, J., & Togans, L. J. (2022). The role of dialecticism in objective and subjective ambivalence. *British Journal of Social Psychology*, 61(3), 826–841. <https://doi.org/10.1111/bjso.12504>
- Luttrell, A., Philipp-Muller, A., & Petty, R. E. (2019). Challenging moral attitudes with moral messages. *Psychological Science*, 30, 1136–1150. <https://doi.org/10.1177/0956797619854706>
- Luttrell, A., & Sawicki, V. (2020). Attitude strength: Distinguishing predictors versus defining features. *Social and Personality Psychology Compass*, 14(8), 1–16. <https://doi.org/10.1111/spc3.12555>
- Maguire, R., Maguire, P., & Keane, M. T. (2011). Making sense of surprise: An investigation of the factors influencing surprise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 176–186. <https://doi.org/10.1037/a0021609>
- Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, 8(6), 423–429. <https://doi.org/10.1111/j.1467-9280.1997.tb00455.x>
- 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). American Psychological Association. <https://doi.org/10.1037/13091-009>
- Noordewier, M. K., Topolinski, S., & Van Dijk, E. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass*, 10(3), 136–149. <https://doi.org/10.1111/spc3.12242>
- O'Connor, K., Effron, D. A., & 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), 540–559. <https://doi.org/10.1037/pspa0000195>
- Petrocelli, J. V., Clarkson, J. J., Tormala, Z. L., & Hendrix, K. S. (2010). Perceiving stability as a means to attitude certainty: The role of implicit theories of attitudes. *Journal of Experimental Social Psychology*, 46(6), 874–883. <https://doi.org/10.1016/j.jesp.2010.07.012>
- Petty, R. E. & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences*. Lawrence Erlbaum.
- Philipp-Muller, A. Z., Wallace, L. E., & Wegener, D. T. (2020). Where does moral conviction fit? A factor analytic approach examining antecedents to attitude strength. *Journal of Experimental Social Psychology*, 86, Article 103900. <https://doi.org/10.1016/j.jesp.2019.103900>
- Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, 11(1), 50–74. <https://doi.org/10.1111/tops.12292>
- Robinson, J. S., Page-Gould, E., & Plaks, J. E. (2017). I appreciate your effort: Asymmetric effects of actors' exertion on observers' consequentialist versus deontological judgments. *Journal of Experimental Social Psychology*, 73, 50–64. <https://doi.org/10.1016/j.jesp.2017.06.005>
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. Pinter & Martin Publishers.
- Rucker, D. D., Tormala, Z. L., Petty, R. E., & Briñol, P. (2014). Consumer conviction and commitment: An appraisal-based framework for attitude certainty. *Journal of Consumer Psychology*, 24(1), 119–136. <https://doi.org/10.1016/j.jcps.2013.07.001>

- Shepperd, J. A., & McNulty, J. K. (2002). The affective consequences of expected and unexpected outcomes. *Psychological Science, 13*(1), 85–88. <https://doi.org/10.1111/1467-9280.00416>
- Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology, 88*(6), 895–917. <https://doi.org/10.1037/0022-3514.88.6.895>
- Skitka, L. J., & Morgan, G. S. (2014). The social and political implications of moral conviction. *Political Psychology, 35*(1), 95–110. <https://doi.org/10.1111/pops.12166>
- 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>
- Spencer-Rodgers, J., Peng, K., & Wang, L. (2010). Dialecticism and the co-occurrence of positive and negative emotions across cultures. *Journal of Cross-Cultural Psychology, 41*(1), 109–115. <https://doi.org/10.1177/0022022109349508>
- Teeny, J. D., & Petty, R. E. (2022). Attributions of emotion and reduced attitude openness prevent people from engaging others with opposing views. *Journal of Experimental Social Psychology, 102*, Article 104373. <https://doi.org/10.1016/j.jesp.2022.104373>
- Teeny, J. D., Siev, J. J., Briñol, P., & Petty, R. E. (2021). A review and conceptual framework for understanding personalized matching effects in persuasion. *Journal of Consumer Psychology, 31*(2), 382–414. <https://doi.org/10.1002/jcpy.1198>
- Thibaut, J. W., & Kelley, H. (1959). *The social psychology of groups*. John Wiley. <https://doi.org/10.4324/9781315135007>
- Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*(5), 1334–1338. <https://doi.org/10.1016/j.jesp.2008.03.010>