

Journal of Personality and Social Psychology

On Being Honest About Dishonesty: The Social Costs of Taking Nuanced (but Realistic) Moral Stances

Elizabeth Huppert, Nicholas Herzog, Justin F. Landy, and Emma Levine

Online First Publication, March 6, 2023. <https://dx.doi.org/10.1037/pspa0000340>

CITATION

Huppert, E., Herzog, N., Landy, J. F., & Levine, E. (2023, March 6). 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://dx.doi.org/10.1037/pspa0000340>

On Being Honest About Dishonesty: The Social Costs of Taking Nuanced (but Realistic) Moral Stances

Elizabeth Huppert^{1, 2}, Nicholas Herzog¹, Justin F. Landy³, and Emma Levine¹

¹ Booth School of Business, University of Chicago

² Kellogg School of Management, Northwestern University

³ Department of Psychology and Neuroscience, Nova Southeastern University

Despite the well-documented costs of word–deed misalignment, hypocrisy permeates our personal, professional, and political lives. Why? We explore one potential explanation: the costs of moral flexibility can outweigh the costs of hypocrisy, making hypocritical moral absolutism a preferred social strategy to admissions of moral nuance. We study this phenomenon in the context of honesty. Across six studies (total $N = 3545$), we find that communicators who take flexible honesty stances (“It is sometimes okay to lie”) that align with their behavior are penalized more than hypocritical communicators who take absolute honesty stances (“It is never okay to lie”) that they fail to uphold. Although few people take absolute stances against deception themselves, they are more trusting of communicators who take absolute honesty stances, relative to flexible honesty stances, because they perceive absolute stances as reliable signals of communicators’ likelihood of engaging in future honesty, regardless of inconsistent behavior. Importantly, communicators—including U.S. government officials—also anticipate the costs of flexibility. This research deepens our understanding of the psychology of honesty and helps explain the persistence of hypocrisy in our social world.

Keywords: hypocrisy, honesty, moral judgment, economic games, communication

Supplemental materials: <https://doi.org/10.1037/pspa0000340.supp>


Publicly espousing one’s absolute commitment to honesty is common. We teach our children that “honesty is the best policy,” we require our students to take oaths of academic honesty, and we assure our spouses that we would never deceive them. Yet, lying is a common behavior (DePaulo & Kashy, 1998) and, under some circumstances, it is even considered more ethical to lie than to tell the truth (Levine, 2021; Levine & Schweitzer, 2014). Thus, there is a puzzling inconsistency in everyday behavior: people frequently lie but publicly endorse an absolute moral stance that forbids this very behavior. This paradox implies that at least some people are endorsing a moral standard that they do not live up to in daily life. In other words, they are taking hypocritical moral stances on honesty.


Existing work on hypocrisy suggests that publicly preaching the importance of honesty, while also privately behaving dishonestly, should bring social costs. Although the endorsement of moral norms is generally well received, hypocrites who fail to practice what they

preach are disliked. In fact, hypocrites are often perceived more negatively than people who never endorsed a moral norm in the first place (Effron, O’Connor, et al., 2018). Hypocrisy may be viewed as particularly bothersome because behavior that violates a previous statement leads to the inference that one’s words were spoken in vain. Impressions of false signaling, or “cheap talk,” often have more adverse consequences than ignoring a norm entirely (Jordan et al., 2017; Jordan & Sommers, 2020).

Given the costs of being labeled a hypocrite, one seemingly reasonable strategy when talking about honesty is to acknowledge that honesty is complicated and lying is sometimes permissible. In this research, we find that many people do admit to lying in everyday life. Therefore, a flexible honesty stance is likely to align with a communicator’s actual behavior and their underlying beliefs. Furthermore, we find that very few people actually believe that lying is never okay (see Pilot Studies 1–3 and Studies 4 and 5).

Elizabeth Huppert  <https://orcid.org/0000-0003-0643-315X>

Justin F. Landy  <https://orcid.org/0000-0002-4082-9505>

Emma Levine  <https://orcid.org/0000-0001-7631-4231>

The authors thank attendees of the Center for Decision Research brownbag and members of the Honesty Opportunity Prosociality and Ethics lab at the University of Chicago Booth School of Business for feedback on this project. They thank Michael White and Solomon Lister for research assistance and the Center for Decision Research for support in data collection. They also thank Ayelet Fishbach and Bernd Wittenbrink for comments on the article. Finally, they thank the Charles E. Merrill Faculty Scholarship for financial support.

Elizabeth Huppert played a lead role in the formal analysis and writing of

original draft and an equal role in conceptualization, methodology, and writing of review and editing. Nicholas Herzog played a supporting role in conceptualization, data curation, formal analysis, and writing of original draft and an equal role in the writing of review and editing. Justin F. Landy played a supporting role in conceptualization, formal analysis, methodology, and writing of original draft and an equal role in writing of review and editing. Emma Levine played a lead role in funding acquisition, a supporting role in formal analysis and writing of original draft, and an equal role in conceptualization, methodology, and writing of review and editing.

Correspondence concerning this article should be addressed to Elizabeth Huppert, Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, United States. Email: elizabeth.huppert@kellogg.northwestern.edu

If acknowledging nuance about the morality of dishonesty is both an accurate expression of one's behaviors and a widely held position on honesty, this could allow communicators to avoid the social costs of hypocrisy when discussing honesty.

However, in the present research, we find that the social costs of taking such realistic, flexible moral stances can be greater than the costs of taking hypocritical, absolute stances. Although endorsing moral flexibility may align with a communicator's actual behavior, and often their private beliefs, doing so also signals lower commitment to moral values and a greater likelihood of dishonest behavior in the future, which can be more costly than hypocrisy. Consider the backlash Hillary Clinton faced after openly discussing the practice of speaking differently to different audiences on the campaign trail during the 2016 U.S. presidential election. Unlike most politicians, Clinton did not take an absolute stance against the practice of varying her messaging across contexts. Rather, Clinton admitted that she believed it was necessary to provide different audiences with different messages to achieve political goals, for example, by publicly proclaiming the importance of banking reform, while privately making alliances with the industry. Clinton came under attack for taking a morally flexible stance regarding consistency and honesty in political messaging. Despite being honest about her behavior and the nuances of political messaging, Clinton was labeled a liar. It seems that voters would have preferred that she be dishonest about the nature of political dishonesty.¹

Clinton's situation is not unique. Outside the political domain, leaders and relational partners frequently confront situations in which they are required to take stances on moral issues. For example, CEOs may be asked about the acceptability of questionable, but common, business practices, like gender wage disparities within a company or the use of fossil fuels in manufacturing pipelines. In these situations, individuals must decide whether to commit to a lofty, but likely untenable, moral stance or be more flexible and realistic in their messaging. Likewise, people may be asked about their moral beliefs explicitly in interpersonal conversations. Although taking an absolute moral stance that one fails to uphold may be penalized relative to taking no stance at all (Jordan et al., 2017), it is not yet clear how people judge hypocritical moral absolutism relative to realistic moral flexibility. We study this question specifically in the domain of honesty because this is a domain in which many people *do* privately hold flexible moral views themselves, making flexible stances truthful representations of private beliefs and behaviors.

The present research sheds new light on the psychology of hypocrisy, honesty, and moral flexibility by comparing judgments of, and behavior toward, people who make absolute versus flexible proclamations about honesty and then behave dishonestly. We define *absolute* moral proclamations as statements that renounce any and all violations of a moral norm and leave no room for ambiguity. For example, a stance that says that it is never okay to lie, no matter the circumstances, is an absolute honesty stance. We define *flexible* (or *nuanced*) moral proclamations as statements that acknowledge moral rules can be complicated. Some behaviors that are often considered wrong, such as lying, may be viewed positively at times, such as prosocial lies. A stance that acknowledges that it is sometimes acceptable to lie is a flexible honesty stance.

We consider how people judge communicators who proclaim, "It is never okay to lie," (absolute) and communicators who proclaim, "It is sometimes okay to lie" (flexible), and then go on to tell equivalent lies. Using both hypothetical vignettes and economic games, with both laypeople and government officials, we find that communicators are often penalized more for taking flexible stances that are consistent with their behavior than for taking absolute stances on honesty that they then fail to uphold. We also find that communicators who take flexible stances on honesty are judged more negatively than communicators who take no stance on honesty, highlighting the costs of taking flexible, yet realistic, moral stances. Importantly, communicators predict these costs. Though we find that most people do hold flexible stances on honesty in private, they also intuit that it is costly to acknowledge these stances in public. Consequently, communicators are likely to inflate their honesty beliefs in public discourse, which can contribute to the persistence of widespread hypocrisy.

Theory

People recognize that honesty is a complicated moral value. People are frequently dishonest (DePaulo & Kashy, 1998), and they often believe that their own and others' dishonesty is justified (Hildreth & Anderson, 2018; Levine & Schweitzer, 2014). For example, people believe it is ethical to tell lies that prevent unnecessary harm, even when they consider being the recipient of such lies (Levine, 2021). Although a growing body of work documents people's nuanced moral judgments of (dis)honesty, the consequences of discussing these nuances remain unclear. When people publicly discuss lying, what are the relative consequences of committing to unrealistic absolutism versus openly acknowledging nuance? Given that most people do lie at times, we propose that there is a trade-off between being labeled a hypocrite and being seen as having low moral standards. We therefore consider the social costs of both hypocrisy and flexibility by examining how both influence moral judgment and trust.

The Costs of Hypocrisy

As previously discussed, hypocrisy typically incurs social costs (Efron et al., 2015; O'Connor et al., 2020). A communicator, for example, who claims that illegally downloading music is wrong, is judged as more unethical for illegally downloading music than a communicator who made no such moral claim before engaging in the same illegal behavior (Jordan et al., 2017). Although committing an illegal act is generally viewed negatively, there is an additional hypocrisy penalty for condemning the same act before committing it, over and above the general disapproval of the act itself. When people behave hypocritically, their words can be discounted completely and perceived as false signals that were used strategically to claim undeserved moral credit. This dynamic is well-documented in a variety of domains, from steroid use to plagiarism to speeding while driving (Efron, O'Connor, et al., 2018; Jordan et al., 2017). Building on this work, we might expect that a communicator who

¹ Consistent with this idea, conservative voters celebrated Donald Trump's condemnation of Clinton's lies, despite likely recognizing that Trump was often dishonest too (article: <https://www.vox.com/2016/10/7/13207286/clinton-speech-transcripts-wikileaks-email>).

claims that it is never okay to lie and then lies is judged more negatively than a communicator who lies but makes no such claim.

Recent work, however, suggests that hypocrisy penalties are not inevitable (Jordan & Sommers, 2022). Inconsistent communicators who admit to their own moral failings—for example, communicators who claim that it is wrong to download music illegally, but who admit that they sometimes do it anyways—are viewed more positively than “traditional” hypocrites who do not acknowledge their own inconsistency between their words and deeds (Jordan et al., 2017). Importantly, these communicators who fail to live up to their own standards *are* seen as more hypocritical than communicators who make no moral claims, but they are not seen as attempting to claim undeserved moral credit. In general, when a communicator’s hypocrisy is attributed to personal moral failings, rather than false signaling, observers tend to judge the word–deed inconsistency less harshly (Jordan & Sommers, 2020). These results suggest that other inferences can sometimes counteract judgments of hypocrisy, making word–deed misalignment seem less immoral than it otherwise would.

The Costs of Flexibility

Although hypocrisy is often costly, particularly when compared to taking no stance at all, there are circumstances in which one *must* take a stance. During conflicts, for example, people are often forced to take sides (Shaw et al., 2017). Similarly, there is often pressure to take moral stances during personal and professional conversations, rather than to stay neutral (Silver & Shaw, 2022). Therefore, to fully understand how to navigate trade-offs involving moral stance-taking, we must also understand how people judge hypocrites relative to those who take more measured, flexible stances.

In general, taking a strong stance on a social issue, or *moralizing* an issue, signals principledness (Kreps et al., 2017; Kreps & Monin, 2014; Van Zant & Moore, 2015), which increases judgments of moral character (Zlatev, 2019). Conversely, taking a more flexible stance on a social issue may signal a lack of genuine concern for the issue, which undermines judgments of morality (Zlatev, 2019). Recent work highlights the costs of moral flexibility in the domain of honesty in particular. Even though most people believe that lies that prevent interpersonal harm are ethical, people tend to penalize others who put themselves in situations that would allow them to tell these lies. Specifically, in a series of economic games, Jensen et al. (2021) find that people judge communicators who avoid finding out how their honesty will affect others (i.e., whether it will help or harm a partner) as more moral than communicators who seek out this information before making communication decisions. This preference for blind honesty is driven by the belief that communicators who seek out information are more likely to tell lies in general. Even if a communicator looks for more information to tell a prosocial lie (rather than a selfish truth), which is seen as moral (Levine & Schweitzer, 2014), observers worry that the communicator’s flexibility may lead them to tell harmful lies as well. Beliefs about a communicator’s propensity to tell harmful lies in the future can significantly undermine moral character evaluations (Jensen et al., 2021). Similarly, observers may view a flexible stance on honesty as a signal that the communicator is more likely to lie—in both moral and immoral ways—in the future. In other words, flexible moral stances may be perceived as credible signals of dishonest future behavior.

This proposition is consistent with research on truth bias. In general, people tend to interpret others’ claims as true (Levine, 2014). Therefore, people are likely to believe that a communicator who takes a flexible stance on honesty sees dishonesty as more justified than a communicator who takes an absolute stance on honesty. By the same logic, flexible honesty stances are likely seen as signaling lower commitment to honesty and a lower likelihood of behaving honestly in the future, relative to absolute honesty stances. Even if a communicator lies after taking an absolute honesty stance, which should lead to perceptions of hypocrisy and consequently detract from expectations of future honesty, we posit that they will still be perceived as caring more about honesty in general than someone who takes a flexible stance on honesty and also lies.

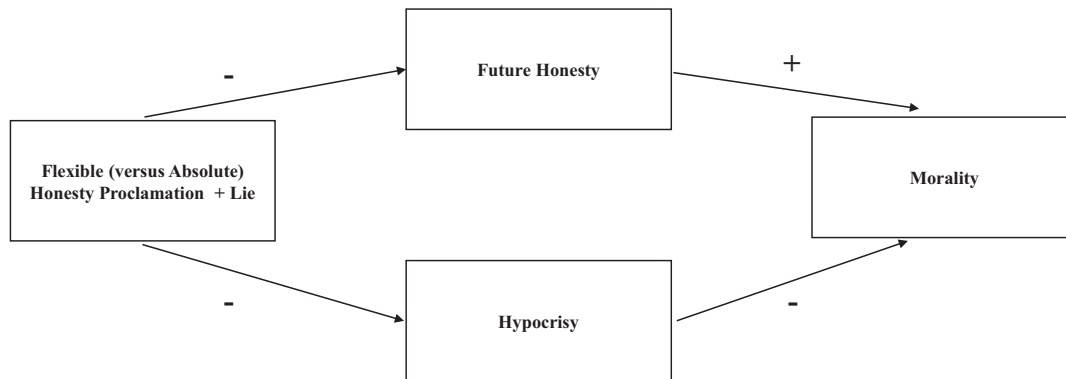
Summary

In the current research, we consider the reputational trade-offs associated with taking absolute, yet hypocritical, stances, versus flexible, yet realistic, stances on honesty. We develop our understanding of moral communication by highlighting an important, yet ignored, force that can overwhelm the costs of hypocrisy: moral flexibility. The goal of the present research is to understand how these different communication strategies influence moral judgments and associated behaviors. We find that admission of moral nuance can reflect even *more* negatively on moral character than hypocrisy itself. Although there may still be situations in which hypocrisy is more costly than flexibility, highlighting the severe costs of moral flexibility is critical for understanding the persistence of hypocrisy in the social world.

Simply put, when communicators must take a stance on the morality of lying, they face a trade-off: absolute stances can bring a hypocrisy penalty if the communicator is later caught lying, but flexible stances bring their own cost, because they are seen as signaling that the communicator is more likely to behave dishonestly in general. In other words, most people do not have a costless way of discussing untenable moral standards, such as absolute honesty. In at least some cases, the cost of flexibility outweighs the cost of hypocrisy, making flexible but realistic stances an even worse social strategy than absolute but hypocritical stances. We depict our theoretical account in Figure 1.

Ultimately, this research helps to explain why hypocrisy is so pervasive in public rhetoric and private interactions, particularly around the value of honesty. According to evolutionary partner choice models of morality (Hoffman et al., 2015; Nowak, 2006a, 2006b; Panchanathan & Boyd, 2004), default patterns of behavior converge over time to match a socially rewarded equilibrium. Existing research suggests that hypocrisy should not arise as a default behavior because of the associated social costs. If hypocrisy is perceived to signal a lack of morality, hypocritical targets should be excluded from cooperative exchanges over time (Barclay, 2004, 2006; Rockenbach & Milinski, 2011). However, this does not seem to be the case; people continue to elect leaders and support individuals and organizations that exhibit word–deed inconsistency (e.g., Kim et al., 2021). Our findings help explain this puzzle by highlighting the costs of moral flexibility, which have been largely ignored in previous hypocrisy research (Effron, O’Connor, et al., 2018; Jordan & Sommers, 2022). As a result, this work helps to explain a social tolerance for hypocritical absolutism, broadly, and to explain inconsistencies in public versus private attitudes toward

Figure 1
A Theoretical Model of How Absolute and Flexible Honesty Stances Influence Moral Judgments



honesty, in particular. Although people do dislike hypocrites, they also strongly dislike realists who acknowledge moral flexibility about honesty.

Overview of the Present Studies

We test our theoretical account across six experiments reported in the main text, as well as four pilot studies and seven supplemental studies (see Online Supplemental Material). All data collection was approved by the institutional review board at The University of Chicago. Five of our experiments in the main text and five of our supplemental studies were preregistered (see Supplemental Material Appendix 4, for all study-specific preregistration links on <https://AsPredicted.org>). Stopping rules for each of our preregistered studies in the main text (Studies 1–5) were decided in advance. Based on our past research, we aimed to recruit 100–150 participants per cell in each of these studies. Our samples had statistical power of greater than 0.99 in Studies 1–4 and 0.92 in Study 5 to detect our critical effect comparing morality judgments of communicators who take absolute versus flexible stances on honesty. These power estimates were computed using post hoc power analyses for independent-samples *t* tests with $\alpha = .05$ in *G**power (Faul et al., 2007). For Studies 2 and 5, we used moral evaluations at Time 2 only. Cohen's *d*s for the effect of the absolute versus flexible proclamation on moral evaluations were 0.49, 0.71, 0.42, 0.67, and 0.34 for Studies 1, 2, 3, 4, and 5, respectively.

Almost all of our studies (17 in total) began with informed consent and ended with basic demographic questions. Study 6 and Pilot Study 3 were the only exceptions, since these survey studies were administered through the nonprofit CivicPulse, and we followed their recruitment and data collection processes. All participants were compensated for their time.

Study 1 uses a hypothetical scenario about a dishonest politician to provide an initial examination of how people evaluate communicators who take flexible and absolute honesty stances and then lie. We find that, relative to a candidate who takes a hypocritical absolute honesty stance, people are less willing to vote for a candidate who takes a flexible honesty stance, and they judge this candidate as less moral. We then directly disentangle the costs of flexibility from the costs of hypocrisy in Studies 2 and 3. In these studies, participants evaluate communicators after learning their honesty stance and again after learning about the communicator's deceptive behavior. In Study 2, we

also add a control condition to compare how judgments of communicators who take absolute and flexible stances differ from judgments of communicators who take no honesty stance at all. In Study 3, we vary the order of proclamations and behavior to determine whether absolute and flexible stances are perceived differently when they follow, rather than precede, a lie. In these studies, we identify costs of hypocrisy, such that communicators who endorse absolute honesty and lie are seen as less moral and less likely to engage in future honesty after their deceptive behavior is revealed. However, these communicators are still seen as more moral and more likely to engage in future honesty overall than the communicators who endorse flexible honesty. The costs of taking a flexible stance outweigh the costs of hypocrisy.

We explore the robustness of these effects in Studies 4 and 5. Specifically, we examine how absolute and flexible stances are evaluated when they are followed by honest behavior (moral consistency), and when they are followed by different types of lies (prosocial vs. selfish). Overall, we find that the costs of taking a flexible, rather than an absolute, stance do not seem to depend on the behavior that follows it. In Studies 4 and 5, we also find that the reputational costs of moral flexibility influence interpersonal trust. Furthermore, we examine judgments of ambitious but nonabsolute moral stances (i.e., "Lying is rarely okay"), which might reflect thoughtful exceptions to behaving honestly, and thus better balance the costs of flexibility and the costs of hypocrisy. We find that communicators who take ambitious, nonabsolute honesty stances are not immune to hypocrisy judgments but also are not penalized as much as communicators who take flexible stances (i.e., "Lying is sometimes okay"); communicators who take ambitious stances are still seen as more likely to be honest in the future, which drives interpersonal trust.

In our final study, we shed light on how these dynamics influence public hypocrisy. In Study 6, we survey publicly elected and appointed officials in the United States and ask whether these officials anticipate the relative costs of hypocrisy and moral flexibility.

Throughout our studies, we find consistent support that inferences about communicators' general tendency toward (dis)honest behavior underlie negative judgments about communicators who take flexible moral stances. Flexible honesty stances are costly relative to not taking an honesty stance, taking an absolute stance, and taking an ambitious stance, despite seeming less hypocritical than these alternative stances. Absolute honesty stances are perceived as

hypocritical when coupled with dishonest behavior, as are ambitious stances and (in some cases) no honesty stances, and this hypocrisy is viewed negatively. However, in our studies, people believe that taking a flexible stance is more diagnostic of future (dis)honest behavior than engaging in hypocrisy, which undermines moral judgments, voting intentions, and trust.

We also rule out a number of alternative mechanisms for our effects. For example, we find evidence that our effects are not solely due to the perceived societal benefits of absolute proclamations. Though absolute proclamations are seen as more beneficial for society than flexible stances—presumably because communicating moral rules can promote moral behaviors in others (e.g., gossip and communication as cultural learning; Baumeister et al., 2004)—this inference does not consistently mediate the effect of moral proclamations on moral judgments (Study 5 and Supplemental Study 7). We also find that communicators who make absolute proclamations are perceived as feeling guiltier when they lie, consistent with the inference that absolute proclamations are seen as communicating one's underlying commitment to honesty and general tendency to behave honestly, but guilt itself does not consistently mediate our effects (see Study 3 Supplemental Results and Supplemental Studies 1–2). We also consider whether participants believe that communicators do not view or construe their deception within our studies as lies per se, which could lead communicators endorsing absolute honesty to seem more moral; however, we do not find evidence for this alternative explanation (Supplemental Study 3). In Studies 4 and 5, we rule out the possibility that people simply prefer others whose moral beliefs align with their own; even participants who themselves believe that lying is sometimes okay make negative judgments of communicators who take this stance (Studies 4 and 5).

Across our studies, we find consistent evidence that public honesty stances are viewed as signals of one's commitment to honesty and likelihood of engaging in honest behaviors in the future, regardless of whether communicators behave consistently with their stance. This inference helps explain the relatively positive moral judgments of hypocritical absolutist communicators, thereby helping to explain the persistence of hypocritical absolutism in everyday communications. All study materials, raw data, and analysis scripts for all studies are available in our R project on the

Open Science Framework (https://osf.io/wbpm7/?view_only=9e4aff282d064ca2910ca0580868fc72).

Pilot Studies 1–3

To motivate our experiments, we examined the actual stances that people take on honesty, in both public and private, in three pilot studies. In Pilot Study 1 ($N = 130$; 46% female, $M_{\text{age}} = 31.05$, $SD_{\text{age}} = 13.70$), which was conducted in local parks in Chicago, Illinois, we assessed attitudes toward honesty by asking participants to choose the one statement that they most agreed with out of five options: *Lying is never/rarely/sometimes/often/always okay*. In Pilot Study 2 ($N = 154$; 78% female, $M_{\text{age}} = 28.83$, $SD_{\text{age}} = 11.54$), which took place in a university laboratory, and in Pilot Study 3, which was conducted with a sample of government officials, we again asked participants to indicate their own attitudes on honesty. However, we also asked participants to indicate the one statement that most reflected their *public* stance on honesty (i.e., the stance they would take if their goal was to gain public trust; *Lying is never/rarely/sometimes/often/always okay*) and to choose the one statement that most reflected their behavior in everyday life (*I never/rarely/sometimes/often/always lie*; see OSF for full materials).

Pilot Study 3 was included as a module in a survey alongside Study 6, which we administered in partnership with the nonprofit CivicPulse (see Study 6 for more details). Participants in Pilot Studies 1 and 2 had to select an answer from the set of choices presented to them, but participants in Pilot Study 3 had the option of not answering the question at all (which we report as “nonapplicable”). The proportion of participants making each choice in all pilot studies is listed in Table 1.

In these studies, we consider the possibility that people might penalize communicators who take flexible stances on honesty because most people hold absolute stances. Our results suggest this explanation is unlikely. A small minority of community members actually believe that lying is never okay, as indicated by the proportion of participants in Pilots 1 and 2 choosing the option, “Lying is never okay” as most reflective of their beliefs. An even smaller minority of participants say that they never lie. Even government officials in Pilot 3, who were more likely to report that “Lying is never okay” than participants in our other samples, admitted that they typically failed to live up to these standards. These results suggest that any social costs

Table 1
Participants' Attitudes Toward Honesty in Pilot Studies 1–3

Sample	Percentage of participants endorsing each item					
	Always	Often	Sometimes	Rarely	Never	NA
How often do people believe it is okay to lie?						
Participants' own attitudes toward honesty						
Pilot 1: park participants	2%	5%	45%	42%	8%	
Pilot 2: laboratory participants	0%	2%	40%	48%	10%	
Pilot 3: government officials	0%	0%	3%	25%	52%	19%
Participants' public stance on honesty						
Pilot 2: laboratory participants	0%	0%	19%	52%	29%	
Pilot 3: government officials	0%	0%	1%	15%	58%	25%
How often do people say they lie?						
Participants' honest behavior						
Pilot 2: laboratory participants	0%	1%	36%	56%	6%	
Pilot 3: government officials	0%	0%	4%	37%	36%	23%

Note. Percentages reflect the percent of participants within each study choosing each response option (Pilot 1, parks: $N = 130$; Pilot 2, laboratory: $N = 154$; Pilot 3, CivicPulse: $N = 166$). Participants in Pilot 3 had the option to skip these questions, resulting in nonapplicable (NA) responses, but participants answered all questions in Pilot 1–2.

for taking the flexible stance, relative to the absolute stance, are not fully driven by preferences for similarity, as many people do not actually endorse nor think they live up to absolute stances. We return to this point in Studies 4 and 5.

When we examine participants' *public* honesty stances, people appear much less flexible. In Pilot Study 2, 29% of participants indicated that they would take an absolute stance against deception in public, which significantly differs from the proportion of participants (10%) who took this stance in private, $\chi^2(1, N = 154) = 16.44, p < .001$. The difference in the proportion of government officials who endorsed the absolute honesty stance in public and private was directionally similar, but not statistically significant ($p = .320$; see Table 1). These results suggest that both community members and government officials recognize the value of endorsing strong honesty standards in public, even if their public stance does not match their actual behaviors or beliefs. In other words, people seem to (correctly) intuit that being honest about dishonesty is socially penalized. These findings highlight that flexible honesty is not a descriptively normative stance, helping to explain why this stance is costly despite being an accurate depiction of many people's private beliefs and behaviors. We test this idea more directly in Study 6.

Study 1: Moral Stances in the Political Context

In Study 1, we provide an initial investigation of our theoretical account in the political domain. We chose this context because politicians are often required to take public stances on moral issues and thus are likely to face trade-offs between acknowledging moral flexibility and potentially being seen as hypocritical. Specifically, we compare voting intentions for, and moral evaluations of, political candidates who were described as publicly taking absolute or flexible honesty stances and telling equivalent lies.

Method

Participants

We aimed to recruit 1,200 participants from Amazon Mechanical Turk (MTurk). A total of 1,207 participants completed the survey in its entirety (partial data were not analyzed). Comprehension check questions were included at the end of the survey to assess understanding of the political scenario. Per our preregistration, 184 participants were excluded before conducting any analyses because they had answered comprehension check questions incorrectly. We also excluded two participants because they had repeated MTurk IDs, resulting in a final analysis sample of 1,021 participants (51.8% female, $M_{\text{age}} = 38.14, SD_{\text{age}} = 11.60$).

Procedure and Materials

Participants were randomly assigned to one of 12 conditions in a 2 (proclamation: absolute vs. flexible) \times 2 (frame: lie vs. truth) \times 3 (perspective: third-person vs. second-person vs. first-person) between-subjects design. Participants were introduced to a hypothetical political candidate named Matthew Johnson, who was described as belonging to the participant's political party and as running for local office. Though our scenario was based on real-life events,² we used a hypothetical political candidate to reduce potential biases that might be related to preexisting political affiliations or attitudes toward real-life politicians.

Participants read a transcript of a televised interview between the candidate and a television reporter. All the candidate's responses during the interview were held constant across conditions, except the candidate's response to the final question. The final question in the interview addressed ethics in politics and was used to manipulate the candidate's stance on honesty. Specifically, the television reporter asked candidate Matthew Johnson to speak about honesty in politics. Mr. Johnson either responded with an absolute proclamation, "Thank you for bringing this up. I take an absolute stance against lying. As a politician, it is never okay to lie," or a flexible proclamation, "Thank you for bringing this up. I see lying as a complicated issue. As a politician, it is sometimes okay to lie." To examine the robustness of our effects, and ensure our results were not due to any particular wording, we also manipulated two additional features of these proclamations. Specifically, we manipulated whether the (absolute or flexible) proclamation was framed in terms of lying or truth-telling (e.g., "As a politician, it is never okay to lie" vs. "As a politician, it is always imperative to tell the truth") and we manipulated whether the (absolute or flexible) proclamation was stated in the third-person (e.g., "As a politician, it is never okay to lie"), the first-person (e.g., "As a politician, I would never lie"), or the second-person (e.g., "As a politician, you should never lie").

In the next phase of the experiment, participants learned that the same candidate engaged in deception for self-interested reasons by making a false claim about the source of his campaign financing.

Dependent Variables. After learning about Johnson's honesty proclamation and subsequent deception, participants indicated how likely they would be to vote for him and evaluated him in terms of morality, future honesty, and hypocrisy.

Voting Intentions. We measured participants' willingness to vote for Johnson using the following item: "What is the likelihood that you would vote for Matthew Johnson?" (1 = *not at all likely* to 7 = *extremely likely*).

Morality. We measured perceptions of Johnson's morality using a three-item scale ($\alpha = .94$), which asked participants to rate him on three qualities: "moral," "ethical," and "trustworthy" (1 = *not at all* to 7 = *extremely*).

Future Honesty. In this study, we also measured our hypothesized mechanism—perceptions of the candidate's tendencies toward honest behavior and general commitment to honesty—using a four-item scale ($\alpha = .75$): "How often does Matthew Johnson lie?" (1 = *not often at all* to 7 = *extremely often*; reverse-coded); "How likely is Matthew Johnson to lie in the future?" (1 = *not at all likely* to 7 = *extremely likely*; reverse-coded); "In Matthew Johnson's mind, how extreme would a situation need to be in order to deem lying 'acceptable'?" (1 = *not at all extreme* to 7 = *the most extreme*); and a rating of agreement with the statement "Matthew Johnson is committed to honesty." (1 = *not at all* to 7 = *extremely*).

Hypocrisy. We measured perceived hypocrisy using a five-item scale consisting of the following statements ($\alpha = .88$): "Matthew Johnson is a hypocrite."; "Matthew Johnson behaved inconsistently with his values."; "Matthew Johnson's interview response conflicted with his behavior during the campaign."; "Matthew Johnson is

² Our hypothetical scenario was inspired by the following article discussing inconsistencies between Beto O'Rourke's words and deeds: <https://www.politifact.com/factchecks/2017/jul/27/beto-orourke/beto-orourke-claims-near-uniqueness-not-taking-cor/>.

inauthentic.”; “Matthew Johnson is disingenuous.” (1 = *strongly disagree* to 7 = *strongly agree*).

Results

We ran a series of 2 (proclamation) \times 2 (frame) \times 3 (perspective) analyses of variance (ANOVAs) on each of our dependent variables. We focus on the effect of primary theoretical interest (the main effect of proclamation) in our discussion and in Table 2 because the manipulations of frame and perspective did not moderate this effect for any of our measures. Full descriptive statistics across all conditions are reported in Supplemental Material Appendix 1.1.

Voting Intentions

Participants indicated that they were less likely to vote for the candidate who behaved consistently with the flexible honesty proclamation than the candidate who violated the absolute honesty proclamation, $F(1, 1009) = 60.35, p < .001, \eta_p^2 = .06$. This effect was robust across frames and perspectives; no other effects were significant at the $p < .05$ level ($ps > .097$).

Morality

Participants also viewed the candidate who behaved consistently with the flexible honesty proclamation as less moral than the candidate who violated the absolute honesty proclamation, $F(1, 1009) = 58.04, p < .001, \eta_p^2 = .05$. No other effects were significant at the $p < .05$ level ($ps > .176$).

Future Honesty

Participants viewed the candidate who behaved consistently with the flexible honesty proclamation as being less honest in general (i.e., less committed to behaving honestly and less likely to be honest in the future) than the candidate who violated the absolute honesty proclamation, $F(1, 1009) = 113.41, p < .001, \eta_p^2 = .10$, even though this candidate was arguably more honest about their actual behaviors. There was also a main effect of frame, $F(1, 1009) = 4.63, p = .032, \eta_p^2 < .01$. Statements about lies conveyed a greater commitment to honesty and likelihood of future honesty than statements about truth-telling. No other effects were significant at the $p < .05$ level ($ps > .183$).

Hypocrisy

Consistent with existing research on word–deed misalignment (e.g., Jordan et al., 2017), participants viewed the candidate who behaved consistently with the flexible honesty proclamation as less

hypocritical than the candidate who behaved inconsistently with the absolute honesty proclamation, $F(1, 1009) = 72.89, p < .001, \eta_p^2 = .07$, despite being less likely to vote for this candidate, and also judging this candidate to be less moral. No other effects were significant at the $p < .05$ level ($ps > .101$).

Mediation Analyses

We conducted a multiple mediation analysis using a bootstrapping procedure with the lavaan package in R (Rosseel, 2012) to examine how perceptions of hypocrisy and future honesty contributed to the relationship between proclamation (0 = absolute, 1 = flexible) and moral evaluations. Both judgments of hypocrisy and future honesty mediated the effect of proclamation on moral evaluations (indirect effect through hypocrisy = 0.114, 95% CI [0.08, 0.15]; indirect effect through future honesty = -0.487 , 95% CI [$-0.58, -0.39$]). Taking a flexible stance on honesty lowered perceived hypocrisy, relative to taking an absolute stance, which improved moral evaluations. However, taking a flexible stance also decreased perceptions of future honesty, which decreased moral evaluations. Notably, the indirect effect through future honesty was over four times larger than the indirect effect through hypocrisy. Thus, the costs of acknowledging moral nuance outweighed the costs of behaving hypocritically.

Discussion

In Study 1, when a political candidate who took a flexible honesty stance engaged in deception, he was seen as less hypocritical than a candidate who took an absolute honesty stance that he then violated. However, he was also seen as less moral, garnered less political support (in the form of voting intentions), and was seen as less likely to engage in future honesty. Consistent with our theoretical account, judgments of lower future honesty associated with the flexible stance undermined moral judgments more than judgments of hypocrisy associated with the absolute stance. These competing effects highlight the trade-off communicators face when taking honesty stances. These results were robust across multiple framings and perspectives used to convey the candidate’s flexible or absolute stance.

Studies 2 and 3: Disentangling the Costs of Flexibility and Hypocrisy

In Studies 2 and 3, we further disentangle the social costs associated with moral flexibility and hypocrisy by examining judgments of honesty stances over time. Specifically, we asked participants to evaluate communicators at two points in time: once after taking a stance on honesty (T1) and once after the stance was coupled with deceptive behavior (T2).

Study 2: Moral Stances in the Political Context Over Time

In Study 2, we used the same political scenario as in Study 1, but we assessed voting intentions and related judgments both before and after the hypocritical behavior was revealed. We also introduced a control condition in which the candidate did not take any stance on honesty. Comparing evaluations of candidates who take absolute and flexible honesty stances to a control candidate sheds light on

Table 2
Main Effect of Proclamation in Study 1

Dependent measure	Proclamation	
	Absolute	Flexible
Voting intentions	3.01 (1.42)	2.36 (1.23)
Morality	2.87 (1.23)	2.32 (1.03)
Future honesty	3.12 (1.06)	2.45 (0.87)
Hypocrisy	5.38 (1.14)	4.75 (1.27)

Note. Marginal means, with standard deviations in parentheses.

whether our effects are driven by costs of flexible stances (as we argue) or rather by benefits of absolute stances.

Method

Participants

We aimed to recruit 600 participants from MTurk. A total of 603 participants completed the survey in its entirety (partial responses were excluded from analysis). Comprehension check questions were included at the end of the survey to assess understanding of the political scenario. Consistent with our preregistration, 39 participants were excluded for answering comprehension check questions incorrectly, resulting in a final sample of 564 participants (44.9% female, $M_{\text{age}} = 41.80$, $SD_{\text{age}} = 12.31$).

Procedure and Materials

The stimuli used in Study 2 were identical to the stimuli used in the third-person perspective, lie frame condition in Study 1 (e.g., “As a politician, it is never okay to lie”), except we added an additional control condition. In the control condition, Mr. Johnson was never asked to speak about ethics in politics and made no proclamation about honesty. Proclamation was manipulated between subjects (proclamation: absolute vs. flexible vs. control), and time was manipulated within subjects (T1 vs. T2).

Dependent Variables. Participants indicated how likely they would be to vote for the candidate and evaluated the candidate’s morality and future honesty at both points in time using the same measures as in Study 1 ($\alpha \geq .81$). Participants also evaluated the candidates in terms of hypocrisy, but only at T2, after the lie had been revealed, using the same measure as in Study 1 ($\alpha = .88$).

Results

We report results from a series of mixed ANOVAs, using proclamation as a between-subjects factor and time as a within-subjects factor, on our dependent variables of voting intentions, morality, and future honesty. We also report the results from a one-way ANOVA examining the influence of proclamation on hypocrisy, since hypocrisy was only evaluated at T2. We followed up significant effects with independent-samples t tests to compare differences in evaluations between conditions and with paired-samples t tests to compare differences in evaluations within a proclamation condition over time. Table 3 presents all descriptive statistics.

Voting Intentions

There was a significant main effect of proclamation, $F(2, 561) = 134.19$, $p < .001$, $\eta_p^2 = .32$, on voting intentions, such that participants were less likely to vote for the candidate who endorsed flexible honesty than the candidate who endorsed absolute honesty, $t(708) = 16.34$, $p < .001$, $d = 1.23$, or the control candidate, $t(786) = 17.37$, $p < .001$, $d = 1.24$. Participants indicated similar levels of willingness to vote for the candidate endorsing absolute honesty and the control candidate, $t(756) = -0.45$, $p = .649$, $d = -0.03$, despite the fact that the candidate endorsing absolute honesty behaved hypocritically and the control candidate did not. There was also a main effect of time, $F(1, 561) = .42$, $p < .001$, $\eta_p^2 = .52$, such that

Table 3
Effect of Proclamation Over Time in Study 2

Dependent variable	Proclamation		
	Absolute	Flexible	Control
A. Voting intentions			
Time 1	4.80 (1.30)	2.41 (1.37)	4.72 (1.42)
Time 2	3.05 (1.47)	1.86 (1.08)	3.24 (1.55)
B. Morality			
Time 1	4.92 (1.27)	2.54 (1.21)	4.65 (1.28)
Time 2	2.79 (1.33)	1.96 (1.01)	2.94 (1.41)
C. Future honesty			
Time 1	5.18 (1.23)	2.70 (1.00)	4.52 (1.18)
Time 2	2.88 (1.13)	2.18 (0.92)	2.88 (1.11)
D. Hypocrisy			
Time 2	5.66 (1.08)	4.97 (1.32)	5.24 (1.10)

Note. Marginal means, with standard deviations in parentheses.

participants were less willing to vote for candidates at T2 ($M = 2.73$, $SD = 1.52$), after the deception was revealed, than at T1 ($M = 3.98$, $SD = 1.76$), before the deception was revealed.

Importantly, these main effects were qualified by a significant interaction between proclamation and time, $F(2, 561) = 48.99$, $p < .001$, $\eta_p^2 = .15$. Though participants indicated lower voting intentions at T2 for all three candidates relative to T1, absolute: $t(169) = 14.61$, $p < .001$, $d_{\text{RM}} = 1.12$; control: $t(208) = 17.73$, $p < .001$, $d_{\text{RM}} = 1.23$; flexible: $t(184) = 9.28$, $p < .001$, $d_{\text{RM}} = 0.68$,³ demonstrating a social penalty for engaging in deception, the decrease in voting intentions over time was largest, in terms of mean differences (see Table 3), for the candidate who endorsed absolute honesty and lied, highlighting the costs of hypocrisy. Nevertheless, this absolutist candidate still received greater voting intentions at T2 than the candidate endorsing flexible honesty (absolute vs. flexible candidates at T2: $t(353) = 8.69$, $p < .001$, $d = 0.92$). The fact that participants were less willing to vote for the candidate endorsing flexible honesty than the candidate endorsing absolute honesty at T2 indicates that the costs of hypocrisy were not sufficient to outweigh the initial cost of taking a flexible stance on honesty (absolute vs. flexible candidates at T1: $t(353) = 16.82$, $p < .001$, $d = 1.79$). Furthermore, we find that the flexible stance was initially quite costly even relative to a control condition (flexible vs. control candidates at T1: $t(392) = 16.36$, $p < .001$, $d = 1.65$), but the absolute stance did not provide an initial advantage relative to the control candidate (absolute vs. control candidates at T1: $t(377) = 0.58$, $p = .562$, $d = 0.06$). These results indicate that flexibility is particularly costly.

Morality

We found a similar pattern for moral evaluations. Overall, there was a main effect of proclamation on morality ratings, $F(2, 561) = 129.40$, $p < .001$, $\eta_p^2 = .32$, such that the candidate endorsing flexible honesty was seen as less moral than the candidate endorsing absolute honesty, $t(708) = 14.97$, $p < .001$, $d = 1.12$, and the control

³ d_{RM} denotes the repeated-measures Cohen’s d , calculated as the mean difference score between T1 and T2 ratings, divided by the standard deviation of difference scores. For more on how to interpret this statistic, and how it relates to the more familiar independent groups Cohen’s d , see Morris and DeShon (2002).

candidate, $t(786) = 15.42, p < .001, d = 1.10$. The candidate endorsing absolute honesty and the control candidate were viewed as similarly moral, $t(756) = 0.52, p = .606, d = 0.04$. There was also a main effect of time, $F(1, 561) = 754.83, p < .001, \eta_p^2 = .57$. Participants were viewed as less moral at T2 ($M = 2.57, SD = 1.34$) than T1 ($M = 4.04, SD = 1.64$), after deception was revealed.

These main effects were qualified by a significant interaction between proclamation and time, $F(2, 561) = 71.96, p < .001, \eta_p^2 = .20$. The biggest decrease in morality ratings over time, in terms of mean differences (see Table 3), was for the candidate who endorsed absolute honesty and lied, absolute: $t(169) = 17.51, p < .001, d_{RM} = 1.34$; control: $t(208) = 18.59, p < .001, d_{RM} = 1.29$; flexible: $t(184) = 10.18, p < .001, d_{RM} = 0.75$, reflecting a penalty for hypocrisy. However, this candidate was still seen as more moral than the more consistent candidate who endorsed flexible honesty and lied at T2 (absolute vs. flexible candidates at T2: $t(353) = 6.69, p < .001, d = 0.71$). The control candidate and the candidate who endorsed absolute honesty and lied were seen as similarly moral, overall, $t(756) = 0.52, p = .606, d = 0.04$, including at T2, $t(377) = -1.01, p = .314, d = -0.10$. Both the control candidate and candidate endorsing absolute honesty were seen as more moral than the candidate who endorsed flexible honesty and lied overall, control versus flexible candidate: $t(786) = 15.42, p < .001, d = 1.10$; absolute versus flexible candidate overall: $t(708) = 14.97, p < .001, d = 1.12$, highlighting the costs of flexibility.

Future Honesty

There was a main effect of proclamation, $F(2, 561) = 147.94, p < .001, \eta_p^2 = .35$, on future honesty. The candidate who endorsed flexible honesty was viewed as less likely to be honest in the future than the candidate who endorsed absolute honesty, $t(708) = 15.66, p < .001, d = 1.18$, and the control candidate, $t(786) = 14.39, p < .001, d = 1.03$. The control candidate was also seen as less likely to be honest in the future than the candidate who endorsed absolute honesty, $t(756) = 2.92, p = .004, d = 0.21$. Although voting intentions and morality were similar for the candidate endorsing absolute honesty and the control candidate, the absolute stance conferred some benefit in future honesty impressions. There was also a main effect of time, $F(1, 561) = 909.41, p < .001, \eta_p^2 = .62$, such that future honesty of all three candidates was lower at T2, after the deception was revealed.

These main effects were qualified by a significant interaction between proclamation and time, $F(2, 561) = 107.83, p < .001, \eta_p^2 = .28$. Although all candidates were seen as less honest at T2 (relative to T1), this difference was the greatest for the candidate endorsing absolute honesty, absolute: $t(169) = 21.03, p < .001, d_{RM} = 1.61$; control: $t(208) = 19.35, p < .001, d_{RM} = 1.34$; flexible: $t(184) = 9.23, p < .001, d_{RM} = 0.68$. Despite the fact that participants adjusted their perceptions of the candidate endorsing absolute honesty more than any other candidates, they still perceived this candidate as more likely to engage in future honesty than the candidate who endorsed flexible honesty and behaved consistently at T2 (absolute vs. flexible candidates at T2: $t(353) = 6.34, p < .001, d = 0.67$), due to the fact that the initial flexible stance was so costly for future honesty impressions (absolute vs. flexible candidates at T1: $t(353) = 20.93, p < .001, d = 2.22$). Examining T1 evaluations also provides evidence that the flexible stance is costly for future honesty impressions relative to the control candidate (control vs. flexible candidates at T1: $t(392) = 16.49, p < .001, d = 1.66$).

Hypocrisy

Hypocrisy, unlike our other measures, was only collected at T2, after the deception was revealed. There was a main effect of proclamation, $F(2, 561) = 15.67, p < .001, \eta_p^2 = .05$, such that participants viewed the candidate who endorsed absolute honesty as more hypocritical than the control candidate, $t(377) = 3.75, p < .001, d = 0.39$, and the candidate who endorsed flexible honesty, $t(353) = 5.39, p < .001, d = 0.57$. The control candidate was also seen as more hypocritical than the candidate who endorsed flexible honesty, $t(392) = 2.20, p = .028, d = 0.22$. It is possible that people expect others to publicly endorse absolute honesty as the “default” stance, and consequently, lying seems hypocritical even when one has not taken any honesty stance at all.

Mediation Analyses

We conducted a multiple mediation analysis using a bootstrapping procedure with the lavaan package in R (Rosseel, 2012) with proclamation entered as the independent variable (1 = flexible, 0 = absolute). We focus on the flexible and absolute stance conditions (excluding the control condition) because they reflect our key conditions of theoretical interest. Judgments of hypocrisy and future honesty at T2 were entered as simultaneous mediators. T2 judgments of morality were entered as the dependent variable. We were interested in T2 evaluations, since they represent participants’ impressions after both the stance and deception were revealed.

Results indicated that the flexible honesty proclamation led to decreased perceptions of hypocrisy and of future honesty, compared to the absolute proclamation. Both of these inferences predicted moral evaluations, in opposite directions. Decreased perceptions of hypocrisy associated with the flexible stance were correlated with more positive perceptions of morality, but the decreased perceptions of future honesty were correlated with more negative perceptions of morality (indirect effect through hypocrisy = 0.104, 95% CI [0.03, 0.19]; indirect effect through future honesty = -0.522, 95% CI [-0.70, -0.35]). As in Study 1, however, the indirect effect through future honesty was larger than the indirect effect through hypocrisy, which helps to explain why flexible, but realistic, stances led to more negative moral judgments than absolute, yet hypocritical, stances.

Discussion

In Study 2, a political candidate who took a flexible honesty stance was viewed as less moral and garnered less political support than a candidate who took a hypocritical absolute stance and a candidate who took no honesty stance at all. Examining T1 evaluations demonstrates that the candidate endorsing flexible honesty was viewed as less likely to engage in honest behaviors in the future relative to both other candidates, which undermined initial moral judgments and political support. Examining T2 evaluations highlights that the candidate endorsing flexible honesty was also seen as the least hypocritical, but nevertheless, the initial costs of flexibility still detracted from moral evaluations at T2. These results also show that candidates endorsing absolute honesty and those taking no honesty stance are viewed relatively similarly, suggesting that participants are not making their moral evaluations merely on the basis of a running tally of good and bad acts: taking *no* stance and then lying was seen as morally similar to taking a *praiseworthy*

stance and then violating it by lying at T2. Acknowledging moral flexibility seems to be uniquely costly (we replicate this result in Supplemental Study 7, see Supplemental Material Appendix 2.7).

Study 3: Moral Stances in the Laboratory Over Time

Studies 1 and 2 establish that flexibility has a negative effect on moral evaluations that competes with the effect of hypocrisy in a political context. To ensure that our results from Studies 1 and 2 are not unique to political domains, nor to the specific scenarios we employed, we examine participants' perceptions of anonymous actors in economic games in Studies 3–5.

In Study 3, as in Study 2, we examine the consequences of taking absolute and flexible stances at two points in time. In Study 3, however, we also vary the order in which information (honesty stance and lie behavior) is revealed to participants. Some participants learned about the communicator's (absolute or flexible) honesty stance first, evaluating the communicators after just this information, and then learning about deception (as in Study 2), but some participants learned about the deception first. This design allows us to examine whether the flexible honesty stance is still costly when it follows a lie—and therefore may be perceived as a justification for the prior behavior—rather than when it precedes a lie.

In Study 3, we also pair absolute and flexible stances with prosocial lies. In Studies 1 and 2, the political candidate lied by misrepresenting the source of his campaign funding. A candidate would tell this lie, presumably, to boost his image among voters (i.e., for selfish reasons). However, most people take flexible honesty stances because they think it is ethical to lie for *prosocial*, not selfish reasons (Levine, 2021). Therefore, a communicator who says it is “sometimes okay to lie” and then tells a prosocial, rather than selfish, lie might be perceived as making thoughtful moral exceptions. Furthermore, a flexible honesty stance that follows a prosocial lie may be viewed as justifying these ethical exceptions *specifically*, and therefore, evaluated more positively. Study 3 investigates these possibilities.

Method

Participants

We aimed to recruit 600 participants on MTurk. Comprehension check questions were included in the middle of the survey, before any dependent measures were collected, to assess understanding of the communicator's behavior in the previous study and to verify understanding of supplemental game instructions (see Supplemental Material Appendix 1.3, for supplemental game results). Participants who failed a minimum of two comprehension checks were automatically prevented from continuing with the survey. Six hundred two participants successfully passed comprehension checks and completed the full survey. From these 602 participants, we excluded two participants with repeated MTurk IDs, resulting in a final analysis sample of 600 participants (48.67% female, $M_{\text{age}} = 37.86$, $SD_{\text{age}} = 12.08$).

Procedure and Materials

Participants were matched with a partner who was a past study participant on MTurk (we henceforth refer to this person as the “communicator”). Participants learned about the communicator's decisions in the previous experiment, including their responses to a

survey about beliefs and preferences and their decision to lie in a deception game (Gneezy, 2005; Levine & Schweitzer, 2014, 2015).

Communicator Information. To manipulate the communicator's stance on honesty, we provided participants with the communicator's response to a survey question asking them to choose the one statement that they most agreed with from the options: *It is never/sometimes/always okay for people to lie*. The communicator chose one of two focal responses: “It is never okay for people to lie” (the absolute proclamation) or “It is sometimes okay for people to lie” (the flexible proclamation). All other survey responses were held constant across conditions.

Participants also learned about the same communicator's behavior in a deception game, in which the communicator had the opportunity to lie or tell the truth about the outcome of a coin flip in exchange for different monetary outcomes. In this game, telling a lie was prosocial; it harmed the communicator and benefited another player in the game, who was not our focal participant. On the other hand, telling the truth was selfish; it helped the communicator but harmed the other player in the game. All communicators were described as telling a prosocial lie. The exact details of the game are available on OSF and summarized in brief in Figure 2.

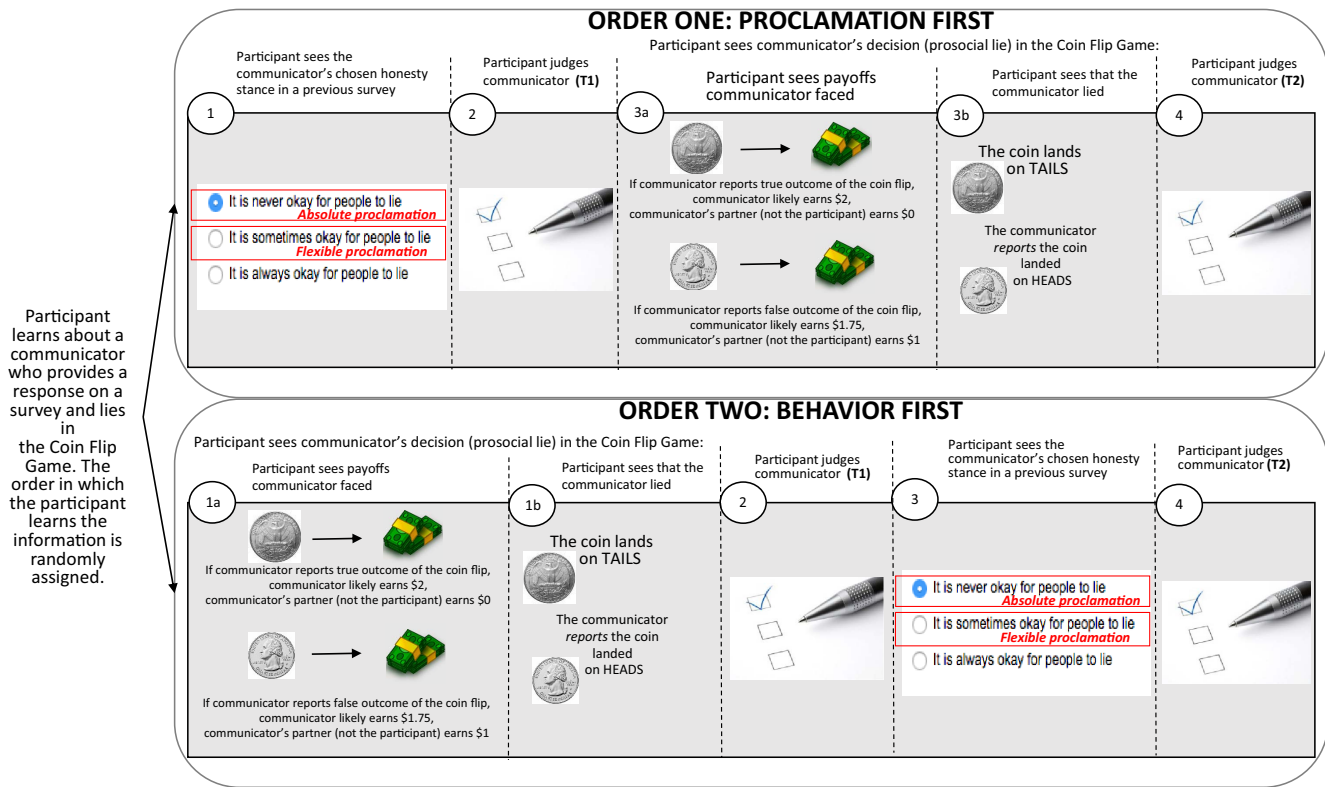
We randomly assigned participants to one cell of a 2 (proclamation: absolute vs. flexible) \times 2 (order: proclamation first vs. behavior first) \times 2 (time: T1 vs. T2) mixed design. Proclamation and order were manipulated between subjects; time was manipulated within subjects. Participants learned about either the survey responses (and proclamation) or the lie behavior in phase one and then completed one set of evaluations about the communicator. In phase two, participants learned about the remaining information (lie behavior or proclamation) and completed a second set of evaluations. This design is depicted in Figure 2.

Dependent Variables. At T1 and T2, participants evaluated the communicator's morality, future honesty, and hypocrisy (all $\alpha \geq .82$). The morality scale was identical to the one used in Studies 1 and 2. For future honesty and hypocrisy, we used measures that were similar to those used in Studies 1 and 2, but the wording was modified to account for differences between the political and deception game context. References to Matthew Johnson were replaced with “person” more generally; for example, participants answered the question, “How often does this person lie?” rather than “How often does Matthew Johnson lie?” (an item in the future honesty scale). We used a four- rather than five-item scale for hypocrisy in this study: “This person's survey responses conflicted with his/her behavior in The Coin Flip Game” was removed from our analyses because participants in the “proclamation first” condition had not yet learned about the communicator's deception at T1.

Results

A series of mixed ANOVAs were conducted to investigate the effects of proclamation (absolute vs. flexible), order (proclamation first vs. behavior first), and time (T1 vs. T2) on each attitudinal variable. In these analyses, to follow-up on significant interactions, we ran a series of independent-samples *t* tests for measures between groups and paired *t* tests for ratings across time. We only report detailed results for the morality variable below since this is our focal variable in this study. We include brief results for the other attitudinal measures in this section. The full results from the ANOVAs on future honesty and hypocrisy and descriptive statistics for all these measures are reported in Supplemental Material Appendix 1.3.

Figure 2
Diagram of Study 3 Design



Note. The top panel depicts the sequence of events for participants in the proclamation first condition. The bottom panel depicts the sequence of events for participants in the behavior first condition. Participants in both order conditions complete two evaluations at different time points (whether T1 is after just the proclamation or after just the deceptive behavior varies based on assignment to proclamation first or behavior first). See the online article for the color version of this figure.

Morality

There was a significant main effect of proclamation, $F(1, 596) = 72.76, p < .001, \eta_p^2 = .11$, such that communicators endorsing the absolute honesty stance were viewed as more moral than communicators endorsing the flexible honesty stance. This reflects an overall social penalty for moral flexibility, as in Study 2. Notably, we observed a significant moral flexibility penalty regardless of when participants learned about the communicator's stance on honesty (see Figure 3). There was also a main effect of order, $F(1, 596) = 13.21, p < .001, \eta_p^2 = .02$, such that communicators were viewed as more moral when the proclamation was revealed prior to the (dishonest) behavior. There was also a main effect of time, $F(1, 596) = 19.07, p < .001, \eta_p^2 = .03$, such that communicators were viewed as more moral at T1 than T2.

Importantly, these findings were qualified by a significant Proclamation \times Time interaction, $F(1, 596) = 19.64, p < .001, \eta_p^2 = .03$. Communicators who endorsed the absolute honesty stance were viewed as less moral at T2 than T1, $t(305) = 5.51, p < .001, d_{RM} = 0.31$. In other words, communicators who endorsed absolute stances did incur a hypocrisy penalty when it was revealed that they also lied (see Figure 3, for illustration). However, moral evaluations toward communicators who endorsed the flexible honesty stance did not change over time, $t(293) = -0.05, p = .960, d_{RM} < 0.01$;

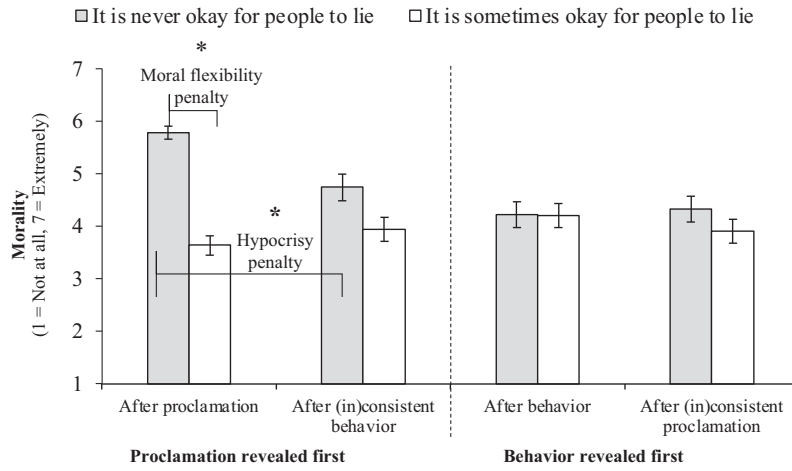
see Figure 3, for illustration; that is, communicators who endorsed flexible stances did not incur a penalty after they lied. Notably, this result is different from what we identified in Study 2, in which communicators who had endorsed the flexible stance were penalized after lying, albeit to a lesser degree than communicators who had endorsed the absolute stance. This difference is likely driven by the nature of the lie: telling a prosocial lie (which is typically seen as ethical) did not lead to a deception penalty when it was consistent with the communicator's stance, whereas the selfish lie in Study 2 did.

There was also a significant Proclamation \times Order interaction, $F(1, 596) = 39.29, p < .001, \eta_p^2 = .06$. Communicators who endorsed the absolute honesty stance were seen as more moral when the proclamation (vs. behavior) was revealed first, $t(610) = -8.36, p < .001, d = -0.68$. This result implies that saying it is "never okay to lie" is perceived more favorably than the action of telling a prosocial lie. The reverse pattern was found for communicators who endorsed the flexible honesty stance. Communicators were seen as more moral when the behavior (vs. proclamation) was revealed first, $t(586) = 2.36, p = .019, d = 0.19$, implying that saying it is "sometimes okay to lie" is perceived less favorably than telling a prosocial lie.

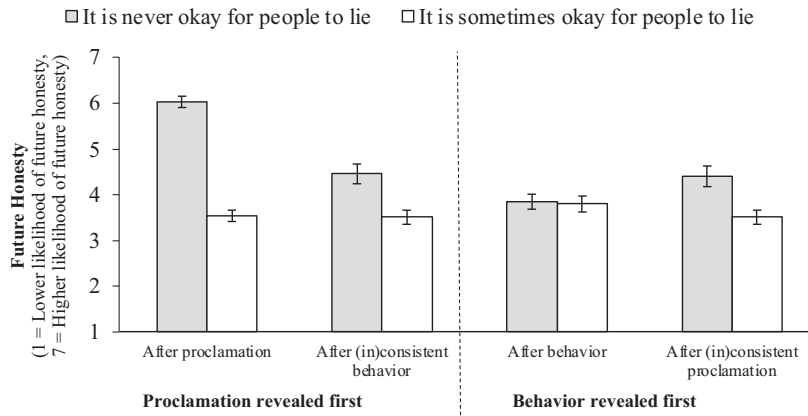
Additionally, there was a significant Order \times Time interaction, $F(1, 596) = 7.13, p = .008, \eta_p^2 = .01$. Communicators were viewed as less moral at T2 than T1 when the proclamation was revealed

Figure 3
Evaluations of Communicators Based on Proclamation, Order, and Time

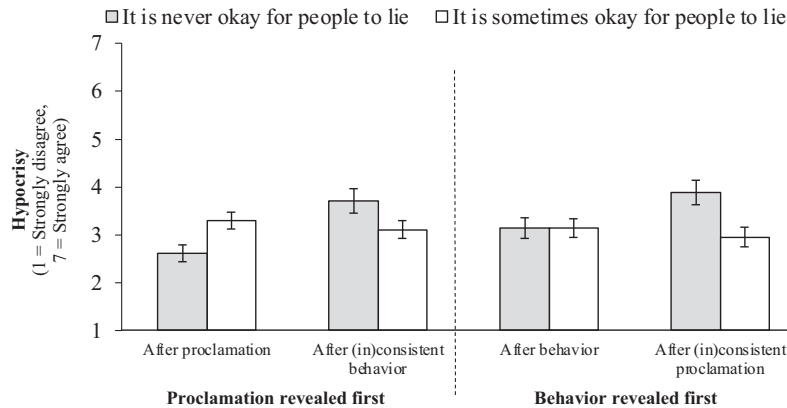
Panel A. Morality



Panel B. Future Honesty



Panel C. Hypocrisy



Note. Error bars reflect 95% confidence intervals. *reflects significant differences between cells.

before, $t(304) = 4.24, p < .001, d_{RM} = 0.24$, but not after, $t(294) = 1.42, p = .156, d_{RM} = 0.08$, the behavior. In other words, revealing deceptive behavior deflated moral evaluations at T2, but revealing moral proclamations after the lie was already known did not.

These effects were further qualified by a significant three-way Proclamation \times Order \times Time interaction, $F(1, 596) = 70.03, p < .001, \eta_p^2 = .11$. There was a hypocrisy penalty for communicators who endorsed the absolute honesty stance, but not communicators who

endorsed the flexible honesty stance, only when the proclamation preceded (vs. followed) behavior.

Additional Attitudinal Measures

In addition to lowering moral evaluations, endorsing the flexible stance also led to lower impressions of future honesty, $F(1, 596) = 220.75, p < .001, \eta_p^2 = .27$, particularly at T1 when the proclamation was revealed first (see Figure 3). Communicators endorsing absolute honesty lost moral credit and were judged as lower in future honesty at T2 when they violated their absolute honesty stance, suggesting a hypocrisy penalty.

Consistent with this notion, communicators who endorsed the absolute honesty stance were also viewed as more hypocritical than communicators who endorsed the flexible honesty stance, especially at T2 when the word–deed misalignment was revealed (comparison of hypocrisy evaluations for communicators endorsing the absolute stance across time: $t(305) = -10.09, p < .001, d_{RM} = 0.58$; see Supplemental Material Appendix 1.3, for further hypocrisy results). Communicators endorsing the flexible honesty stance, on the other hand, were actually seen as *less* hypocritical across time, $t(293) = 2.78, p = .006, d_{RM} = 0.16$ —they did not incur a hypocrisy penalty and did not lose moral credit over time for their prosocial deception. However, these communicators were also seen as less likely to engage in future honesty at T2 than T1 (see Supplemental Material Appendix 1.3, for further future honesty results). Though communicators endorsing absolute and flexible honesty stances were both rated as less likely to engage in future honesty at T2 than T1, the discrepancy in future honesty evaluations across time was larger for the communicators endorsing absolute honesty, $t(305) = 5.55, p < .001, d_{RM} = 0.32$, versus flexible honesty, $t(293) = 2.98, p = .003, d_{RM} = 0.17$. Taken together, communicators who endorsed the flexible honesty stance and lied suffered less severe consequences in interpersonal evaluations across time (in hypocrisy, morality, and future honesty evaluations) than communicators who endorsed the absolute honesty stance and lied. Even so, communicators endorsing the flexible honesty stance were viewed as less moral and less likely to engage in future honesty *overall*. In other words, the costs of hypocrisy were not enough to offset the costs of the flexible stance, as in Studies 1 and 2 (see Figure 3).

Mediation Analyses

We conducted a multiple mediation analysis using a bootstrapped procedure with the lavaan package in R (Rosseel, 2012). In this analysis, we examined whether hypocrisy and future honesty ratings mediate the relationship between proclamation (0 = absolute, 1 = flexible) and T2 moral evaluations, collapsed across order. Results provide further evidence that communicators who endorsed absolute proclamations lost moral credit for behaving hypocritically, but the signal of future honesty associated with the absolute proclamation counteracted this negative effect (indirect effect through hypocrisy = 0.307, 95% CI [0.20, 0.42]; indirect effect through future honesty = -0.563, 95% CI [-0.71, -0.43]).

Discussion

Study 3 largely replicates the results of Study 2, using a very different paradigm. As in Study 2 (and consistent with much prior

work), we find that hypocrisy does have reputational costs; when proclamations were revealed first, communicators endorsing absolute honesty were penalized at T2 relative to T1. However, the initial costs of the flexible stance (T1 evaluations of communicators endorsing absolute vs. flexible stances) were severe enough that flexibility was still costly at T2 relative to absolutism. Unlike Study 2, we find that prosocial lying was *not* penalized when it was paired with a flexible stance (T1 vs. T2 evaluations for communicators endorsing flexibly honesty), presumably because prosocial lies are seen as relatively ethical, whereas selfish lies—the focus of Studies 1 and 2—are not. We replicate these across-time patterns in Supplemental Studies S3 and S5 (see Supplemental Material Appendices 2.3 and 2.5).

Study 3 also sheds light on how the order in which a communicator's stance versus their behavior is revealed influences the costs of hypocrisy and flexibility. It is reasonable to wonder whether honesty stances might be interpreted as a justification (in the case of a flexible stance) or an apology (in the case of an absolute stance) for deception when these stances follow rather than precede deception. However, we find no evidence for this; the order in which a communicator's stance versus their behavior was revealed did not moderate our key effects.

Overall, across Studies 2 and 3, the costs of admitting to moral flexibility outright were greater than the costs of hypocrisy. Communicators who endorsed absolute honesty and lied were still seen as hypocritical, which is costly, but our mediation results provide evidence that the negative effect of moral flexibility on future honesty competes with the negative effect of moral absolutism on hypocrisy in driving moral evaluations, consistent with our theoretical account.

Studies 4 and 5: Absolute and Flexible Stances Paired With Lies and Truths

Studies 4 and 5 build on the paradigm used in Study 3 and extend our investigation in four ways. First, we examine how absolute versus flexible moral proclamations influence a behavioral measure of trust. Although participants might evaluate communicators who make flexible honesty proclamations more negatively than communicators who make absolute honesty proclamations, these judgments may not translate to greater trust in hypocritical communicators.

Second, we further explore the mechanisms underlying our effects. Specifically, we consider whether the perceived social benefit of the proclamation itself explains the relationship between proclamation and morality. We also examine a mechanism of shared beliefs by considering whether participants' own stances on honesty moderate our effects. We find inconsistent support in favor of these alternative mechanisms while continuing to find evidence that absolute honesty proclamations are viewed more positively than flexible honesty proclamations because they signal greater future honesty.

However, high honesty standards are not unique to absolutism. In these studies, we extend our investigation in a third way by considering whether communicators can make ambitious, nonabsolutist claims (“Lying is rarely okay”) that can convey similar commitment to future honesty. We compare evaluations of communicators who take these ambitious, but nonabsolute, honesty stances to those who take absolute (“Lying is never okay”) or flexible (“Lying is sometimes okay”) proclamations about honesty. We expected that the ambitious stance would signal greater future

honesty than the flexible stance, and either lower or similar future honesty than the absolute stance. However, we also expected that the ambitious stance would be seen as less hypocritical than the absolute stance, and either more or similarly hypocritical relative to the flexible stance, once deception was revealed. Because we expected judgments of future honesty to be more predictive of trust and morality than perceived hypocrisy, we expected the ambitious stance to be perceived more similarly to the absolute stance than the flexible stance for measures of trust and morality.

Finally, we orthogonally manipulated communicators' proclamations about honesty and their lying or truth-telling behavior. Doing so allows us to examine judgments of moral consistency (i.e., taking an absolute stance and then telling *the truth*) and test how much moral proclamations influence interpersonal judgment and trust, relative to actual behavior. In Study 4, we examine the behaviors of prosocial lying and selfish truth-telling. In Study 5, we examine the behaviors of prosocial truth-telling and selfish lying.

Study 4: Prosocial Lies and Selfish Truths

Study 4 uses the same setup as Study 3; participants observe the behavior of an anonymous communicator who completed a previous study. However, in Study 4, we also included communicators who took ambitious honesty stances and communicators who told the selfish truth. Unlike Study 3, participants in Study 4 completed all evaluations at one time (after both the stance and behavior were revealed) and then played a trust game with communicators.

Method

Participants

We aimed to recruit 600 participants on MTurk. Comprehension check questions were included in the middle of the survey, before any dependent variables were collected, to assess comprehension of the communicator's behavior and the instructions to a trust game. Participants who failed a minimum of two comprehension checks were automatically prevented from continuing with the survey. Six hundred two participants successfully passed comprehension checks and completed the full survey. From these 602 participants, we excluded four participants with repeated MTurk IDs, resulting in a final analysis sample of 598 participants (43.31% female, $M_{\text{age}} = 35.49$, $SD_{\text{age}} = 10.56$).

Procedure and Materials

We randomly assigned participants to one cell of a 3 (proclamation: absolute vs. ambitious vs. flexible) \times 2 (behavior: prosocial lie vs. selfish truth) between-subjects design.

We used the same paradigm from Study 3 to manipulate the communicator's stance and their behavior. Participants either learned about a communicator who had told a prosocial lie (identical to Study 3) or a selfish truth in the deception game. That is, participants either learned about a communicator who told the truth to receive more money in a game at the expense of another player or who lied to benefit another player at their own expense. (Study 4 uses the same setup depicted in the "Order One: Proclamation First" panel of Figure 2.)

Dependent Variables. After learning about the communicator's honesty stance and previous behavior in the deception game, participants evaluated the communicator in terms of morality, future honesty, and hypocrisy using the same measures as in Study 3 (all $\alpha_s \geq .84$).

The Rely-or-Verify Game. After evaluating the communicator, participants played the Rely-or-Verify (RV) game, a behavioral measure of integrity-based trust, with the communicator. The RV game (Levine & Schweitzer, 2015) is a two-party strategic game in which one party (the participant) must decide whether to trust another party's (the communicator's) claim or verify that claim at a cost. First, the communicator sends a message to the participant. The payoffs of the game are structured such that the communicator has a personal incentive to send a dishonest message. The participant decides whether to rely on this message or verify the accuracy of this message before reporting the message to the experimenter. Participants have an incentive to rely on the message if they believe that the communicator is telling the truth but to verify the content of the message if they believe that the communicator is lying. The decision to rely on the communicator reflects trust in the veracity of the communicator's message. The mixed-strategy equilibrium in the game is to rely 40% of the time.

Attitudes Toward Honesty. At the end of the study, we also measured participants' own attitudes toward honesty. Participants answered the same question that the communicator had received in the previous survey about beliefs and preferences: "Please choose the one statement that you most agree with from the options below: *Lying is always/often/sometimes/rarely/never okay.*"

Results

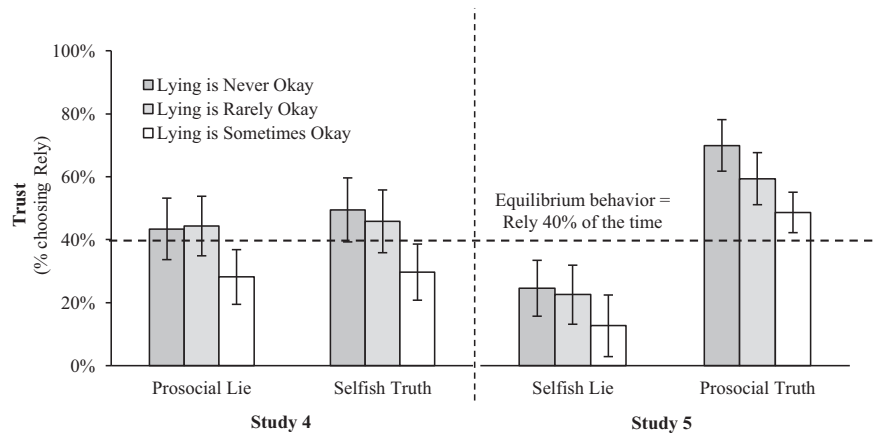
Trust in the RV Game

We conducted a series of logistic regressions to examine the impact of the communicator's proclamation and behavior on participants' decisions to rely on the communicator's words in the RV game. We regressed the decision to rely (1 = rely, 0 = verify) on dummy variables for the flexible (1 = flexible, 0 = ambitious or absolute) and ambitious (1 = ambitious, 0 = flexible or absolute) proclamations, behavior (1 = prosocial lie, 0 = selfish truth), and their interaction terms.

Results revealed a significant main effect of the flexible proclamation ($b = -0.84$, $p = .005$), but no main effect of the ambitious proclamation ($b = -0.15$, $p = .618$) or behavior ($b = -0.24$, $p = .403$). Participants were significantly less likely to rely on communicators who chose the proclamation "Lying is sometimes okay" (29% rely) compared to communicators who chose the proclamation "Lying is never okay" (46% rely), even after telling equivalent lies. Participants did not significantly differ in their willingness to rely on communicators who chose the proclamation "Lying is never okay" and "Lying is rarely okay" (45% rely). The communicator's honesty proclamation was more predictive of trust than their honest or dishonest behavior, indicating that the signal of future honesty associated with both the absolute and ambitious moral proclamation is quite strong; a communicator's honesty stance can be a more important determinant of trust than actual deceptive behavior, at least when deception is prosocial (see Figure 4).

There was no Flexible Proclamation \times Behavior ($b = 0.17$, $p = .693$) or Ambitious Proclamation \times Behavior ($b = 0.18$, $p = .653$) interaction, suggesting that the communicator's honesty stance

Figure 4
Percentage of Participants Who Trust Communicators (as Measured by the Decision to Rely on a Communicator's Words)



Note. Error bars reflect 95% confidence intervals. The dashed line at 40% trust reflects the equilibrium behavior in the game (to rely 40% of the time).

predicted similar levels of trust regardless of whether a flexible, ambitious, or absolute stance was paired with actual truth-telling or deceptive behavior. In other words, communicating high future honesty standards (by one's stance) was more important for trust than moral consistency.

Choosing the flexible stance also undermined trust in communicators relative to the rational equilibrium in the RV game (see Appendix B in Levine & Schweitzer, 2015, for equilibrium solution); participants relied on communicators who chose "Lying is sometimes okay" significantly less often than 40% of the time (29% vs. 40%), $\chi^2(1, N = 204) = 10.43, p = .001$.

Attitudinal Measures

A series of two-way ANOVAs were conducted to investigate the effects of proclamation (absolute vs. ambitious vs. flexible) and behavior (selfish truth vs. prosocial lie) on each attitudinal variable. Significant effects were followed up with independent-samples *t* tests, as were preregistered. Descriptive statistics for Study 4 are presented in Table 4.

Morality. There was a main effect of proclamation, $F(2, 592) = 20.83, p < .001, \eta_p^2 = .07$, on moral evaluations. Communicators who endorsed the absolute honesty proclamation were viewed as more moral than communicators who endorsed the ambitious honesty proclamation, $t(392) = 3.05, p = .002, d = 0.31$, and communicators who endorsed the ambitious honesty proclamation were viewed as more moral than communicators who endorsed the flexible honesty proclamation, $t(404) = 3.30, p = .001, d = 0.33$. Neither the main effect of behavior nor the interaction was significant at the $p < .05$ level ($ps > .272$).

Future Honesty. There was a main effect of proclamation, $F(2, 592) = 78.91, p < .001, \eta_p^2 = .21$, such that communicators who endorsed the absolute honesty stance were viewed as more likely to engage in future honesty than communicators who endorsed the ambitious honesty stance, $t(392) = 3.55, p < .001, d = 0.36$, and communicators who endorsed the ambitious honesty stance were

viewed as more likely to engage in future honesty than communicators who endorsed the flexible honesty stance, $t(404) = 8.31, p < .001, d = 0.82$. There was also a main effect of behavior, $F(1, 592) = 33.83, p < .001, \eta_p^2 = .05$, such that communicators who told selfish truths were viewed as more likely to engage in future honesty than communicators who told prosocial lies.

These effects were qualified by a significant Proclamation \times Behavior interaction, $F(2, 592) = 3.60, p = .028, \eta_p^2 = .01$. Regardless of whether they told a selfish truth or prosocial lie, communicators who endorsed the absolute or ambitious honesty stance were viewed as more likely to engage in future honesty than communicators who endorsed the flexible honesty stance ($ts > 4.98, ps < .001$). However, this effect was larger if they subsequently told the truth.

Hypocrisy. We found a main effect of behavior, $F(1, 592) = 7.44, p = .007, \eta_p^2 = .01$, on hypocrisy ratings, such that communicators who told prosocial lies were viewed as more hypocritical than communicators who told selfish truths. There was no main effect of proclamation ($p = .874$).

These results were qualified by a significant Proclamation \times Behavior interaction, $F(2, 592) = 9.36, p < .001, \eta_p^2 = .03$. After telling a prosocial lie, communicators who endorsed the absolute or ambitious honesty stance were viewed as more hypocritical than communicators who endorsed the flexible honesty stance but similar to each other, absolute versus ambitious, prosocial lie: $t(203) = 0.63, p = .528, d = 0.09$; absolute versus flexible, prosocial lie: $t(200) = 3.07, p = .002, d = 0.43$; ambitious versus flexible, prosocial lie: $t(207) = 2.54, p = .012, d = 0.35$. The fact that ambitious stances seem hypocritical when paired with deception indicates that the ambitious stance communicates high honesty standards and is interpreted quite differently from the flexible stance. In other words, the ambitious stance does not protect against hypocrisy judgments to the same degree as the flexible stance—even though both stances may suggest that some lies are acceptable.

After telling a selfish truth, communicators who endorsed the absolute honesty stance were viewed as less hypocritical than

Table 4
Descriptive Statistics in Studies 4 and 5

Variable	Behavior	Proclamation		
		Absolute	Ambitious	Flexible
Study 4				
Morality	Prosocial lie	4.91 (1.47) _a	4.57 (1.63) _{a,b}	4.19 (1.49) _b
	Selfish truth	5.16 (1.38) _a	4.56 (1.51) _b	3.96 (1.36) _c
Hypocrisy	Prosocial lie	3.38 (1.65) _a	3.24 (1.52) _a	2.75 (1.25) _b
	Selfish truth	2.54 (1.42) _a	2.77 (1.45) _{a,b}	3.11 (1.19) _b
Future honesty	Prosocial lie	4.66 (1.20) _a	4.33 (1.25) _a	3.55 (0.99) _b
	Selfish truth	5.50 (1.19) _a	4.92 (1.30) _b	3.77 (0.94) _c
Study 5				
Morality	Selfish lie	2.62 (1.56) _{a,b}	2.89 (1.46) _a	2.39 (1.05) _b
	Prosocial truth	5.94 (0.99) _a	5.56 (0.99) _b	4.83 (1.12) _c
Hypocrisy	Selfish lie	5.42 (1.72) _a	4.79 (1.51) _b	3.49 (1.12) _c
	Prosocial truth	1.73 (0.95) _a	1.95 (0.93) _a	2.65 (1.09) _b
Future honesty	Selfish lie	2.86 (1.45) _{a,b}	3.03 (1.23) _a	2.60 (0.90) _b
	Prosocial truth	5.93 (0.96) _a	5.50 (0.90) _b	4.59 (0.93) _c
Social benefit of the proclamation	Selfish lie	4.90 (1.49) _a	4.78 (1.57) _a	2.69 (1.27) _b
	Prosocial truth	5.39 (1.33) _a	5.11 (1.35) _a	2.81 (1.27) _b

Note. Cells within the same row that share a subscript letter do not significantly differ. Cells that do not share a subscript letter significantly differ at $p < .05$.

communicators who endorsed the flexible honesty stance and similar to those who endorsed the ambitious honesty stance, absolute versus flexible, selfish truth: $t(192) = -3.05, p = .003, d = -0.44$; absolute versus ambitious, selfish truth: $t(187) = -1.09, p = .279, d = -0.16$. In other words, not only is it considered hypocritical to tell a lie after endorsing an absolute or ambitious stance on honesty but it is also considered hypocritical to (selfishly) tell the truth after endorsing a flexible stance on honesty. Presumably, if one has committed to the idea that lying is acceptable under some circumstances, then lying to help another person ought to be one

of those circumstances, leading to perceptions of hypocrisy when one opts *not* to lie, in order to benefit oneself.

Moderated Mediation Analysis

Bootstrapped moderated mediation analyses were conducted to examine mediators of the relationship between proclamation and trust using Hayes' PROCESS Macro for SPSS, Model 7 (Hayes, 2017). Results are presented in Table 5.

Table 5
Moderated Mediation Results on Trust From Studies 4 and 5

Study 4							
Mediator	Behavior						Index of moderated mediation
	Prosocial lie			Selfish truth			
	$X \rightarrow M$	$M \rightarrow Y$	Indirect effect	$X \rightarrow M$	$M \rightarrow Y$	Indirect effect	
Hypocrisy	-0.56**	-0.29**	0.23 [0.10, 0.37]	0.46**	-0.08	-0.19 [-0.32, -0.06]	0.42 [0.23, 0.62]
Future honesty	-0.94**	0.27*	-0.50 [-0.67, -0.25]	-1.43**	0.59**	-0.76 [-0.96, -0.58]	0.26 [0.07, 0.47]
Study 5							
Mediator	Behavior						Index of moderated mediation
	Selfish lie			Prosocial truth			
	$X \rightarrow M$	$M \rightarrow Y$	Indirect effect	$X \rightarrow M$	$M \rightarrow Y$	Indirect effect	
Hypocrisy	-1.63**	-0.21	0.17 [-0.12, 0.46]	0.81**	0.03	-0.08 [-0.23, 0.06]	0.25 [-0.17, 0.66]
Future honesty	-0.34*	0.55**	-0.19 [-0.37, -0.05]	-1.13**	0.33*	-0.64 [-0.94, -0.41]	0.45 [0.22, 0.75]
Social benefit of the proclamation	-2.16**	0.11	-0.19 [-0.51, 0.11]	-2.44**	0.15	-0.22 [-0.57, 0.13]	0.02 [-0.03, 0.10]

Note. All items in the first column were entered as simultaneous mediators, the dummy-coded flexible proclamation as the independent variable (1 = flexible, 0 = ambitious or absolute), trust in the RV game as the dependent variable (1 = rely, 0 = verify), and behavior as the moderator (Study 4: 1 = prosocial lie, 0 = selfish truth; Study 5: 1 = selfish lie, 0 = prosocial truth). We ran separate linear and logistic regressions to compute the $X \rightarrow M$ and $M \rightarrow Y$ pathways. Each "indirect effect" signifies the lower level and upper level 95% confidence intervals around the indirect effect for the corresponding mediator. Bold numbers indicate confidence intervals that do not contain zero. RV = rely-or-verify.
* $p < .05$. ** $p < .01$ for coefficients in these regressions.

We examined the role of hypocrisy and future honesty in mediating the relationship between a dummy-coded flexible proclamation (1 = flexible, 0 = ambitious/absolute) and the decision to rely on the words of a communicator in the RV game. Behavior was entered as a moderator (1 = prosocial lie, 0 = selfish truth) of the effects of proclamation on the mediators (i.e., the “a paths”). We used the flexible proclamation (vs. ambitious and absolute proclamations) as the comparison group because we saw more differentiation in interpersonal evaluations between communicators endorsing the flexible “sometimes” stance and those endorsing the ambitious “rarely” or the absolute “never” stance. The ambitious stance was interpreted more as an indicator of high honesty standards than as an indicator of flexibility.

When the communicator told a prosocial lie, we found that perceived hypocrisy mediates the effect of a flexible proclamation on trust, as does future honesty. When the communicator told a selfish truth, we also found that perceived hypocrisy mediates the effect of a flexible proclamation on trust, but in the opposite direction, as does future honesty. Therefore, word–deed (mis)alignment influenced how participants viewed hypocrisy, such that the communicators who endorsed absolute and ambitious proclamations were seen as more hypocritical when they went on to tell a prosocial lie but *less* hypocritical when they went on to tell the selfish truth. However, regardless of behavior type, hypocrisy was always detrimental for trust. The ambitious and absolute stance always signaled more future honesty relative to the flexible stance, which fostered trust. Overall, we find evidence of moderated mediation through perceived hypocrisy and future honesty (see Table 5, for indices of moderated mediation). We find similar results when we conduct the same analysis with moral evaluations entered as the dependent variable instead of trust (see Supplemental Material Appendix 1.4, for additional moderated mediation results).

Attitudes Toward Honesty

More participants indicated that lying is “sometimes” (33%) or “rarely” (50%) okay than “never” okay (16%). In our preregistered analysis, in which we added attitudes toward honesty to our ANOVAs on morality, hypocrisy, and future honesty, we found that attitudes toward honesty did not change the significance of any of our effects in our models (see Supplemental Material Appendix 1.4 for detailed results).⁴

Discussion

Study 4 provides evidence that flexible honesty proclamations undermine trust and moral judgments relative to both absolute and ambitious honesty proclamations because they signal lower future honesty. Further, ambitious stances, which allow for some degree of flexibility, did not significantly mitigate hypocrisy judgments relative to absolute stances when communicators told prosocial lies. Though lying after endorsing a strong stance on honesty (both absolute and ambitious) versus a flexible stance seems more hypocritical, it does not discredit high honesty standards entirely. We find evidence that perceptions of future honesty reflect positively on communicators, even hypocritical ones, and lead to more interpersonal trust, while showing that shared beliefs between the participant and communicator cannot fully explain our results.

Study 5: Selfish Lies and Prosocial Truths

In Study 5, we compare evaluations of communicators who tell selfish lies and prosocial truths to examine whether our finding in Study 4—that participants weighed the proclamation more heavily than the lie when deciding whom to trust—generalizes to selfish lies. We also examine whether beliefs that the flexible proclamation is less valuable for society than the absolute and ambitious proclamations helps explain the social costs of flexible honesty proclamations.

Method

Participants

We aimed to recruit 600 participants on MTurk. As in Study 4, comprehension check questions were included in the middle of the survey, before any dependent measures were collected, to assess understanding of the communicator’s behavior in the previous study and to verify understanding of instructions for the RV game. Participants who failed a minimum of two comprehension checks were automatically prevented from continuing with the survey. Five hundred ninety-six participants successfully passed comprehension checks and completed the full survey. We did not exclude any other participants, since there were no repeated MTurk IDs, resulting in the final sample of 596 participants (48.32% female, $M_{\text{age}} = 38.02$, $SD_{\text{age}} = 11.58$).

Procedure and Materials

We randomly assigned participants to one cell in a 3 (proclamation: absolute vs. ambitious vs. flexible) \times 2 (behavior: selfish lie vs. prosocial truth) between-subjects design.

The stimuli for Study 5 were identical to the stimuli used in Study 4 except that the deception game instructions were altered to vary the communicator’s motives for lying or truth-telling during the deception game. Telling a lie in the deception game helped the communicator and harmed the other player in the game (i.e., lying was selfish). Telling the truth harmed the communicator but helped the other player in the game (i.e., telling the truth was prosocial).

After learning about the communicator’s honesty stance and behavior, participants evaluated the communicator using the same measures of morality, future honesty, and hypocrisy as in Study 4 (all $\alpha \geq .92$). We also added two items to measure perceived social benefits of the proclamation. We asked participants about their beliefs

⁴ In further exploratory analyses, we find that participants’ own attitudes toward honesty did moderate the strength of our effects, though they did not reverse them (see Supplemental Material Appendix 1.4, for detailed results). Participants who endorsed the absolute honesty stance had stronger preferences for absolutism in others. Yet, even participants who endorsed the flexible honesty stance judged communicators who endorsed absolute honesty stances as more moral and more likely to engage in future honesty than communicators who also endorsed flexible honesty stances. Although we do find some evidence of similarity effects for trust (i.e., elevated trust toward communicators who hold the same belief as the participant), we still find relatively low levels of trust in communicators endorsing flexible honesty, even among participants who held flexible honesty stances themselves. Alongside our mediation results, this suggests that moral judgments of communicators who fail to live up to their ambitious or absolute stances on honesty are driven by expectations that communicators who take these stances are likely to engage in future honesty. However, trust may be driven both by beliefs about future honesty and a preference for communicators with similar views to one’s own.

about the communicator's proclamation using two questions, which were matched to condition: "To what extent do the following statements, '[absolute, ambitious, or flexible proclamation, depending on condition]' encourage positive behaviors in society?" and "To what extent do the following statements, '[absolute, ambitious, or flexible proclamation, depending on condition]' help people uphold valuable societal norms?" The two items were combined into one composite measure, $r(594) = .84, p < .001$. Participants answered these questions before playing the RV game and answered a question on their own attitudes toward honesty at the end of the survey.

Results

Trust in the RV Game

We regressed the decision to rely (1 = rely, 0 = verify) on dummy variables for the flexible (1 = flexible, 0 = ambitious or absolute) and ambitious (1 = ambitious, 0 = absolute or flexible) proclamations, behavior (1 = selfish lie, 0 = prosocial truth), and their interaction terms.

Results revealed a main effect of the flexible proclamation ($b = -0.92, p = .002$). Consistent with the results from Study 4, we found that participants were significantly less likely to rely on communicators who selected "Lying is sometimes okay" (31% rely) compared to communicators who selected "Lying is never okay" (47% rely), even after the communicators told equivalent lies. Participants did not significantly differ in their willingness to rely on communicators who selected "Lying is never okay" and "Lying is rarely okay" (42% rely).

There was also a significant main effect of behavior ($b = -2.00, p < .001$); prosocial truth-tellers were trusted more (60% rely) than selfish liars (20% rely). Although selfish lying was generally detrimental for trust, consistent with Studies 1 and 2, communicators who endorsed an absolute or ambitious stance on honesty were trusted more than actors who endorsed a flexible stance, even after telling a selfish lie (see Figure 4).

The main effect of the ambitious proclamation ($b = -0.48, p = .113$), Flexible Proclamation \times Behavior interaction ($b = 0.11, p = .821$), and the Ambitious Proclamation \times Behavior interaction ($b = 0.37, p = .416$) were not significant.

Attitudinal Measures

A series of ANOVAs were conducted to investigate the effects of proclamation (absolute vs. ambitious vs. flexible) and behavior (prosocial truth vs. selfish lie) on each attitudinal variable. Significant effects were followed up with independent-samples t tests, as were preregistered. Descriptive statistics for Study 5 are presented in Table 4.

Morality. There was a main effect of proclamation, $F(2, 590) = 18.36, p < .001, \eta_p^2 = .06$, on moral evaluations. Communicators who endorsed absolute or ambitious honesty were viewed similarly, $t(396) = -0.18, p = .860, d = -0.02$, and as more moral than communicators who endorsed the flexible honesty stance, absolute versus flexible proclamation: $t(397) = 3.35, p < .001, d = 0.34$; ambitious versus flexible proclamation: $t(393) = 3.85, p < .001, d = 0.39$. There was a main effect of behavior, $F(1, 590) = 792.31, p < .001, \eta_p^2 = .57$, such that communicators who told prosocial truths were viewed as more moral than communicators who told selfish lies.

These effects were qualified by a significant Proclamation \times Behavior interaction, $F(2, 590) = 6.97, p = .001, \eta_p^2 = .02$. The

flexible honesty proclamation was penalized relative to both absolute and ambitious honesty when communicators told a prosocial truth, absolute versus flexible, prosocial truth: $t(192) = 7.29, p < .001, d = 1.05$; ambitious versus flexible, prosocial truth: $t(196) = 4.84, p < .001, d = 0.69$. When telling a selfish lie, the flexible stance was only penalized relative to the ambitious stance, ambitious versus flexible, selfish lie: $t(195) = 2.75, p = .006, d = 0.39$; absolute versus flexible, selfish lie, $t(203) = 1.23, p = .220, d = 0.17$. Though the flexible stance was seen as no worse than the absolute stance when communicators told a selfish lie, it was also not seen as any better, even in the face of clear and unabashed hypocrisy.

Future Honesty. There was a significant main effect of proclamation, $F(2, 590) = 30.94, p < .001, \eta_p^2 = .09$, such that communicators who endorsed flexible honesty were viewed as less likely to engage in future honesty than communicators who endorsed absolute, $t(397) = 4.52, p < .001, d = 0.45$, or ambitious, $t(393) = 4.78, p < .001, d = 0.48$, honesty. There was also a main effect of behavior, $F(1, 590) = 793.17, p < .001, \eta_p^2 = .57$. Communicators who told prosocial truths were viewed as more likely to engage in future honesty than communicators who told selfish lies.

These effects were qualified by a significant Proclamation \times Behavior interaction, $F(2, 590) = 12.42, p < .001, \eta_p^2 = .04$. Unsurprisingly, the absolute stance signaled the greatest likelihood of future honesty, followed by the ambitious stance, when communicators told a prosocial truth, absolute versus flexible, prosocial truth: $t(192) = 9.89, p < .001, d = 1.42$; ambitious versus flexible, prosocial truth: $t(196) = 7.08, p < .001, d = 1.01$. However, after telling selfish lies, only the ambitious honesty stance conveyed greater future honesty than the flexible honesty stance, absolute versus flexible, selfish lie: $t(203) = 1.51, p = .132, d = 0.21$; ambitious versus flexible, selfish lie: $t(195) = 2.75, p = .007, d = 0.39$. In the context of selfish lies, the absolute stance was discredited to some degree. Notably, however, even in a case where a communicator is being dishonest about their dishonesty for *selfish* reasons, the absolute stance did not lower expectations of future honesty relative to the flexible stance.

Hypocrisy. There was a main effect of proclamation, $F(2, 590) = 8.21, p < .001, \eta_p^2 = .03$, such that communicators who endorsed absolute honesty were seen as more hypocritical than communicators who endorsed flexible honesty, $t(397) = 3.05, p = .002, d = 0.31$. However, communicators who endorsed ambitious honesty were not judged significantly differently than actors who endorsed absolute honesty, $t(396) = 1.44, p = .150, d = 0.14$, or flexible honesty, $t(393) = 1.62, p = .105, d = 0.16$. There was also a main effect of behavior, $F(1, 590) = 569.36, p < .001, \eta_p^2 = .49$. Communicators who told selfish lies were seen as more hypocritical than communicators who told prosocial truths.

These main effects were qualified by a significant Proclamation \times Behavior interaction, $F(2, 590) = 67.15, p < .001, \eta_p^2 = .19$. Communicators who endorsed absolute honesty and lied were seen as more hypocritical than communicators endorsing both ambitious and flexible honesty prior to lying, absolute versus ambitious, selfish lie: $t(198) = 2.72, p = .007, d = 0.38$; absolute versus flexible, selfish lie: $t(195) = 6.93, p < .001, d = 0.99$. However, after telling a prosocial truth, communicators who endorsed *flexible* honesty were seen as more hypocritical than both other types of communicators, absolute versus flexible, prosocial truth: $t(192) = -6.23, p < .001, d = -0.89$; ambitious versus flexible, prosocial truth: $t(196) = -4.89, p < .001, d = -0.69$.

The fact that truth-telling appears hypocritical after taking a flexible honesty stance adds further support to the notion that flexible proclamations convey a low likelihood of future honesty.

Social Benefit of the Proclamation. There was a main effect of proclamation on perceived social benefit, $F(2, 590) = 182.62, p < .001, \eta_p^2 = .38$. The absolute, $t(397) = 17.62, p < .001, d = 1.76$, and ambitious, $t(393) = 15.95, p < .001, d = 1.61$, stances were both thought to carry more positive social influence than the flexible stance, and they did not significantly differ from each other in perceived social benefit, $t(396) = 1.29, p = .198, d = 0.13$. There was also a main effect of behavior, $F(1, 590) = 7.72, p = .006, \eta_p^2 = .01$. Although our measures of perceived social benefit specifically asked about the proclamation *itself*, participants thought that communicators who told prosocial truths would have more positive social influence than communicators who told selfish lies. The Proclamation \times Behavior interaction was not significant ($p = .427$).

Moderated Mediation Analysis

We conducted similar analyses to those in Study 4 (note the behavior dummy is now coded 1 = selfish lie, 0 = prosocial truth) but added an additional mediator, the social benefit of the proclamation. Results are depicted in Table 5.

When the communicator told a selfish lie, future honesty mediated the effect of a flexible proclamation on trust. Neither hypocrisy nor the social benefit of the proclamation significantly mediated the relationship between proclamation and trust after a selfish lie. When the communicator told a prosocial truth, future honesty mediated the effect of a flexible proclamation on trust. Neither hypocrisy nor the social benefit of the proclamation mediated the relationship between the proclamation and trust.

Consistent with our findings in Study 4, we find evidence of moderated mediation through future honesty (see Table 5, for indices of moderated mediation), and the indirect effect of future honesty is over three times larger in the prosocial truth condition compared to the selfish lie condition. However, we find no evidence of moderated mediation through hypocrisy or the social benefit of the proclamation. When we conduct the same analysis with moral evaluations entered as the dependent variable, we find evidence of moderated mediation through both future honesty and hypocrisy, but still no evidence of moderated mediation through the social benefit of the proclamation (see Supplemental Material Appendix 1.5, for additional results and Supplemental Study 7, for further inconsistent results related to the social benefit of the proclamation).

Attitudes Toward Honesty

Consistent with our results from Pilot Studies 1–3 and Study 4, we found that a relatively small minority of participants believe that lying is never okay. Participants indicated that lying is “sometimes” (24%) or “rarely” (55%) okay more often than they indicated that lying is “never” okay (18%). In our preregistered analyses, adding attitudes toward honesty to the ANOVA models did not change the pattern of results for any of the character evaluations. Most participants viewed communicators who endorsed the absolute or ambitious honesty stance as the most moral, and even those who endorsed flexible honesty themselves tended to evaluate the communicators endorsing ambitious honesty most positively (see Supplemental Material Appendix 1.5, for more details).

Discussion

Study 5 established the robustness of our effects across various types of lies and truth-telling behaviors, including selfish lies. Flexible moral stances led to less trust than absolute and ambitious moral stances, suggesting that people view moral proclamations as credible signals of one’s actual likelihood of behaving honestly in the future. In Study 5, however, only ambitious honesty stances signaled greater future honesty and morality than flexible honesty stances when paired with (selfish) lies. This finding suggests that some degree of nuance (i.e., *rarely*) can be ideal, depending on the relative costs of flexibility and hypocrisy. Notably though, the blatantly hypocritical absolute stance was not seen as less moral and was actually seen as more trustworthy than the flexible stance.⁵

Taken together, Studies 4 and 5 emphasize the importance of communicating high honesty standards. Both ambitious and absolute honesty stances convey higher honesty standards than the flexible stance, which in turn predicts trust. In Study 4, the communicator’s professed stance on honesty predicted trust more than moral consistency. In Study 5, behavior did matter in that selfish lying detracted from trust, but nevertheless, communicators endorsing ambitious and absolute honesty were still trusted more than communicators endorsing flexible honesty. We find consistent support that inferences of future honesty are predictive of trust and morality, while also finding that relative preferences for more ambitious moral stances do not seem to be driven entirely by similarity preferences nor by the perceived social benefits of the proclamation.

Study 6: Do Leaders Anticipate the Consequences of Hypocritical Absolutism?

Studies 1–5 provide evidence that moral flexibility can be more costly than hypocrisy. In Study 6, we ask if communicators, and government leaders in particular, anticipate these effects. If leaders recognize the costs of moral flexibility, relative to hypocritical absolutism, this can help to explain the persistence of hypocritical absolutism in the public domain. To examine this possibility, we surveyed publicly elected or appointed government officials in the United States to determine whether these leaders predict that endorsing moral flexibility will reduce trust to a greater extent than engaging in moral hypocrisy.

Method

Participants

We surveyed government officials through the nonprofit CivicPulse, which recruits officials from all 50 states and across the political spectrum. We asked CivicPulse to recruit 100 officials from the following 12 local government positions: head building official,

⁵ We explore reactions to selfish liars who make absolute or flexible honesty proclamations further in Supplemental Studies 2 and 5 (see also Supplemental Studies 4 and 7, which feature selfish lies in the political context, similar to Studies 1 and 2). We find that communicators who tell selfish lies after making absolute or flexible honesty proclamations are not viewed as differentially moral in Supplemental Study 2, though the communicators endorsing absolute honesty are more moral in Supplemental Study 5, similarly to Studies 1 and 2 (see Supplemental Appendices 2.2 and 2.5). Behaving hypocritically for selfish reasons can be more costly than behaving hypocritically for prosocial reasons (Jordan & Sommers, 2022), but even so, hypocritical selfish behaviors are not *more* costly than proclamations of moral flexibility.

head of communications, head of human resources, head of finance, head of clerk's office, head of purchasing/procurement, head of fire protection services, head of public works, head of law enforcement, top appointed officials, governing board members, and top elected officials. Our final participant sample represents a range of political positions and affiliations ($N = 166$) and was larger than expected, partly due to the number of partial survey completions in our sample (CivicPulse recruited 100 fully completed surveys, but some participants only answered a subset of questions). We analyze partial responses in this study, unlike in Studies 1–5. Of the 166 participants included in this sample, 112 participants submitted fully completed surveys with no “nonapplicable” answers to any questions. The additional 12 full completes were the result of a higher-than-expected response rate to the final survey invitation. Of the full sample, 30% identified as female (10% did not answer the question about gender), and 25% identified as Democrat and 24% identified as Republican (33% did not answer the question about political affiliation and 18% were “other” or “independent”).

Procedure and Materials

The experimental items we report here were asked alongside the attitudes toward honesty measures that were assessed in Pilot Study 3.⁶ Participation in the survey, as well as for each individual question was voluntary, and some participants chose to only answer a subset of questions in the survey. Consequently, the final number of participants included in each analysis varies across measures.

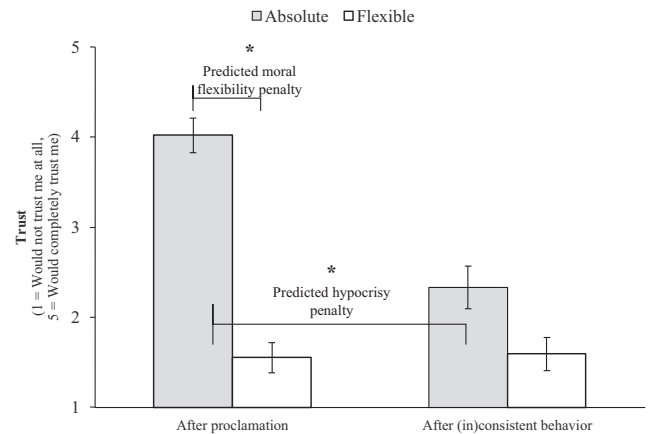
Participants were asked to predict the extent to which their constituents would trust them if a media source revealed that they had engaged in deception after they took an absolute or a flexible stance on honesty. This was a within-subjects design, and the order in which participants were asked to imagine each scenario (absolute vs. flexible proclamation) was randomized across participants.

The description of each scenario was very similar to the stimuli in Studies 1 and 2, but instead of describing a hypothetical candidate named Matthew Johnson and asking participants to evaluate Mr. Johnson, we asked government officials to imagine that they had taken the absolute or flexible honesty stance during a television interview. We used the exact same language for the absolute and flexible stances as in Study 2. Participants were asked to predict constituent trust following each honesty stance. Then, we asked participants to imagine that a media source had published an article stating that they (the participant) had lied in office after taking this honesty stance. We then asked participants to predict constituent trust a second time. Overall, participants made four predictions of trust: 1a) Prediction of trust after taking the absolute honesty stance, 1b) Prediction of trust following the revealed deception after taking the absolute honesty stance, 2a) Prediction of trust after taking the flexible honesty stance, 2b) Prediction of trust following the revealed deception after taking the flexible honesty stance. Trust was measured with a single item, using a five-item scale (1 = *would not trust me at all* to 5 = *would completely trust me*).

Results

A series of paired t tests were conducted to investigate the effects of the honesty stance (absolute vs. flexible) and deception on trust. A total of 145 participants answered both items assessing trust in the absolute honesty scenario and 145 participants answered both items

Figure 5
Predicted Trust Following an (Absolute or Flexible) Proclamation and Revealed Deception



Note. Error bars reflect 95% confidence intervals. *reflects significant differences between cells.

assessing trust in the flexible honesty scenario. However, the same 145 participants did not respond to all four items, and therefore the sample used to compare predictions of trust between the absolute and flexible honesty conditions included only the 138 participants who answered all four trust items across the absolute and flexible honesty scenarios.

Participants predicted that constituents would trust them more if they took an absolute honesty stance than if they took a flexible honesty stance before any deception was revealed, $t(137) = 19.48$, $p < .001$, $d_{RM} = 1.66$. Participants also predicted that constituents would trust them less when deception was revealed, relative to before deception was revealed, if they had previously taken an absolute stance on honesty, $t(144) = 12.38$, $p < .001$, $d_{RM} = 1.03$. This finding suggests that participants recognize that hypocrisy is costly; the much higher trust attributed to the initial absolute stance (before any hypocrisy) indicates that participants think that absolute stances are particularly beneficial if they can actually live up these statements (or at least, if they can avoid getting caught when they violate them). Conversely, participants did not think that constituent trust would change when deception was revealed if they had previously taken a flexible stance, $t(144) = -0.49$, $p = .628$, $d_{RM} = -0.04$. However, participants still thought constituents would trust them more if they lied after taking a hypocritical, absolute stance than a consistent, flexible stance, T2 ratings: $t(137) = 5.26$, $p < .001$, $d_{RM} = 0.45$; see Figure 5, highlighting that people recognize the value of communicating high honesty standards even when hypocrisy is known and publicized.

Discussion

Leaders in the United States predicted that the costs of moral flexibility would be greater than the costs of moral hypocrisy. These

⁶ Originally, the module containing Study 6 was asked before the module containing Pilot Study 3. However, this order was reversed after the first nine participants. Therefore, 95% of participants completed Pilot Study 3 prior to Study 6, and 5% completed the items for Study 6 prior to Pilot Study 3.

officials still recognized that hypocrisy would undermine public trust, but nevertheless, they expected that lying after endorsing absolute honesty would lead to greater overall trust than behaving consistently with a flexible honesty stance. Leaders expected this even when it was made clear that the hypocrisy would be publicized by a media source, suggesting that leaders thought there was a benefit to communicating high honesty standards even if they failed to live up to this standard and this failure was widely broadcast. The fact that leaders correctly intuit the costs of being honest about dishonesty helps to explain why people continue to invoke absolute moral language in their public stances, even if their behaviors, and often their private beliefs, do not live up to these lofty declarations. It is both expected to be, and *is*, true that the costs of admitting to moral flexibility are at least as great as the costs of hypocrisy, making it clear to communicators that publicly endorsing flexibility is an unwise strategy when asked about one's stance on honesty.

General Discussion

Across six experiments, we find that taking flexible stances on honesty is detrimental for trust, voting intentions, and moral evaluations. In Study 1, political candidates who made absolute honesty proclamations and lied for selfish reasons were seen as more moral, and received greater political support, than candidates who made more flexible proclamations that were consistent with their behavior. We find similar results in Study 2, and by comparing judgments of candidates who took absolute and flexible stances over time, demonstrate that hypocrisy is penalized once it is revealed, but the initial costs of taking a flexible stance overwhelm this penalty. Furthermore, candidates who endorsed flexibility in Study 2 were evaluated more negatively than candidates taking no honesty stance, highlighting the absolute costs of flexibility. Study 3 builds on Study 2, replicating these across-time results with a new paradigm in which a communicator told a prosocial lie rather than a selfish lie. Interestingly, even when flexible stances could be interpreted as morally motivated because flexible stances allow for ethical, prosocial lies, they are still penalized. We find this to be the case regardless of the order in which a communicator's stance and deception is revealed. Across these three studies, mediation results reveal that perceptions of communicators' future honest behavior offset the costs of hypocrisy, leading to more negative judgments of communicators who took flexible, rather than absolute, stances, regardless of the lies that followed.

In Studies 4 and 5, we further explored the mechanisms underlying, and boundaries of, our effects. First, we find additional support for the mechanism of future honesty by highlighting that ambitious, nonabsolute stances that also signal a high likelihood of future honesty are viewed positively. As long as one commits to a high moral standard (even if it is not truly absolute), one can avoid a moral flexibility penalty. Moreover, we rule out shared beliefs as a potential mechanism in these studies; most participants in our studies (and Pilot Studies 1–3) did not endorse the absolute stance themselves, despite judging communicators who endorse absolute honesty favorably. In Study 5, we also examine whether participants reward absolute honesty stances because they promote important norms in society (i.e., because they are socially beneficial). However, we do not find that this inference mediates our effects (consistent with Supplemental Study S7, see Supplemental Material Appendix 2.7).

In addition to considering these mechanisms, we also ruled out two other potential explanations, guilt and construal of deception, in our

Supplemental Studies. Previous work suggests that feelings of guilt can reduce moral hypocrisy (Polman & Ruttan, 2012) and may even promote positive behaviors as a form of moral cleansing (e.g., Tetlock et al., 2000). Consistent with this work, we considered whether communicators who endorsed absolute honesty and behaved inconsistently were perceived as feeling guilty for their deception, and whether this inference also increased moral evaluations. However, we find inconsistent evidence that guilt mediates the relationship between proclamation and morality (see Supplemental Results for Study 3 and Supplemental Studies 1 and 2 in Supplemental Material Appendix 1.3, Supplemental Material Appendix 2.1 and 2.2). Likewise, we thought participants might have believed that communicators who endorsed absolute stances were less likely to construe their deceptive actions as actual lies; however, we do not find support for this notion in Supplemental Study 3 (see Supplemental Material Appendix 2.3). Rather, our studies reveal that flexible moral stances lead to negative judgments of morality, conditional on taking a stance of any kind, because moral proclamations influence perceptions of future honesty.

In Studies 4 and 5, we found that both the ambitious and absolute stances were viewed as a strong signal of future honesty regardless of the ultimate behavior these stances were paired with. We found that preferences for absolute and ambitious honesty proclamations persisted across prosocial truths, selfish truths, prosocial lies, and even selfish lies. Although communicators endorsing absolute honesty stances were not seen as significantly more moral than communicators endorsing flexible honesty stances when telling selfish lies in Study 5 (also consistent with results in Supplemental Studies 2 and 7, see Supplemental Material Appendices 2.2 and 2.7), they were also not seen as significantly *less* moral, and they were still trusted more. These results suggest that flexible honesty stances are rarely beneficial, relative to taking a more ambitious stance, but the strength of these effects may still depend on the nature of the hypocrisy. Studies 4 and 5 also assessed the behavioral consequences of absolute proclamations. Communicators who took an absolute stance on honesty and then lied could be seen as having lied about their stance on lying. Although participants could have discounted these communicators' moral proclamations entirely after encountering such word–deed misalignment, they did not. Even with real bonuses at stake, participants were more likely to rely on the words of communicators who hypocritically claimed that lying was never okay than they were to rely on the words of communicators who consistently claimed that lying was sometimes okay.

To determine the robustness of our findings, we conducted a meta-analysis on all of our studies that compared moral evaluations of communicators endorsing absolute or flexible honesty. Our meta-analysis included $k = 10$ studies: Studies 1–5 from the main article, and Supplemental Studies 2–5 and 7 (reported in Supplemental Material Appendices 2.2–2.5, 2.7). Study 6, Supplemental Studies 1 and 6, and Pilot Studies 1–4 were not included because these studies do not compare moral evaluations of communicators endorsing absolute versus flexible honesty. From each study, we included only the conditions that examined judgments of communicators endorsing absolute and flexible honesty stances, with additional control conditions or conditions examining ambitious stances removed, and collapsing across other manipulations (e.g., manipulations of frame in Study 1, manipulations of behavior in Studies 4 and 5). For Studies 2 and 3 and Supplemental Study 5, we only included the results from T2, when participants had full information about the communicator's proclamation and behavior, because T2 judgments

are more directly comparable to our other studies. Last, we included only the American sample from Supplemental Study 5, excluding the Indian sample, because it was our only study with a cross-cultural sample. The meta-analysis, therefore, provides an estimate of the overall size of our primary finding in the United States, but as we discuss in Supplemental Study 5, this result may differ across cultures. The meta-analysis was conducted using the metafor package for R (Viechtbauer, 2010). We fit a random-effects model using inverse-variance weighting (Lipsey & Wilson, 2001) and the between-subjects Cohen's d as our effect size measure (Morris & DeShon, 2002). Overall, across the 10 studies (total $N = 4,300$), we found consistent evidence that, even in the presence of hypocrisy, taking flexible honesty stances led to more negative judgments of morality than taking absolute stances, meta-analytic $d = 0.30$, $p = .001$, 95% CI [0.12, 0.48].

In Study 6, we provide real-world evidence that leaders in the United States anticipate these effects and recognize that moral nuance is costly. In doing so, Study 6 helps to explain the persistence of hypocrisy in social exchanges. Although behaving in line with an absolute moral stance may not always be attainable, it is still beneficial to endorse these lofty norms rather than taking a more flexible moral stance.

Contributions

This work has a number of important implications. First, it expands our theoretical and practical understanding of hypocrisy. Past hypocrisy research suggests that word–deed misalignment should contribute to perceptions of hypocrisy, which is consistent with our findings. The vast majority of this work also suggests that perceptions of hypocrisy should predict social punishment (Effron, O'Connor, et al., 2018), which is somewhat inconsistent with our findings. Communicators who endorsed absolute or ambitious honesty stances and then exhibited word–deed misalignment *were* viewed as hypocritical, but they were still trusted more than communicators who admitted that lying is sometimes okay. Though some recent work has recognized that word–deed misalignment may not be penalized if it is not interpreted as hypocrisy (Effron, O'Connor, et al., 2018), does not involve public moral engagement (Jordan & Sommers, 2020), or if communicators admit to their own failings (Jordan et al., 2017), we find that word–deed misalignment can be preferred to consistency even when it *is* interpreted as hypocrisy, does involve public moral proclamations, and communicators do *not* acknowledge their own inconsistencies. Even though hypocrisy does have costs, a previously overlooked alternative to hypocrisy—admitting to nuance—can have even greater costs, allowing hypocrisy to persist as a superior social strategy, at least in some circumstances. Communicators who made absolute and ambitious honesty proclamations were viewed positively not because they were not seen as hypocritical, but rather, *despite* being seen as hypocritical. Though hypocrisy was not costless, moral flexibility was worse.

Ironically, our results also suggest that communicators' moral stances may be the least informative when observers are most eager to learn them. People likely inquire about communicators' moral stances when they are particularly motivated to understand the communicator's underlying character (e.g., when interviewing a job candidate, or questioning a politician). However, by simply asking the question, the asker puts any imperfect communicator in a

double bind: the communicator can either take a more absolute moral stance that contradicts their behavior or the communicator can take a more honest, realistic, and consistent stance. Our results from Pilot Studies 2 and 3, and the results of Study 6, suggest that communicators know that they should inflate their moral stances rather than admitting to moral nuance, at least in the context of honesty, and they also are likely to recognize that their own words and deeds do not always align. Therefore, stated moral stances are likely uninformative in this context. Nonetheless, we find that people in this context do *not* discredit absolute stances entirely. Although hypocrisy is costly when actors freely offer their moral stance or engage in condemnation, relative to neutral information (e.g., Jordan et al., 2017), hypocrisy is relatively beneficial when the alternative is admitting moral flexibility.

This work also helps to resolve two puzzles in the psychology and communication of honesty. First, people publicly claim lying is wrong, while privately practicing it, and even justifying it. Study 6 offers one explanation for these divergent preferences regarding the communication and execution of honesty: leaders and communicators correctly anticipate that admitting to moral nuance will lower trust. Second, people want *others* to claim lying is wrong, but they also want others to occasionally lie to them (Levine, 2021). For example, people may want their spouse to say that they would never lie, but these same people might still prefer that their spouse offer them false praise on a new haircut. Similarly, in Study 4, we find that participants relied on communicators' honesty stances more than their (dis)honest behavior when deciding whether to trust the communicator to be honest in the future. Our results suggest that people do tolerate flexibility in (dis)honest behavior, but not in moral claims, because they see moral claims as more diagnostic of future behavior. Therefore, this research also highlights a potential difficulty in moral education and rhetoric regarding honesty. Divergent preferences for honest behavior and communication about honesty likely preclude open discussion of moral nuance in both personal and public life.

Limitations and Future Directions

Our investigation focused specifically on the context of honesty. Future work should investigate attitudes toward absolutism and flexibility in other moral domains. As discussed, absolute hypocritical stances on honesty may go unpunished—in part—because people believe that some dishonesty is unavoidable, or even ethical at times. People expect others to endorse and idealize absolute honesty because this is a strong norm but also are likely to expect that people do not behave perfectly in line with this norm. This dynamic can help explain why hypocrites' words are not entirely discredited in the case of honesty. However, past hypocrisy research has focused on word–deed misalignment for more discretionary behaviors that can realistically be totally avoided, such as illegal drug use or bicycle theft (Alicke et al., 2013; Polman & Ruttan, 2012). Hypocrisy may be more costly when it comes to discretionary behaviors, if people expect that alignment between words and deeds is relatively feasible, whereas realistic proclamations of nuance may be especially costly in domains in which people *expect* some degree of word–deed misalignment anyway.

Likewise, for claims about moral behavior that are more specific and, consequently, easier to uphold, hypocrisy may also be more costly. We tested this possibility in Supplemental Study 6 by comparing evaluations of political candidates who claimed, "It is never okay to

lie about accepting political action committee (PAC) donations,” (our scenario referenced political action committee donations, which the communicator refers to as “PAC” donations) before lying about accepting PAC donations to evaluations of candidates who broadly stated, “It is never okay to lie,” before lying about accepting PAC donations. Avoiding deception about PAC donations, specifically, is much easier than avoiding deception in *all* aspects of life, and therefore, we thought hypocrisy between this specific claim and behavior may seem especially immoral and self-serving. Although we expected that this very specific word–deed misalignment might be penalized to a greater degree than word–deed misalignment between a generalized moral norm and a specific instance of contradictory behavior, we did not find support for this hypothesis (see Supplemental Study 6 in Supplemental Material Appendix 2.6). However, we cannot rule out the possibility that generalized moral norms might function differently than specific moral statements based on this one example, and future research should further explore the potential costs and benefits of more specific moral stances.

It is also possible that word–deed misalignment will be penalized more when the transgressions are more frequent or more severe. Communicators in our main studies told one lie, and these lies were relatively minor. In Supplemental Study 3 (see Supplemental Material Appendix 2.3), we also examined the impressions of communicators who told repeated lies. Interestingly, we found that participants still inferred greater future honesty from communicators who took absolute stances on honesty before telling many lies, relative to communicators who took flexible stances and told many lies. It is possible, however, that absolute claims paired with a very large amount of lies, or lies of a greater magnitude, would eventually be punished more than admissions of moral nuance.

It will also be important for future work to examine when flexible stances might escape social penalty. Though we find that participants were accepting of a very small degree of flexibility in others, as evidenced in the evaluations of communicators who took the ambitious stance (“Lying is rarely okay”), these ambitious stances were interpreted more in line with the absolute stance than the flexible stance, since they conveyed high honesty standards. More work is needed to understand the boundaries between full flexibility (“Lying is sometimes okay”) and more ambitious statements. Moral flexibility might be seen as more acceptable when paired with specific conditionals that highlight *why* lying is sometimes acceptable. For example, we had thought that saying “It is sometimes okay to lie if it prevents harm,” might be viewed more favorably than broader flexible moral proclamations. We began to investigate this possibility in Pilot Study 4 (reported in detail in Supplemental Material Appendix 3) by comparing evaluations of communicators who endorsed absolute honesty and those who endorsed flexible honesty with different caveats (i.e., “It’s sometimes okay to lie, if you have a good reason”; “It’s sometimes okay to lie, but only in really exceptional circumstances”; “It’s sometimes okay to lie, if it prevents harm”). We find that adding some of these caveats (“It’s sometimes okay to lie, but only in really exceptional circumstances”; “It’s sometimes okay to lie, if it prevents harm”) does make communicators seem more moral than endorsing flexibility without a caveat (“It’s sometimes okay to lie”), but even so, communicators endorsing absolute honesty still seem more moral and more likely to engage in future honesty than these communicators endorsing flexibility with a caveat. Furthermore, adding these caveats to the *never* statement, which we thought might be a way of introducing

flexibility that seems more tolerable (e.g., “It’s never okay to lie, except in really exceptional circumstances”), does not elevate impressions of communicators relative to simply taking the position that lying is never okay (and it can even detract at times). Although we only examined initial impressions of communicators based on their stances (there was no deceptive behavior in this pilot study), this pattern of evaluations suggests that the costs of flexibility are quite robust. Furthermore, a comparison of just the absolute and flexible stances to a control communicator in this pilot lends additional support to our theoretical model: the absolute honesty stance made the communicator seem similarly moral to the control communicator when no behavior was known, $t(56) = 1.72, p = .090, d = 0.45$, but a flexible honesty stance made a communicator seem less moral than a control communicator, $t(58) = -3.80, p < .001, d = -0.98$. This pattern shows that flexibility, specifically, is costly. A deeper examination of moral proclamations across domains and phrasing can help identify whether making broad, moralized claims functions differently than other types of proclamations, and whether the relative benefits of absolute stances are specific to stances on honesty.

Future work should also examine whether these results replicate in more varied contexts, as well as in more naturalistic settings. Past research suggests that hypocrisy is viewed more negatively in independent cultures than interdependent cultures (Effron, Markus, et al., 2018). Even though our participants came from predominantly independent cultures, we did explore potential cultural moderation in Supplemental Study 5 by recruiting participants from both the United States and India. We asked participants to evaluate communicators who took absolute or flexible stances on honesty and told selfish lies, at two points in time (once after just the stance and a second time after the selfish lie was later revealed), as in Study 2. Consistent with our other studies, participants viewed the communicators who endorsed absolute honesty as more moral and more likely to engage in future honesty, compared to communicators who endorsed flexible honesty, at T1. When participants learned that the communicators endorsing absolute honesty then told selfish lies, they viewed these communicators as hypocritical, which detracted from evaluations of morality and future honesty. This overall pattern emerged in both countries (see Supplemental Material Appendix 2.5 for full results, and OSF for full materials); however, the American participants penalized communicators endorsing absolute honesty for their hypocrisy to a greater extent than the Indian participants. As a result, moral evaluations of the hypocritical absolutist communicators at T2 were lower in the American sample, suggesting that participants from the more independent culture (the United States) were more bothered by word–deed misalignment relative to the participants from the more interdependent culture (India), consistent with previous findings (Effron, Markus, et al., 2018). However, we are hesitant to draw strong cultural conclusions given that we did not include measures of individual mindset or cultural identification in this study and only compared participants from two countries. Therefore, future work should continue to consider cross-cultural differences in perceptions of hypocrisy. In addition to considering cultural context, moral proclamations and accompanying behavior may be evaluated differently across settings based on the features of the communicator. For example, people may interpret moral proclamations differently when made by a politician, a parent, or a peer based on their expectations of how these people should communicate moral values. Future research should explore the importance of these contextual factors on the perceived credibility of moral proclamations.

Conclusion

Although more work remains to be done, the present studies identify a robust social cost to taking a flexible stance on honesty, compared to taking an absolute or ambitious stance. Taking a flexible stance on honesty that is consistent with one's behavior is seen as less hypocritical than taking an absolute stance, yet it also undermines perceptions of one's future honesty, which leads to negative judgments of moral character and decreased trust. As a result, being honest about the acceptability of dishonesty is—and is expected to be—an ineffective social strategy.

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>
- Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the “tragedy of the commons”. *Evolution and Human Behavior*, 25(4), 209–220. <https://doi.org/10.1016/j.evolhumbehav.2004.04.002>
- Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution and Human Behavior*, 27(5), 325–344. <https://doi.org/10.1016/j.evolhumbehav.2006.01.003>
- 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>
- Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology*, 8(2), 111–121. <https://doi.org/10.1037/1089-2680.8.2.111>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142. <https://doi.org/10.1006/game.1995.1027>
- DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, 74(1), 63–79. <https://doi.org/10.1037/0022-3514.74.1.63>
- Effron, D. A., Lucas, B. J., & O'Connor, K. (2015). Hypocrisy by association: When organizational membership increases condemnation for wrongdoing. *Organizational Behavior and Human Decision Processes*, 130, 147–159. <https://doi.org/10.1016/j.obhdp.2015.05.001>
- Effron, D. A., Markus, H. R., Jackman, L. M., Muramoto, Y., & Muluk, H. (2018). Hypocrisy and culture: Failing to practice what you preach receives harsher interpersonal reactions in independent (vs. interdependent) cultures. *Journal of Experimental Social Psychology*, 76, 371–384. <https://doi.org/10.1016/j.jesp.2017.12.009>
- Effron, D. A., O'Connor, K., Leroy, H., & Lucas, B. J. (2018). From inconsistency to hypocrisy: When does “saying one thing but doing another” invite condemnation? *Research in Organizational Behavior*, 38, 61–75. <https://doi.org/10.1016/j.riob.2018.10.003>
- 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(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Gino, F., & Schweitzer, M. E. (2008). Blinded by anger or feeling the love: How emotions influence advice taking. *Journal of Applied Psychology*, 93(5), 1165–1173. <https://doi.org/10.1037/0021-9010.93.5.1165>
- Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, 95(1), 384–394. <https://doi.org/10.1257/0002828053828662>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Hildreth, J. A. D., & Anderson, C. (2018). Does loyalty trump honesty? Moral judgments of loyalty-driven deceit. *Journal of Experimental Social Psychology*, 79, 87–94. <https://doi.org/10.1016/j.jesp.2018.06.001>
- Hoffman, M., Yoeli, E., & Nowak, M. A. (2015). Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences of the United States of America*, 112(6), 1727–1732. <https://doi.org/10.1073/pnas.1417904112>
- Jensen, S. L., Levine, E. E., White, M., & Huppert, E. (2021). *Lying is ethical, but honesty is the best policy: Introducing and testing a theory of moral error avoidance*. PsyArXiv. <https://doi.org/10.31234/osf.io/t5czy>
- Jordan, J. J., & Sommers, R. (2020, December 25). *False signaling and personal moral failings: Two distinct pathways to hypocrisy with unequal moral weight*. PsyArXiv. <https://doi.org/10.31234/osf.io/87txd>
- Jordan, J., & Sommers, R. (2022). When does moral engagement risk triggering a hypocrite penalty? *Current Opinion in Psychology*, 47, Article 101404. <https://doi.org/10.1016/j.copsyc.2022.101404>
- 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>
- Kim, P. H., Wiltermuth, S. S., & Newman, D. T. (2021). A theory of ethical accounting and its implications for hypocrisy in organizations. *Academy of Management Review*, 46(1), 172–191. <https://doi.org/10.5465/amr.2018.0161>
- 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>
- Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33(4), 378–392. <https://doi.org/10.1177/0261927X14535916>
- Levine, E. (2021, February 19). *Community standards of deception: Deception is perceived to be ethical when it prevents unnecessary harm*. PsyArXiv. <https://doi.org/10.31234/osf.io/g5trb>
- Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology*, 53, 107–117. <https://doi.org/10.1016/j.jesp.2014.03.005>
- Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, 126, 88–106. <https://doi.org/10.1016/j.obhdp.2014.10.007>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1), 105–125. <https://doi.org/10.1037/1082-989X.7.1.105>
- Nowak, M. A. (2006a). *Evolutionary dynamics: Exploring the equations of life*. Harvard University Press. <https://doi.org/10.2307/j.ctv9ghw98>
- Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563. <https://doi.org/10.1126/science.1133755>
- 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>
- Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 432(7016), 499–502. <https://doi.org/10.1038/nature02978>
- Polman, E., & Ruttan, R. L. (2012). Effects of anger, guilt, and envy on moral hypocrisy. *Personality and Social Psychology Bulletin*, 38(1), 129–139. <https://doi.org/10.1177/0146167211422365>
- Rockenbach, B., & Milinski, M. (2011). To qualify as a social partner, humans hide severe punishment, although their observed cooperativeness is decisive. *Proceedings of the National Academy of Sciences of the United States of America*, 108(45), 18307–18312. <https://doi.org/10.1073/pnas.1108996108>

- Rossee, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Shaw, A., DeScioli, P., Barakzai, A., & Kurzban, R. (2017). Whoever is not with me is against me: The costs of neutrality among friends. *Journal of Experimental Social Psychology*, 71, 96–104. <https://doi.org/10.1016/j.jesp.2017.03.002>
- Silver, I., & Shaw, A. (2022). When and why “staying out of it” backfires in moral and political disagreements. *Journal of Experimental Psychology: General*, 151(10), 2542–2561. <https://doi.org/10.1037/xge0001201>
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78(5), 853–870. <https://doi.org/10.1037/0022-3514.78.5.853>
- Van Zant, A. B., & Moore, D. A. (2015). Leaders’ use of moral justifications increases policy support. *Psychological Science*, 26(6), 934–943. <https://doi.org/10.1177/0956797615572909>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Zlatev, J. J. (2019). I may not agree with you, but I trust you: Caring about social issues signals integrity. *Psychological Science*, 30(6), 880–892. <https://doi.org/10.1177/0956797619837948>

Received April 28, 2021

Revision received December 12, 2022

Accepted December 14, 2022 ■