



## Faking It Is Hard to Do: Entrepreneurial Norm Enforcement and Suspicions of Deviance

Minjae Kim, Ezra W. Zuckerman Sivan

Massachusetts Institute of Technology

**Abstract:** Recent research suggests that many norms may be upheld by closet deviants who engage in enforcement so as to hide their deviance. But various empirical accounts indicate that audiences are often quite sensitive to this ulterior motive. Our theory and experimental evidence identify when inferences of ulterior motive are drawn and clarify the implications of such inferences. Our main test pivots on two contextual factors: (1) the extent to which individuals might try to strategically feign commitment and (2) the contrast between “mandated” enforcement, where individuals are asked for their opinions of deviance, and “entrepreneurial” enforcement, where enforcement requires initiative to interrupt the flow of social interaction. When the context is one where individuals might have a strategic motive and enforcement requires entrepreneurial initiative, suspicions are aroused because the enforcers could have remained silent and enjoyed plausible deniability that they had witnessed the deviance or recognized its significance. Given that the mandate for enforcement might be rare, a key implication is that norms might frequently be underenforced.

**Keywords:** social norms; rational choice; social psychology; deviance; enforcement; signaling

‘Doesn’t your position as a delegate to the Political Consultative Conference give you some protection?’ I asked my friend.

‘I hear the Maoists want to abolish that organization. *They call it a collection of radishes, red on the outside but white inside.* They claim that while all the delegates talked as if they support the Communist Party, in actual fact they oppose the party,’ she said. (Cheng 1986:73–74, emphasis added)

ONE of the foundational ideas of sociology is that norms are a primary source of social action and order (Coleman 1990; Homans 1950; Parsons 1937). But an enduring question is why members of a group might expend effort to ensure that others conform. Rational choice theorists have observed that “enforcement” (i.e., sanctioning norm violators and/or encouraging others to conform) is a public good and that the rational strategy is to let others bear the cost of providing this public good (Head 1974; Oliver 1980; Olson 1965:2; Rowley and Peacock 1975; Samuelson 1954). But more recent research suggests that individual actors might gain private utility (i.e., “selective incentives”; Olson 1965) from enforcement (Adut 2004, 2012; Becker 1963; Kuran 1995; Sunstein 1996:910–14). In particular, Willer, Kuwabara, and Macy (2009; cf. Centola, Willer, and Macy [2005] and Jordan et al. [2016]) argue that audiences regard an enforcer as more sincere in her compliance than those who do not enforce even though the enforcer is actually more likely to have privately deviated from the norm and that the “illusion of sincerity” available

**Citation:** Kim, Minjae, and Ezra W. Zuckerman Sivan. 2017. “Faking It Is Hard to Do: Entrepreneurial Norm Enforcement and Suspicions of Deviance” *Sociological Science* 4: 580-610.


**Received:** July 17, 2017

**Accepted:** September 13, 2017

**Published:** October 25, 2017

**Editor(s):** Jesper Sørensen, Mario Small

**DOI:** 10.15195/v4.a24

**Copyright:** © 2017 The Author(s). This open-access article has been published under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited. 

from enforcement explains why individuals are often motivated to enforce. The larger implications of this argument are twofold. First, an enforcer will appear more sincere in her compliance than a “bystander” who is not incurring a personal cost to enforce. Second, norms will be enforced as long as groups have insecure members who have reason to portray their conformity as sincere.

A puzzle emerges from a consideration of this social logic, however. Insofar as there might be ulterior motives behind norm enforcement, and insofar as audiences may themselves engage in enforcement to mask deviance in other social situations, it is odd that audiences do not recognize possibilities for such ulterior motives and discount for them. In fact, various empirical accounts suggest that rather than enhancing a perception of sincerity, enforcement often does raise audience suspicion that the enforcer is masking deviance. Consider research on purges such as the Chinese Cultural Revolution (Walder 2006, 2009; also, see Goode and Ben-Yehuda [1994]), whereby patterns of cross-accusation of deviance between different factions are documented. As suggested by the epigraphical quotation, such enforcement often raised suspicions that avowals of fealty to Mao were driven by the ulterior motive of wanting to protect themselves from allegations of “revisionism.” Similarly, when contemporary firms engage in social activism to fend off suspicions of unethical practices raised by social movement activists (Ingram, Yue, and Rao 2010; McDonnell and King 2013), enforcement often elicits doubts about their commitment and true motives (Yoon, Gurhan-Canli, and Schwarz 2006; cf. Carlos and Lewis 2017). The implication of these accounts then runs counter to the one summarized above: norms might in fact be underenforced because individuals are wise to such risk of enforcement. So which is it? Why and when might enforcement elicit suspicion of deviance while it at other times successfully creates an illusion of sincerity?

We ground our approach to this question in an analysis of a general problem faced by social observers—that of inferring whether an interactant’s motives are sincere or instead she is driven by an ulterior motive. The “sincere motive” inference often seems plausible because observers know that the enforcer is incurring a personal cost to help maintain the norm (Becker 1960; Schelling 1980). But when audiences recognize that there might be possible strategic reasons to feign commitment—in situations we label “high accusability”—they are wise to the possibility that an individual might try to avoid accusations of deviance by engaging in norm enforcement. We argue that to solve the inferential contest between the “sincere motive” and “ulterior motive” that arises in such contexts, audiences look for situational cues. In particular, social situations vary in the degree to which actors are charged with responding to normative deviance. At one end of the continuum, they might be mandated to assess one another’s commitment to norms (e.g., Jordan et al. 2016; Willer et al. 2009). Someone who is given such a mandate cannot claim to be unaware of the deviance—that is, she does not enjoy “plausible deniability.” As such, to fail to enforce the norm is to undermine it. By contrast, when there is no mandate and thus enforcement requires an interruption of social interaction—that is, one must take “entrepreneurial” initiative (Becker 1963; Sunstein 1996)—the alternative to enforcement is to do nothing and remain a bystander. Such a bystander enjoys plausible deniability of the deviance in that she can credibly claim that she

either did not observe the deviance or did not appreciate its significance. As such, it is natural to question why someone would step out of the bystander role and go out of her way to punish someone: the possibility that she is masking deviance thus emerges as highly salient.

After developing our theoretical framework, we test it using a series of online vignette experiments. In the experiments, subjects read about a fictitious group of college students who express their opinions on alcohol use on a college campus. In these studies, we demonstrate that audiences perceive an enforcer to be sincere when there is a mandate to respond to deviance, but audience suspicions are aroused when (1) audiences see possible strategic motives for appearing committed (i.e., in a situation of high accusability) and (2) an enforcer takes entrepreneurial initiative to enforce. We conclude by noting several implications of our theory and findings. Most notably, whereas recent research sees the pursuit of reputational benefits as straightforwardly stimulating prosocial behavior (e.g., Jordan et al. 2016; Willer 2009; see Simpson and Willer [2015] for review), our analysis indicates that such benefits (and thus the prosocial behavior) may be limited or enabled by common situational cues and audience perceptions of strategic motive. Our analysis also implies that potentially destabilizing undercurrents of suspicion might lurk even in many social settings where norms are superficially defended.

## Theory

### *Deepening the Puzzle*

Willer et al. (2009) offer two possible explanations for why enforcement creates an illusion of sincerity despite rational suspicions that the enforcer might be trying to hide her own deviance.<sup>1</sup> One idea comes from research that suggests that human beings often err by projecting their own motivations onto others (Miller and McFarland 1991; Prentice and Miller 1993). The other possible explanation is “correspondence bias,” a cognitive heuristic by which actors assume that others’ public behaviors accurately reflect their private beliefs (Gilbert and Malone 1995; Jones and Harris 1967). These explanations essentially suggest that cognitive difficulties prevent audience members from taking the enforcer’s perspective even though the shoe is often on the other foot.<sup>2</sup>

The experimental evidence presented by Willer et al. (2009) is consistent with this suggestion. In their experiment, subjects were informed of a situation where three actors (1) were asked to privately evaluate a purportedly high-quality but actually nonsensical text, (2) had a public discussion in which one turned out to be a deviant and two turned out to be conformists, and (3) were asked to publicly evaluate their fellow group members. Subjects were then told of one (randomly selected) conformist’s evaluation of fellow group members, either in a condition in which the conformist enforces the norm by criticizing the deviant or in a condition in which the conformist does not enforce the norm by evaluating all fellow group members equally. Subjects were more likely to think that the enforcer was more sincere in his compliance than the nonenforcer, which suggests that enforcement can occur without eliciting suspicions of ulterior motives.<sup>3</sup> The larger implication is

that many norms might be enforced even if there is significant private dissent (see Study 1 and 2 in Willer et al. [2009]).

But there is reason to question how enforcers can appear more sincere in their compliance with the norm than do bystanders (i.e., those who do not enforce). Put differently, even if audiences face cognitive challenges in recognizing ulterior motives in others, it seems that these challenges are often overcome. Accordingly, numerous empirical accounts document instances where (1) sincerity in compliance with the norm is considered important in maintaining one's status in the group (Willer 2009) and (2) norm enforcers are viewed with suspicion that they themselves broke the norm. Notable examples include the following cases.

*The Chinese Cultural Revolution* (see especially Walder [2009]). In this well-known political purge, Mao declared that there were hidden "revisionists" and demanded self-regulation. In the tumult that followed, high school and university students in Beijing (then spreading to the rest of the country) began accusing party officials and one another of being revisionists. Such accusations fueled a general environment of suspicion (Walder 2006).

*Socially responsible activities by corporations.* Firms in the contemporary market often engage in socially responsible activities through which they encourage others to comply with a norm, but such activities often elicit suspicion that they are engaging in those activities in order to cover up their own deviant deeds (Yoon et al. 2006; cf. Carlos and Lewis 2017).<sup>4</sup>

*Homophobia.* Individuals who publicly ridicule and attack gay individuals often elicit suspicion that they are overcompensating for their insecurity in their masculinity (cf. Humphreys 1970; Willer et al. 2013). The very term "homophobia" implies that bias against lesbian, gay, bisexual, and transgender (LGBT) individuals is often driven by the attacker's insecurities and fears.<sup>5</sup>

These examples suggest that audience members often view norm enforcers with more suspicions of deviance than they view bystanders who do not enforce.<sup>6</sup> Our challenge then is to develop a theory that addresses why and when acts of norm enforcement elicit suspicions of deviance in some cases, whereas in other situations, audiences deem enforcers to be sincere in their compliance.

### *Theoretical Framework*

Our theory focuses on the two possible motives that might be inferred from any act of enforcement (cf. Hahl and Zuckerman 2014).<sup>7</sup> One possible inference is that the enforcer is highly committed to the norm and is motivated to protect it: the fact that the enforcer is apparently willing to bear a personal cost in order to contribute to the public good, bolsters the impression of commitment. We will call this the "sincere motive" inference. In the instance of the Chinese Cultural Revolution, one who accuses another of being a revisionist might elicit less suspicion that she is a closeted revisionist because she is willing to speak up and bear the cost of having to navigate the social situation. But the other possible inference that audiences can draw is that enforcement is motivated by the desire to fabricate the impression of commitment to the norm. The key assumption here is that audience members are aware that the actor has an incentive to bolster the impression that she is sincere in

her conformity. Given the general awareness of the possibility of such a strategy, an enforcer might seem like a closeted deviant who has an ulterior motive to mask deviance. We will call this the “ulterior motive” inference. Lastly, note that even if audiences suspect an ulterior motive, this does not mean they are certain of the enforcer’s lack of commitment. This distinction will become important when we discuss the implications of suspicion elicited by enforcement.

### *(1) Situations of High Accusability versus Situations of Low Accusability*

Let us now clarify the contextual conditions under which audiences are most likely to face a contest between sincere and ulterior motive inferences. The most general condition is one in which any participant in the situation has a reasonable fear of being credibly accused—what we term as a situation of “high accusability.” Such a situation has three basic features: (a) deviance is known to have occurred to the point that there is reason to fear that social support for the norm will unravel, (b) it is widely believed that there are additional deviants whose identities are unknown, and (c) credible accusations of closeted deviance are likely. Whenever condition (a) is in place, someone who sincerely believes in the norm has an incentive to sanction a public deviant and thereby bolster the norm. And if it is either not common knowledge that there are closeted deviants (condition b) or there is little basis for thinking that closeted deviance might face a credible accusation (condition c), there is no reason to suspect such a norm enforcer of having an ulterior motive because closeted deviants have no reason to fear accusation (cf. Horne 2001, 2007). But insofar as conditions (b) and (c) are both in place, thus creating what we call a situation of high accusability, the sincere motive inference competes with the ulterior motive inference in the minds of an audience.

The cases reviewed above illustrate common examples of high-accusability situations. For instance, contemporary firms are increasingly exposed to the threat of being branded as ethically compromised by social movement activists (e.g., King and Pearce 2010). Situations of moral panic (Goode and Ben-Yehuda 1994) and purges often feature high accusability as well. It is also important to highlight that in a situation of high accusability, the fear of being accused can spread to all actors, including truly sincere conformists: even if one is free of deviance, allegations might nevertheless stick if there is no definitive proof of innocence. This engenders two key features of high-accusability situations as contrasted with low-accusability situations: (i) enforcement is an attractive strategy for actors (cf. Centola et al. 2005:1031), and (ii) audiences face an inferential contest (because a closeted deviant might be dissimulating to avoid credible accusations).

### *(2) Entrepreneurial Enforcement versus Enforcement in Response to a Mandate*

Yet some evidence suggests that enforcement might still successfully deflect suspicion of closeted deviance even in high-accusability situations (Adut 2004; Jordan et al. 2016; Reilly 2016; Willer et al., 2009). Thus, although high accusability may

be a necessary condition for suspicion to be aroused, it is not sufficient. Hence, we identify a second condition (and another quite general context of enforcement) that is necessary for audience suspicion to be aroused (i.e., entrepreneurial enforcement). Situations (and roles within them) vary in the extent to which someone who did not enforce—a bystander—would enjoy plausible deniability with regard to the existence and/or significance of the deviance. At one end of the continuum are situations that mandate actors to express their views on the norm violation. The situations presented in existing experimental studies (see especially Simpson, Harrell, and Willer [2013] and Willer et al. [2009]; also, see Jordan et al. [2016] for work in evolutionary psychology) tend to induce enforcement via a mandate. In these situations, individuals are specifically asked what they think of others and thus effectively asked whether they are willing to uphold group norms or not. In such a context, to refrain from enforcing a norm or standard is effectively to endorse deviance or poor performance. As such, even someone who does not have an ulterior motive to enforce would have no choice but to enforce, lest she endorse the deviance. This implies that the audience will have no reason to infer anything but a sincere motive.

However, situations that mandate actors to express their views on the norm violation seem relatively rare. More typical cases of norm enforcement seem to be those in which the choice is either to stand by and let the flow of social interaction continue or to interrupt this flow and call out deviance. An enforcer in such a context is a “moral entrepreneur,” in Becker’s (1963) classic phrasing, or a “norm entrepreneur,” in the more recent coinage by Sunstein (1996; e.g., Adut 2004, 2005, 2008; Reilly 2016; cf. Fine 1996).<sup>8</sup> The question of motive is highly salient for such an entrepreneur in a way that it is not when there is a mandate. In particular, were one to remain a bystander, one could plausibly claim that one did not witness the deviance or did not recognize its significance. A bystander thus does not seem to endorse the deviance. Conversely, someone who does take action should elicit suspicion: what motivated her to sanction the deviant when she did not have to say anything? Accordingly, when an enforcer takes entrepreneurial initiative to publicly identify wrongdoing, the voluntary nature of her enforcement serves to raise suspicion that the enforcer is driven by the ulterior motive of trying to cover up her own deviance. But if she were to remain silent then she would be able to plausibly claim that she did not know about norm violation or that she did not think the norm was threatened. As such, entrepreneurial enforcement should raise suspicion under situations of high accusability. Thus, as depicted in Table 1, we predict that when the situation is both one of high accusability and there is no mandate for enforcement, an enforcer will attract suspicion; otherwise, the enforcer will seem sincere. This theoretical framework is summarized in the following proposition:

Proposition: Insofar as a social situation is one where (1) closeted deviants have a reasonable fear of being credibly accused and (2) norm enforcement occurs via entrepreneurial initiative on the part of enforcers, enforcement elicits doubts about the enforcer’s commitments. Conversely, when either of these conditions is absent, enforcement will signal commitment to the norm.



**Table 1:** Theoretical predictions of when norm enforcers seem sincere or suspicious.

Does enforcement occur in the presence of a mandate?	Mandate present	Entrepreneurial enforcement
Low accusability	<i>Sincere</i>	<i>Sincere</i>
High accusability	<i>Sincere ("Illusion of Sincerity")</i>	<i>Suspicious</i>

In sum, our theory explains why enforcement can project an illusion of sincerity (Willer et al. 2009; cf. Jordan et al. [2016]) but also why it more often elicits suspicion. In particular, whereas insecure actors might wish to hide behind the illusion of sincerity under situations of high accusability, this seems realistic only in (seemingly rare) cases where actors are given a mandate to enforce the norm but not where enforcement requires entrepreneurial action. Consequently, the suspicion bred by enforcement might generally result in underenforcement of norms, whereby actors refrain from enforcing lest they appear suspicious. We elaborate on this implication of our theory (and notable exceptions to this implication) in the General Discussion section below.

## Empirical Validation

We provide empirical validation for our proposition with online experiments. Subjects for each of our studies were recruited through Amazon Mechanical Turk and limited to those with American internet protocol addresses.<sup>9</sup> This online experimental setting is useful because we were looking for subjects whose characteristics reflected the general (American) population instead of an audience that has specific knowledge or skills (Berinsky, Huber, and Lenz 2012; Buhrmester, Kwang, and Gosling 2011). However, subjects may pay less attention when participating in online experiments compared to those who are in laboratory settings in universities. As has become standard, we addressed this challenge in two ways. First, we broke down our experiments into different steps. The idea here is for subjects to receive a small amount of information at each click instead of reading too much information at once. Second, we asked them a series of attention questions to track whether they were understanding our instructions correctly (Mason and Suri 2011).

### *Studies 1 and 2: Demonstrating Suspicions of Deviance Cast upon Norm Entrepreneurship*

Studies 1 and 2 are designed to test three predictions applied to situations in which there is a publicly endorsed norm (i.e., there is no direct information on whether it is privately endorsed or disapproved of; cf. Prentice and Miller 1993; Kim 2017; Willer et al. 2009). The predictions emerge from the two dimensions captured in our proposition and Table 1, reflecting the two contextual conditions that animate our theory: (1) whether enforcement is entrepreneurial or mandated and (2) whether the situation is one that is of high versus low accusability. The first prediction is

a baseline prediction that derives from study 3 of Willer et al. (2009), whereby the illusion of sincerity is created by a norm enforcer. Our theory includes this prediction as well, though it explicitly limits it to situations in which enforcement occurs in response to a mandate.<sup>10</sup> Thus, the objective of the first prediction is to see how someone who enforces in response to a mandate appears as compared to someone who does not enforce in response to a mandate.

Hypothesis 1: An enforcer responding to a mandate is more likely to be seen as sincerely endorsing the norm than someone who is given the same mandate but does not enforce.

By contrast, when high accusability and the absence of a mandate raise the salience of the ulterior motive inference, an entrepreneurial enforcer enjoys less plausible deniability that she did not know about or appreciate the norm violation. Entrepreneurial enforcement is thus suspect.

Hypothesis 2: In a situation where accusability is high, an enforcer taking entrepreneurial initiative is less likely to be seen as sincerely endorsing the norm than a bystander.

Finally, we test whether a situation of high accusability (and not just entrepreneurial initiative) is a necessary condition to make the ulterior motive inference salient.

Hypothesis 3: Entrepreneurial enforcers are more likely to be viewed as sincerely endorsing the norm in a situation of low accusability than in a situation of high accusability.

In order to test these hypotheses, we designed two studies. In study 1, we simulated a situation of high accusability and tested hypotheses 1 and 2. In study 2, we compared enforcement in a situation of high accusability to that in a situation of low accusability and thereby tested hypothesis 3. In addition to confirming that the suspicions aroused by entrepreneurial enforcement indeed derive from inferences of strategic motive (hypothesis 3), study 2 plays another important role by testing whether interrupting the flow of social interaction in and of itself raises suspicion of deviance. Our focus in these studies was on distinguishing between norm entrepreneurship, at one end of the continuum, and enforcement in response to a mandate, at the other end. One might argue that an entrepreneurial enforcer is suspected of deviance not because audiences inferred an ulterior motive for enforcement but because audiences suspect anyone who interrupts the flow of social interaction in a situation when normative deviance is known to have occurred. But insofar as entrepreneurial enforcement does not raise suspicion in situations of low accusability (study 2), this implies that interruption in situations of deviance does not necessarily raise significant suspicion.

## Study 1 Design and Procedures

### *Recruitment and Conditions*

Five hundred and nine United States-based subjects<sup>11</sup> were recruited via Amazon Mechanical Turk in November and December of 2015. Each subject was paid \$0.75



for participating. After consenting to participate in the experiment, subjects were told that they would participate in a “social perception task” that was designed to “assess [their] ability to accurately infer the responses of individuals from a group of college freshmen.” Subjects were further told that “[t]his group of freshmen recently participated in an orientation session at their college and submitted their responses to an anonymous questionnaire.”

Subjects were randomly assigned to one of four sets of high-accusability conditions—“*High-Accusability/Mandated and Enforce*,” “*High-Accusability/Mandated and Not Enforce*,” “*High-Accusability/Entrepreneurial*,” and “*High-Accusability/Baseline*.” Subjects, therefore, were not aware of what was happening in other conditions (or, for that matter, that there were other conditions). The level of accusability indicates the extent to which the threat of accusation looms over actors in the situation. The *Mandated* conditions (*Mandated and Enforce* and *Mandated and Not Enforce*) presented a setup whereby the fictive group members were specifically asked to assess a public deviant (but we told subjects only about the assessment of the first group member who was asked to assess the public deviant). In the *Not Mandated* conditions (i.e., *Entrepreneurial* conditions and *Baseline* condition), no such mandate was given, and enforcement required initiative. In the *Entrepreneurial* conditions, we told subjects about the assessment by the (one and only) entrepreneurial enforcer on the public deviant. And in the *Baseline* condition, neither was the mandate given nor did the enforcement take place. These conditions are summarized in Table 2.

Finally, note that each of the conditions besides the *Baseline* condition has three subconditions based on the specific manner by which enforcers invoke or do not invoke normative principles. This is more exploratory in nature but is important because we have no baseline for knowing how various enforcement activities are perceived, at least in this experimental paradigm. We elaborate on this issue below.

### *Study 1 Vignette*

We first lay out the *Baseline* condition, in which neither mandate nor enforcement takes place. After doing so, we describe how other conditions in study 1 are different.

*High-Accusability/Baseline condition.* After the introduction, subjects were told the following story in four consecutive screens. First, they learned that (1) in the prior month, a new cohort of college freshmen arrived at a college campus in the United States and that (2) six of these students who “had met each other only briefly before” and were about to spend the entire year in the same dorm (labeled “students 1 through 6”) had participated in an orientation session. Subjects were then given details about that orientation session and were told that they would be asked to evaluate the students’ behavior. In particular, subjects learned that the six students had participated in a “Bloomsbury Ethics Roundtable,” an academic tradition in the philosophy of ethics in which participants express their opinions on an ethical question, and that this process had two stages—first, writing their opinions in a private, anonymous questionnaire, and second, having a public discussion.<sup>12</sup> Subjects learned that the students were informed about these stages in advance. In particular, the six students were asked to fill out an anonymous questionnaire

**Table 2:** Experimental conditions of studies 1 and 2.

Study 1		
High-accusability situation	Mandate to respond	No mandate to respond
Enforcer	<i>High-Accusability/ Mandated and Enforce</i> (N = 149)	<i>High-Accusability/ Entrepreneurial</i> (N = 150)
Bystander (i.e., nonenforcer)	<i>High-Accusability/ Mandated and Not Enforce</i> (N = 160)	<i>High-Accusability/ Baseline</i> (N = 50)
Study 2		
Low-accusability situation	Mandate to respond	No mandate to respond
Enforcer	<i>Low-Accusability/ Mandated and Enforce</i> (N = 165)	<i>Low-Accusability/ Entrepreneurial</i> (N = 168)
Bystander (i.e., nonenforcer)	<i>Low-Accusability/ Mandated and Not Enforce</i> (N = 183)	<i>Low-Accusability/ Baseline</i> (N = 58)

Note: Because the *Baseline* conditions do not have subconditions for different wording, the *Baseline* conditions have about one-third of the subjects of other conditions.

on “whether or not the college should ban alcohol on its campus.” Subjects also learned that after all the students had filled out their anonymous questionnaires, student 2 asked the administrator “why [the administrator] asked [the students] to fill out those questionnaires” when they “have to give their opinions publicly anyway.” The administrator replied by saying that “some students [in previous years] were influenced by other students when they debated the issue publicly” and that the administrator is “interested in whether the discussion influences students’ opinions.” At this point, subjects were asked some attention-checking questions.<sup>13</sup>

Afterward, subjects learned about stage 2 in three consecutive screens. First, subjects learned that students 2, 4, 1, 3, and 6 had been called upon (in allegedly random order) to express their opinions, and each of them said that the college should not ban alcohol. Next, subjects learned that before the administrator could call the last student (student 5), “an accident” occurred. In particular, the administrator who was collecting the anonymous questionnaires dropped student 5’s response, and everyone saw student 5’s response that said the college should ban alcohol on its campus. Subjects also learned that after the accident, student 5 publicly expressed an opinion that was in line with his private, accidentally revealed response. In particular, subjects learned that student 5 had said that “underage students may feel pressure to drink” if alcohol was allowed on campus, and he added that “drinking is harmful for the development of students’ mind and body.” In this fashion, the first of the three conditions for high accusability was introduced:

student 5 was revealed publicly as a deviant. At that point, subjects were asked some attention-checking questions.

In the last set of screens, subjects were told the remainder of the story. First, they learned something that did not vary by condition. In particular, they learned that after the discussion, the administrator had announced the results from the initial anonymous questionnaire and noted that two students had initially indicated that the college should ban alcohol. This introduces the second element of a high-accusability situation, in which it is widely known or suspected that there are closeted deviants present. At this point, we added the third element for a high-accusability situation by creating expectations of credible allegations of closeted deviance. We did this by having student 5 say the following in a low voice: "Hmm. . . if anyone is interested, I'm quite sure I know who initially agreed with me and then wouldn't apparently admit it! That person and I have talked about this stuff before." In addition to the obvious threat of accusation expressed in the comment, we made the threat of accusation more credible by having the known deviant (i.e., student 5) raise the possibility of accusation. Afterward, subjects in the *Baseline* condition were told that the orientation ended here.

*High-Accusability/Mandated and Enforce conditions.* Next, to test our theory, our conditions varied in the degree to which a mandate was provided to assess the deviance. In all *Mandated and Enforce* conditions, subjects were initially told that the "Bloomsbury Ethics Roundtable" format entailed one additional stage in which all students would be called upon in random order to evaluate one another. Accordingly, after student 5 hinted that he knew who the other deviant was (which is when the story ended for subjects in the *Baseline* condition), subjects were told that student 6 was randomly called first to evaluate other students' opinions and that he said the opinions of students 1, 2, 3, and 4 were "very valid," whereas student 5's opinion was "not valid at all" (see below for more detail on wording). The story ended here for subjects in the *Mandated and Enforce* conditions (i.e., subjects were not told how other students responded to the mandate that was presumably given), and subjects were then asked questions about how they perceived student 6 (see more on the questions below).

*High-Accusability/Mandated and Not Enforce conditions.* Subjects in the *Mandated and Not Enforce* conditions were also told that students were asked to evaluate one another in random order and that student 6 was called first. In these conditions, subjects were told that student 6 said the opinions of students 1, 2, 3, 4, and 5 are all "equally valid." Just as in the *Mandated and Enforce* conditions, the story ended here for subjects in the *Mandated and Not Enforce* conditions, and as in the *Mandated and Enforce* conditions, subjects were then asked questions about how they perceived student 6.

*High-Accusability/Entrepreneurial conditions.* By contrast, subjects in the *Entrepreneurial* conditions were not told anything about an additional stage. Instead, after the administrator announced the results from the anonymous questionnaire and student 5 hinted that he knew who the other deviant was (when the story ended for subjects in the *Baseline* condition), subjects were told that "something happened." Subjects were further told that student 6 "suddenly spoke up in a loud voice" and said student 5's opinion was "not valid at all." Afterward, subjects

were told the same thing as those in the *Baseline* condition—that the orientation ended and the students were dismissed. The story ended here for the *Entrepreneurial* conditions, and subjects in these conditions were then asked questions about how they perceived all students but student 5 (i.e., students 1 through 4 and 6).

### *Different Wordings of the Four Conditions*

The wording used to enforce a norm may influence what motive audiences attribute to an enforcer. Our main intuition was that it may make a difference whether the enforcer simply expresses disapproval for the deviant or whether the enforcer invokes a normative principle. The logic is that it is odd and potentially suspicious for a norm entrepreneur to not justify his or her own enforcement. Therefore, for all sets of conditions other than the *Baseline* condition, we designed three subconditions with different wordings. These wordings reflect varying degrees of justification for enforcement and are drafted so that each wording subcondition has parallel subconditions. Overall, subjects were randomly assigned to one of the 10 conditions (3 [*High-Accusability/Mandated and Enforce*, *High-Accusability/Mandated and Not Enforce*, and *High-Accusability/Entrepreneurial*]  $\times$  3 [*Simple*, *Stating Principle*, and *Activist*] + 1 [*High-Accusability/Baseline*]). Table 1 in the online supplement presents enforcement wordings for each condition. We do not state formal hypotheses on how the wording of enforcement matters. Our intuition, however, was that the Simple wordings of enforcement are the least common and plausible, whereas the Activist wordings may be the most plausible and common in norm entrepreneurship.

### *Dependent Variable*

Before reading the story, subjects in all conditions had not been told that they would be asked to answer about any specific students. As said before, subjects had only been told that the task is to “assess [their] ability to accurately infer the responses of individuals from a group of college freshmen.” After they finished reading the story, subjects in all *Mandated* conditions were asked the following question: “How likely is it that student 6 wrote the following statements on his anonymous questionnaire?” Subjects were asked to evaluate the statement, “student 6 wrote that ‘the college should ban alcohol on its campus’” on scales ranging from 1 (very unlikely) to 10 (very likely). The higher the rating subjects gave to this question the more they doubted student 6’s sincerity in his compliance. Because subjects in the *Mandated* conditions were not told of the other students’ (i.e., students 1 through 5) evaluations of one another (that presumably happened), subjects were not asked about their guesses on the other students’ anonymous questionnaires. That is, subjects in the *Mandated* conditions were explicitly told before answering the questions that they would be asked to only guess student 6’s private response and not any other students’. Following the logic explicated below, testing hypothesis 1 involved comparing subjects’ responses on student 6 within the *Mandated* conditions (i.e., the *Mandated and Enforce* conditions versus *Mandated and Not Enforce* conditions).

Subjects in the *Entrepreneurial* and *Baseline* conditions were asked to answer the same question but for all five other students (i.e., excluding student 5) because there was no difference in how the administrator had treated those students. Subjects

were asked to answer about all five other students in the same screen. Asking about all students in these conditions is especially important because (as explicated below) there are two comparisons that are covered by hypothesis 2: (1) the between-subject comparison (i.e., between student 6 in the *Entrepreneurial* condition and student 6 in the *Baseline* condition) and (2) the within-subject comparison, (i.e., between the enforcer [i.e., student 6] and the bystanders [i.e., students 1, 2, 3, and 4] within the same *Entrepreneurial* conditions). Responses to these questions were used as direct measures of how committed each student appeared because the questions asked subjects about how likely it appeared that the students in question changed their responses. After subjects answered additional questions, including free-response items and demographic items, they were thanked and given the code to be paid.<sup>14</sup> Finally, because subjects were asked about all five students in both the *Entrepreneurial* and the *Baseline* conditions, we are less concerned that subjects' responses in the *Entrepreneurial* conditions reflect a desire to spot the insincere student any more than subjects' responses in the *Baseline* condition.

### *Hypothesis Testing*

Our theoretical framework suggests that subjects draw inferences about others' motives by taking into account a particular feature of social context—that is, whether an explicit mandate for norm enforcement is present or absent. We thus tested hypothesis 1 by comparing the perceived sincerity of enforcers versus nonenforcers within *Mandated* conditions, and we tested hypothesis 2 by comparing the perceived sincerity of enforcers versus nonenforcers within *Not Mandated* conditions. The results of these tests are, therefore, down each column in Table 2 and never diagonal or within each row. Note that the very contextual feature we vary also changes how enforcers may be compared with nonenforcers. In particular, when there is no mandate, all individuals have exactly the same opportunity to engage in norm enforcement; as such, any entrepreneurial enforcer's sincerity can be meaningfully compared to all others who had the same opportunity to interrupt the flow of social interaction and enforce but did not do so. Therefore, for hypothesis 2, we expect that subjects will perceive the enforcer in the *Entrepreneurial* conditions (i.e., student 6 in the *Entrepreneurial* conditions) to be less sincere than bystanders in any condition that lacks a mandate: either the bystanders in the *Entrepreneurial* conditions (i.e., students 1 through 4 in the *Entrepreneurial* conditions) or the bystanders in the *Baseline* condition (e.g., student 6 in the *Baseline* condition).

By contrast, insofar as mandates are separately given to each would-be enforcer, their opportunities to enforce are not equivalent. In particular, any responses to a mandate that presumably followed student 6's response to the mandate (i.e., responses to the mandate by the students 1 through 5) might be interpreted as influenced by student 6's response to the mandate. As such, the appropriate way to compare enforcers and nonenforcers in the *Mandated* conditions is through a between-subject test, whereby the sincerity of the first person in a situation who was given a mandate and who did enforce (i.e., student 6 in the *Mandated and Enforce* conditions) is compared to the sincerity of a person who was also given the first mandate in an equivalent situation but did not enforce in response to that mandate

(i.e., student 6 in the *Mandated and Not Enforce* conditions). Tests for hypothesis 1, therefore, involve only between-subject comparisons.

## Study 1 Results

### *Validating Hypothesis 1*

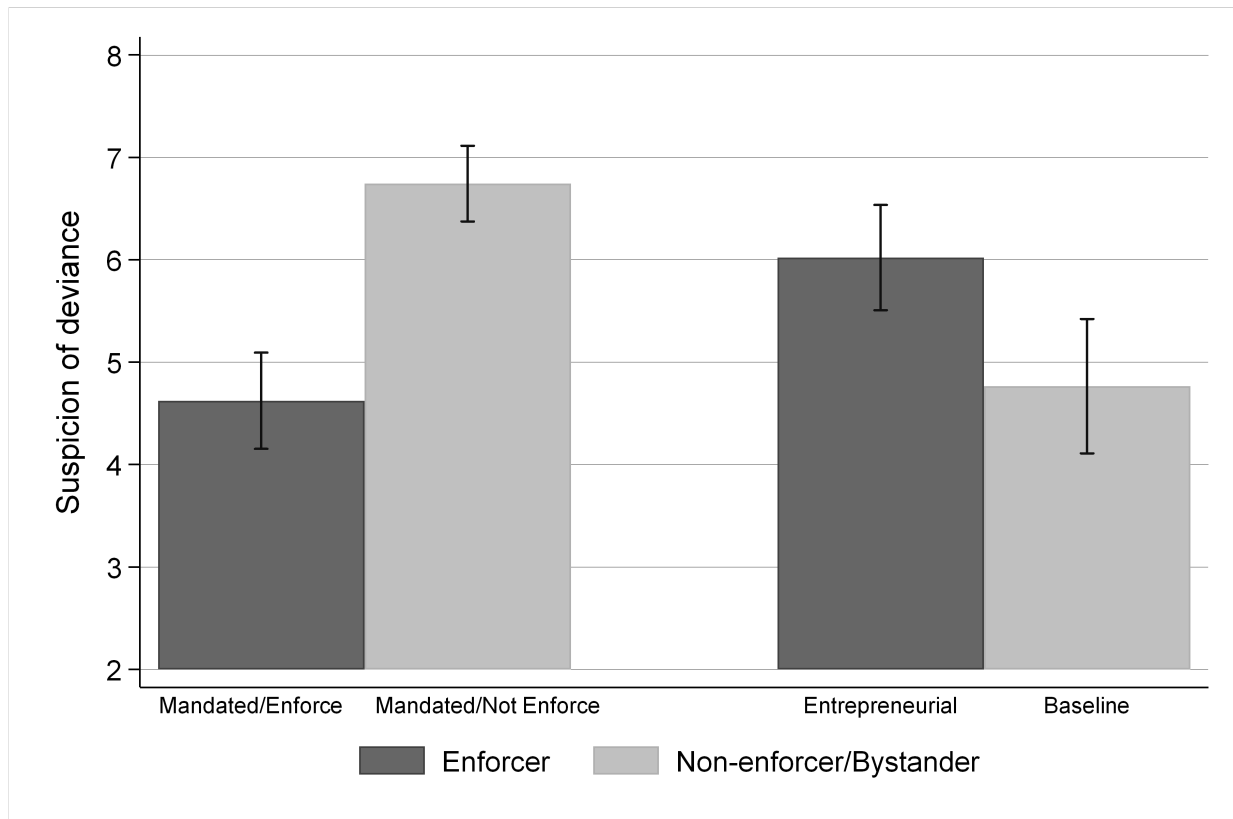
We argued that the illusion of sincerity would be generated when enforcement is in response to a mandate because nonenforcers would lack plausible deniability that they had witnessed or recognized the significance of the deviance. This is reflected in hypothesis 1. In order to validate this prediction, we developed a design that should replicate results from Willer et al. (2009), which showed a higher perception of sincerity for the enforcer than for someone who was also given a mandate but did not enforce. Our results successfully replicated the findings of Willer et al. (2009). Figure 1 shows that student 6 in the *High-Accusability/Mandated and Enforce* conditions appeared less likely to have changed his response ( $M = 4.62$ ) than did student 6 in the *High-Accusability/Mandated and Not Enforce* condition ( $M = 6.74$ ;  $t = 5.87^{15}$ ;  $p < 0.001$ ; degrees of freedom [df] = 307).<sup>16</sup> In addition, an analysis of variance (ANOVA) indicates that suspicion was elicited mainly by the refusal to enforce and not by a particular wording subcondition ( $F[1,301] = 35.33$ ;  $p < 0.001$ ; see Table 2 in the online supplement).

### *Validating Hypothesis 2*

Having replicated the finding (from Willer et al. 2009) that the illusion of sincerity can be elicited by those who enforce in response to a mandate, we then tested whether audience suspicion is aroused when enforcers lack such a mandate. As the main test of our argument, hypothesis 2 predicted that because an entrepreneurial enforcer could have ignored the deviance and enjoyed plausible deniability, he would face greater suspicion of deviance than would bystanders. The results validate this prediction. Student 6 in the *High-Accusability/Entrepreneurial* conditions was perceived as more likely to have changed his response than students 1, 2, 3, and 4 on average ( $M = 6.02$  vs.  $M = 3.83$ ;  $t = 5.86$ ;  $p < 0.001$ ;  $df = 298$ ). Results are substantively the same in each *High-Accusability/Entrepreneurial* wording subcondition. By contrast, there is no evidence that student 6 in the *High-Accusability/Baseline* condition seemed any more suspicious than students 1, 2, 3, and 4 on average ( $M = 4.77$  vs.  $M = 4.16$ ;  $t = 1.28$ ;  $p = 0.20$ ;  $df = 98$ ), indicating that it is entrepreneurial enforcement (and not other characteristics of student 6) that made him appear more suspicious. Figure 2 shows these results visually.

The between-condition comparison (i.e., student 6 in the *High-Accusability/Entrepreneurial* conditions vs. student 6 in the *High-Accusability/Baseline* condition) further demonstrates that entrepreneurial enforcement invites suspicions of deviance. This comparison is a particularly conservative test because the experiment was set up so that student 6 in the *High-Accusability/Baseline* condition (as well as in other conditions) went last in the public discussion. Insofar as subjects were told there was a closeted deviant and asked to report their suspicions of deviance, student 6 in the



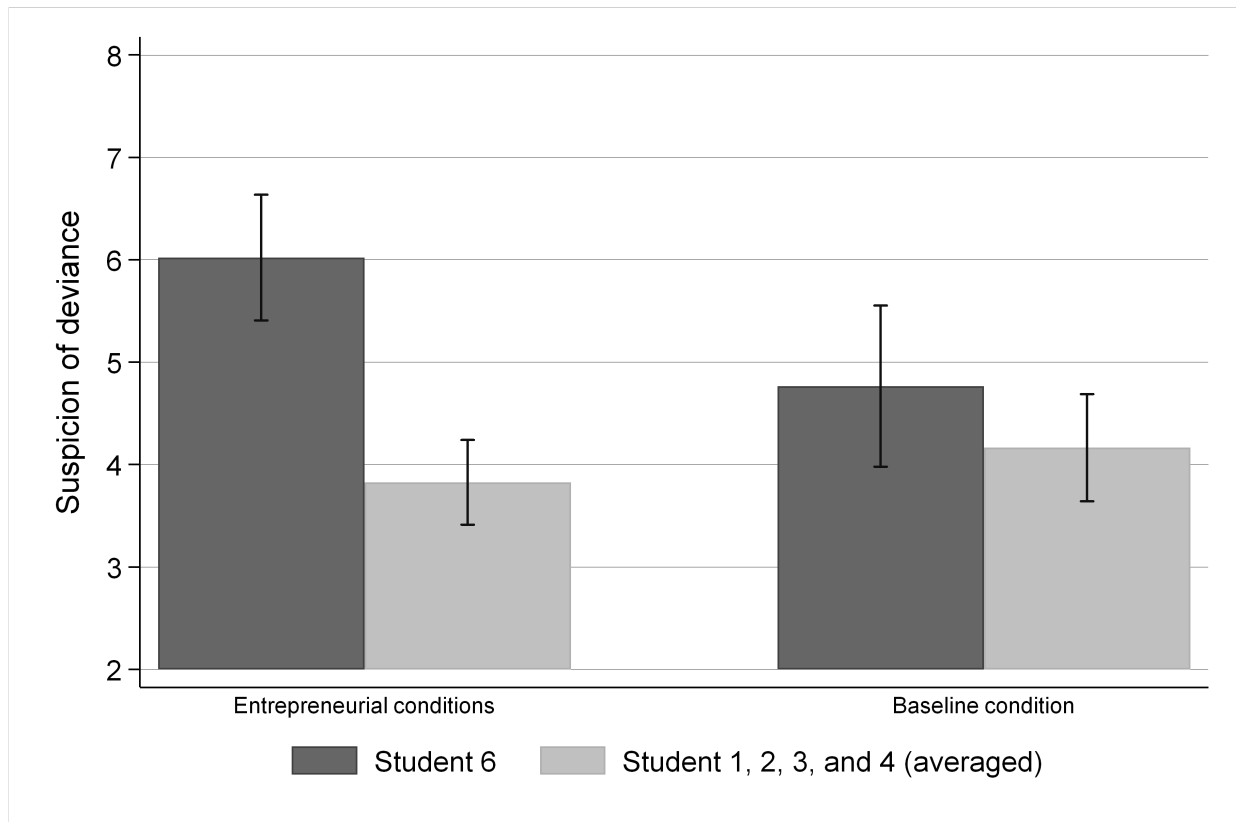


**Figure 1:** Between-subject comparison of perceived likelihood that student 6 changed his response in study 1. The error bars indicate the 95 percent confidence interval. Possible values of the  $y$  axis range from 1 (very unlikely) to 10 (very likely). The measure is the response to the question, “How likely is it that student X wrote the following statements on his anonymous questionnaire?: student X wrote that ‘the college should ban alcohol on its campus.’”

*High-Accusability/Baseline* condition might suffer from higher suspicion of deviance to begin with, especially because subjects are implicitly trying to pinpoint one closeted deviant. Nevertheless, when student 6 remained silent like those before him, he elicited significantly less suspicion than in the condition in which he engaged in entrepreneurial enforcement (M for student 6 in all *High-Accusability/Entrepreneurial* conditions = 6.02 vs. M for student 6 in the *High-Accusability/Baseline* condition = 4.77;  $t = 2.51$ ;  $p = 0.01$ ;  $df = 198$ ). The ANOVA results further confirm that entrepreneurially enforcing elicits suspicion ( $F[1,183] = 4.23$ ,  $p < 0.05$ ; see Table 3 in the online supplement) and that different wording subconditions matter little. Again, Figure 1 visually shows these results.

## Study 1 Discussion

In study 1, we validated hypothesis 1 by replicating the illusion of sincerity effect (Willer et al. 2009), showing that when enforcement occurred in response to a



**Figure 2:** Within-subject comparisons of likelihood that the student in question changed his response in the *High-Accusability/Entrepreneurial* conditions and the *High-Accusability/Baseline* condition in study 1. The error bars indicate the 95 percent confidence interval. Possible values of the *y* axis range from 1 (very unlikely) to 10 (very likely). Separate measures for students 1, 2, 3, and 4 were averaged in all the *High-Accusability/Entrepreneurial* conditions and the *High-Accusability/Baseline* condition (Cronbach's alpha = 0.96). The measure is the response to the following question: "How likely is it that student X wrote the following statements on his anonymous questionnaire?: Student X wrote that 'the college should ban alcohol on its campus.'"

mandate; in such situations, the ulterior motive inference is less salient even in a situation of high accusability. But supporting hypothesis 2, study 1 also shows that suspicions are significantly aroused when enforcement occurs via entrepreneurial initiative. In the absence of a mandate, plausible deniability that one did not see or appreciate the deviance is more available to a bystander; as such, entrepreneurial enforcement is suspicious.

It is worth clarifying that although results from both the within-condition comparison and between-condition comparison validate hypothesis 2, the more useful comparison is likely the within-condition comparison. We limited the comparison of the *High-Accusability/Mandated* conditions to be between student 6 in each of the two different conditions in order to control for factors that are less applicable to the real world (i.e., the unlikely random order in which one is mandated to enforce).<sup>17</sup> Yet counterfactual scenarios, by definition, do not exist in the real world. Instead,

one might more often be rewarded or punished on the basis of one's appearance of commitment in comparison to others' in the same situation. On the basis of such reasoning, we regard the within-condition effect shown in Figure 2 from the *Entrepreneurial* conditions as the strongest evidence in support of hypothesis 2. Consequently, the problem with entrepreneurial enforcement under high accusability is that one invites suspicion of oneself in comparison with others in the same situation.

However, study 1 alone cannot test whether high accusability is a necessary situational feature for enforcement to elicit audience suspicion. Insofar as our theory hinges on the assumption that audiences perceive ulterior motives of covering deviance from entrepreneurial enforcement, we need to directly test that suspicion would be attenuated when there is little reason to defend oneself. Such a test is especially needed because one might argue that the audience suspicion, elicited by entrepreneurial enforcement in study 1, can be aroused by any interruptions and not necessarily by acts suspected as deriving from ulterior motives. That is, insofar as subjects could have a desire to spot the student most likely to be insincere, any attention generated by the interruption (including entrepreneurial enforcement) might have elicited suspicion. Study 2, therefore, is designed to test the idea that even when subjects might have the same level of desire to spot the student most likely to be insincere, entrepreneurial enforcement does not elicit suspicion in the absence of accusability. Therefore, we now move to study 2, in which we introduced conditions that parallel those of study 1 but in low-accusability situations. This allowed us to test hypothesis 3, which holds that entrepreneurial enforcement is more likely to create the impression that the enforcer endorses the norm in low-accusability situations (as implemented in study 2) than in high-accusability situations (conditions from study 1).

## Study 2 Design and Procedures

### *Recruitment and Conditions*

Five hundred seventy-four subjects were recruited via Amazon Mechanical Turk in the same time period as study 1 and paid the same as those in study 1. Study 1 and study 2 were conducted within days of each other, and participants in study 1 were excluded, making them two comparable samples—essentially two sets of subconditions of the same experiment. Because the primary purpose of adding study 2 was to reduce accusability and, therefore, the strategic value behind enforcement, conditions in study 2 are parallel to study 1 but only in low-accusability situations. Therefore, the four sets of conditions of study 2 are as follows: *Low-Accusability/Mandated and Enforce*, *Low-Accusability/Mandated and Not Enforce*, *Low-Accusability/Entrepreneurial*, and *Low-Accusability/Baseline*. Again, subjects were assigned to only one of these conditions, and they were not aware of the presence of other conditions; the full set of conditions is again summarized in Table 2. Enforcement wordings for each condition remain exactly the same as in study 1.

## Study 2 Vignette

Because the only difference between the study 1 and study 2 conditions is the level of accusability, the conditions of study 2 are different from their corresponding conditions of study 1 in only two subtle ways. First, whereas subjects in study 1 were told that students “had met each other only briefly before,” subjects in study 2 were told that students “had never met” one another before the orientation. This change subtly informs subjects in study 2 that students could not make educated guesses about one another’s true commitment on the basis of their experiences, thereby lowering the threat of credible accusation. Second, subjects no longer heard one of the students (i.e., student 5) claim that he knew who privately supported banning alcohol (i.e., who does not endorse the norm). To recall, subjects in all conditions of study 1 were told the following: after the public discussion in which all students but student 5 expressed their opposition to banning alcohol, the students were informed that there was another student who had initially favored banning alcohol on campus; then, student 5 said he was quite sure who was the closeted deviant, thereby raising the threat of accusation. By contrast, subjects in study 2 were not told of any comment after the public announcement and instead assumed that the students proceeded right to the next stage. In other words, right after the public announcement about a closeted deviant by the administrator, subjects in the *Low-Accusability/Mandated* conditions were told that student 6 was called in random order to evaluate the other students. After student 6 evaluated the other students, subjects were asked the same questions as ones subjects in the *High-Accusability/Mandated* conditions of study 1 were asked. In the *Low-Accusability/Entrepreneurial* conditions, subjects were told that student 6 suddenly spoke up to condemn student 5 (i.e., the known deviant), and subjects were also told that the orientation ended there. Subjects in the *Low-Accusability/Baseline* condition were told that the orientation ended right after the administrator made the announcement. Subjects in each condition of study 2 were asked the exact same questions as subjects in the corresponding condition of study 1.

## Study 2 Results

### *Validating Hypothesis 3*

The main objective of study 2 was to test whether an entrepreneurial enforcer in a situation of low accusability (i.e., student 6 in *Low-Accusability/Entrepreneurial* conditions in study 2) would appear less suspicious than the equivalent one in a situation of high accusability (i.e., student 6 in *High-Accusability/Entrepreneurial* conditions in study 1). Our results support this prediction. That is, suspicion is considerably greater for the entrepreneurial enforcer in high-accusability situations as compared to the same enforcer in low-accusability situations ( $M$  for student 6 in all study 1 *High-Accusability/Entrepreneurial* conditions = 6.02 vs.  $M$  for student 6 in all study 2 *Low-Accusability/Entrepreneurial* conditions = 4.49;  $t = 3.66$ ;  $p < 0.001$ ;  $df = 316$ ). We also used ANOVA to identify the effect of the main manipulation, and results confirm the effect of accusability ( $F[1,302] = 13.49$ ;  $p < 0.001$ ; see Table 4 in the online supplement).

In addition, in the *Low-Accusability/Entrepreneurial* conditions, there is no evidence that audiences suspected any problematic commitment from student 6, who took entrepreneurial initiative to enforce the norm ( $M = 4.49$ ) compared to students 1, 2, 3, and 4 on average ( $M = 4.26$ ;  $t = 0.72$ ;  $p = 0.48$ ;  $df = 334$ ). It is also important to note that the additional suspicion directed toward student 6 (compared to the suspicion directed toward students 1, 2, 3, and 4) in the *Low-Accusability/Baseline* condition ( $M$  for student 6 = 5.47 vs.  $M$  for students 1, 2, 3, and 4 = 4.73;  $t = 1.63$ ;  $p = 0.11$ ;  $df = 114$ ) was as large as or larger than the additional suspicion directed toward student 6 in the *Low-Accusability/Entrepreneurial* conditions, indicating that entrepreneurial enforcement in low-accusability situations did little to elicit additional suspicion. Figure 3 shows these results visually.

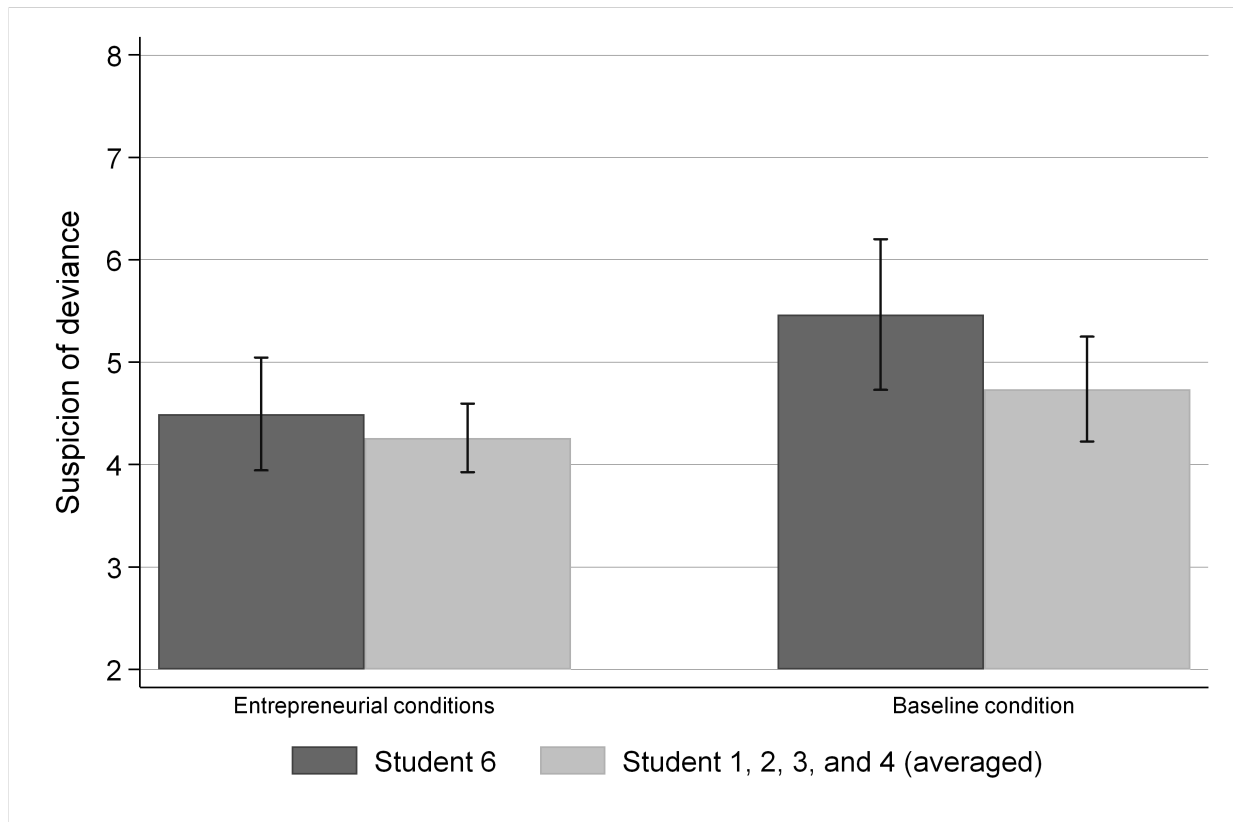
The between-condition comparisons (i.e., student 6 in the *Low-Accusability/Entrepreneurial* conditions vs. student 6 in the *Low-Accusability/Baseline* condition) further validate that entrepreneurial enforcement did not elicit further doubts about the enforcer's commitments in a situation of low accusability. Figure 4 shows this comparison visually. In fact, if student 6 enforced entrepreneurially and did so in what appeared to be a plausible manner, he aroused significantly less suspicion than student 6 in the *Baseline* condition. In particular, student 6 in the *Low-Accusability/Entrepreneurial–Stating Principle* subcondition appeared significantly less suspicious ( $M = 4.16$ ) than student 6 in the *Baseline* condition ( $M = 5.47$ ;  $t = 2.20$ ;  $p < 0.05$ ;  $df = 226$ ). Student 6 in the *Low-Accusability/Entrepreneurial–Activist* subcondition also appeared significantly less suspicious ( $M = 3.84$  vs.  $M = 5.47$ ;  $t = 2.73$ ;  $p < 0.01$ ;  $df = 228$ ). The ANOVA test further confirmed these results ( $F[1,222] = 6.61$ ;  $p = 0.01$ ; see Table 5 in the online supplement<sup>18</sup>).

### *Replicating Willer et al. (2009) in a Situation of Low Accusability*

Lastly, it is useful to confirm that enforcement in response to a mandate in a situation of low accusability can still fend off potential suspicion that the entrepreneurial enforcement in a situation of high accusability elicits, as depicted in Table 1. Study 2 does this. Subjects in general thought that student 6 in the *Low-Accusability/Mandated and Enforce* conditions was less likely to have changed his response ( $M = 4.23$ ) than student 6 in the *Low-Accusability/Mandated and Not Enforce* conditions ( $M = 6.32$ ;  $t = 6.26$ ;  $p < 0.001$ ;  $df = 346$ ; see Figure 4). The ANOVA that compared the two mandated conditions in a situations of low accusability further validated that the effect is significant at the conventional level ( $F[1, 342] = 36.33$ ;  $p < 0.001$ ; see Table 6 in the online supplement).

## Study 2 Discussion

Study 2 tests hypothesis 3 and demonstrates that suspicions that are elicited by entrepreneurial enforcement in the general situation of high accusability can be “turned off” by lowering the level of accusability. Qualitative responses support this part of our theory as well. After asking questions for our main dependent variable and the suspicion score, we asked subjects “why [they] evaluated student 6 the way [they] did.” Their answers suggest that they recognize entrepreneurial enforcement

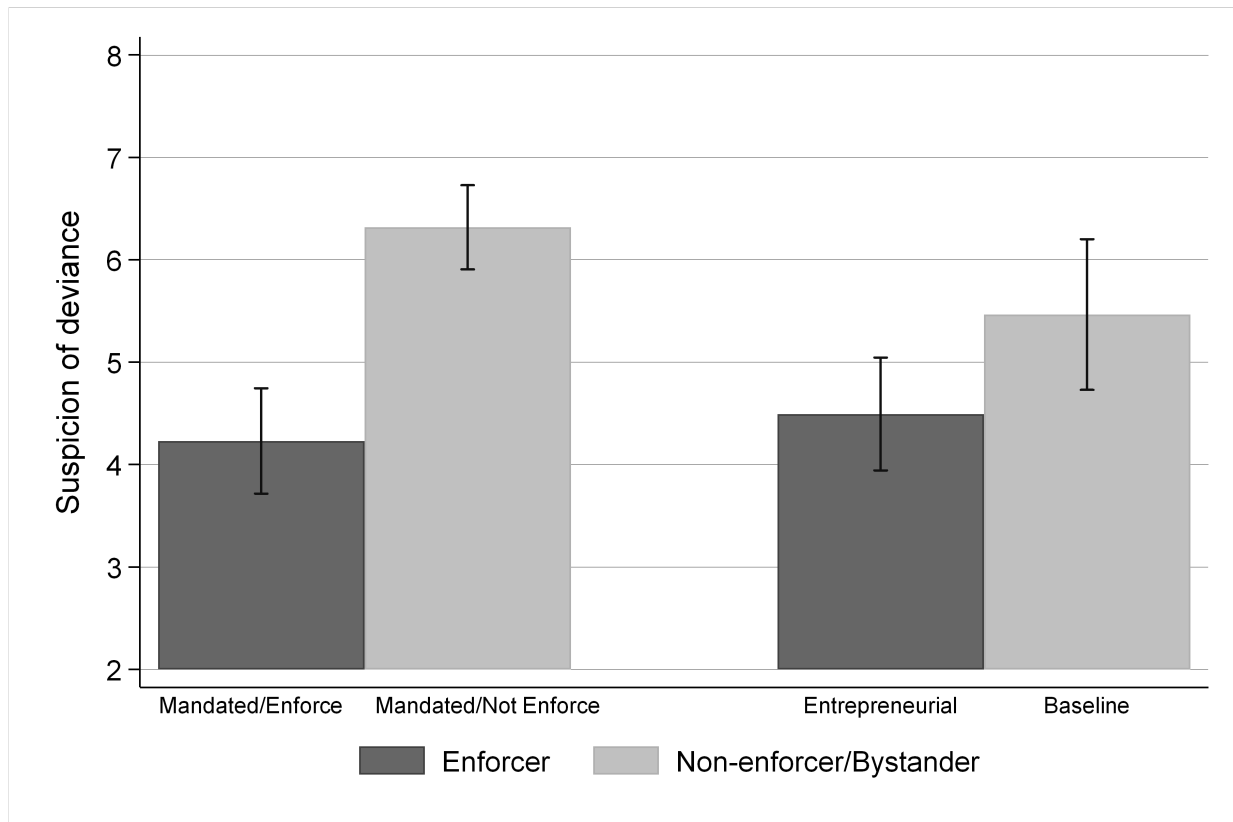


**Figure 3:** Within-subject comparisons of likelihood that the student in question changed his response in the *Low-Accusability/Entrepreneurial* conditions and the *Low-Accusability/Baseline* condition in study 2. The error bars indicate the 95 percent confidence interval. Possible values of the  $y$  axis range from 1 (very unlikely) to 10 (very likely). Separate measures for students 1, 2, 3, and 4 were averaged in all the *Low-Accusability/Entrepreneurial* conditions and the *Low-Accusability/Baseline* condition (Cronbach's  $\alpha = 0.88$ ). The measure is the response to the following question: "How likely is it that student X wrote the following statements on his anonymous questionnaire?: Student X wrote that 'the college should ban alcohol on its campus.'"

has no particular strategic value in a low-accusability situation: as one subject put it, "[student 6] seemed very passionate about his opinion. There was no need to speak out against student 5 unless it really bothered him."

Another important implication of study 2 is that simply interrupting the flow of social interaction does not raise the suspicion of deviance. In fact, the between-condition results showed that when one interrupted the flow of social interaction and enforced in a plausible manner in a situation of low accusability (i.e., in the *Low-Accusability/Entrepreneurial–Stating Principle* and the *Low-Accusability/Entrepreneurial–Activist* conditions), one appeared less suspicious than someone who remained silent. These results then demonstrate that the suspicion elicited by entrepreneurial enforcement in *High-Accusability* conditions in study 1 was not simply an experimental byproduct of study 1 subjects' desires to spot the insincere student—both high accusability and entrepreneurial enforcement are needed to elicit suspicion.





**Figure 4:** Between-subject comparison of perceived likelihood that student 6 changed his response in study 2. The error bars indicate the 95 percent confidence interval. Possible values of the  $y$  axis range from 1 (very unlikely) to 10 (very likely). The measure is the response to the question “How likely is it that Student X wrote the following statements on his anonymous questionnaire?: Student X wrote that ‘the college should ban alcohol on its campus.’”

Yet although these results from a situation of low accusability show that enforcement through interruption alone is not sufficient to raise suspicion, they do not necessarily mean that all interruptions are the same. In fact, we designed a posttest in which student 5 was in a situation of high accusability and said, “You know, I really like this Bloomsbury Roundtable—this is exactly the kind of experience I wanted to have in college. I’m so glad that I got to be part of it, and I’m looking forward to the rest of the orientation.” The objective of this test was to see whether mere interruption using an orthogonal statement might raise suspicion as much as the interruption via enforcement. Results show that entrepreneurial norm enforcement indeed elicited less suspicion than did this irrelevant comment. This suggests that although entrepreneurial enforcement in a situation of high accusability does make the ulterior motive inference more salient, it still keeps the sincere motive inference viable by at least claiming sincerity, whereas someone who interrupts the flow of social interaction without any defensive acts (e.g., enforcement) might not keep afloat even that much of the sincere motive inference. Results from study 2

and the posttest then highlight the situational features that enable enforcement to generate the illusion of sincerity.

## Limitations

Before we discuss and draw implications from our theory and results, it is worth clarifying an important limitation of our study design. Our vignette experiments entailed informing subjects about a hypothetical scenario in which normative deviance and enforcement occurred; we then asked subjects to judge the enforcer's (and bystanders') sincerity. Although results from these vignette studies demonstrated when the enforcer appears suspicious of having an ulterior motive, our subjects were not themselves in the position to enforce (or remain silent). Such a setup is difficult to implement in online experiments in which subjects are anonymous (i.e., when subjects do not have reasons to care about how sincere their conformity appears). Without such a design, however, we do not yet know if would-be enforcers act in the manner implied by the pattern of audience response about which we have theorized and provided experimental results. Therefore, we focused on addressing why the perception of suspicion is elicited by enforcement (and not necessarily why norms are underenforced), although a possible implication of our theory is that norms will be underenforced.

At the same time, the experimental paradigm we developed does provide important indirect evidence as to why norms might be underenforced when entrepreneurial initiative is required to enforce that norm. As Willer et al. (2009:477) put it, "[i]f enforcement is widely regarded as a telltale sign of personal insecurity, social anxiety, and conformity, then it is unlikely that people will enforce to prove to others that they are true believers." Sociologists have indeed long documented how perceptions by others drive individuals' behaviors (e.g., Bourdieu 1984; Goffman 1959), and our theory and empirical results similarly identify conditions under which suspicions aroused by entrepreneurial norm enforcement might lead to the underenforcement of norms. We discuss this issue further in the General Discussion section below.

It is also worth emphasizing the weaknesses and strengths of our sampling strategy. Given that our studies feature the issue of drinking on college campuses, undergraduate participants might be more knowledgeable about the vignette setting. However, we were also concerned that their close involvement in campus life might influence their responses and constrain us from controlling our treatments. Insofar as our sample is familiar (either through experience or social learning) with the setting presented in our vignette and is more representative of the general population than college students in lab settings (Berinsky et al. 2012; Buhrmester et al. 2011), conducting vignette experiments on this online platform seems particularly well suited to test our theory while maximizing our control over our treatments of interest (Parigi, Santana, and Cook 2017).

Finally, a limitation of our sampling strategy is that the samples were limited to the United States-based subjects. We do not have theoretical reasons for entrepreneurial enforcement to arouse suspicion only among the U.S.-based subjects, yet we are limited by substantive (e.g., knowledge around a setting in which such

contestation of a norm is common) and practical (e.g., tools with which to collect samples that are as general as those of U.S. equivalents on the Amazon Mechanical Turk platform) constraints. Empirical accounts documented from political purges (e.g., the Chinese Cultural Revolution) and lay beliefs from different cultures (e.g., sayings analogous to “Whoever smelt it dealt it”) suggest that this theoretical mechanism is prevalent even though we do not validate this theory with samples other than the U.S.-based ones in this article.

## General Discussion

Recent research has suggested how one might use “false enforcement” (Willer et al. 2009) to mask hidden deviance and that this is an important basis for upholding norms. However, various empirical accounts and lay beliefs suggest that audiences are often suspicious of the ulterior motives that this strategy implies. The questions that emerge from this tension then are why and when enforcement can create an illusion of sincerity and why and when it elicits suspicions of ulterior motives.

The main contribution of this article is to identify two situational features that explain variation in suspicion that a norm enforcer is a hidden deviant: (1) whether actors face a significant threat of facing credible accusations of deviance and thus pressure to feign commitment and (2) whether the context is such that actors have no mandate to engage in enforcement but must take entrepreneurial initiative to do so. Our experiments replicated past results showing that deviance is successfully masked when actors are given a mandate to engage in enforcement. But our results also support our argument that suspicion is highest in the more common case when actors have no mandate and precisely when false enforcement is attractive as a strategy to mask hidden deviance (i.e., under high accusability). Again, the general logic is that someone who takes entrepreneurial initiative to sanction deviance could have ignored the deviance without having endorsed it; this invites the suspicion that she has a special motive to signal her commitment to the norm. By contrast, someone who fails to condemn a norm violation when given a mandate to do so effectively endorses the deviance and thus contributes to the undermining of the norm.

We now put the contribution of our article in a broader context by considering our theory’s implications for how norm enforcement might evolve over time in light of situations documented in previous studies. We then conclude by highlighting the tension between rewards and punishments faced by actors who undertake prosocial action.

### *Accounting for Persistence of Norm Entrepreneurship and Its Implications*

Our theory and empirical tests pointed to conditions under which norm enforcers can escape suspicions of ulterior motive even in a situation of high accusability (i.e., when assessment is mandated) as documented by Willer et al. (2009). Although our conjecture is that such instances are rare, it is nevertheless important to note empirical cases that meet this condition reaffirming the literature’s and our evidence’s

external validity. In particular, Adut's (2004) analysis of the French investigating magistrates' anticorruption campaign is a case of enforcement in response to a mandate. Because the job of the investigating magistrates was to legally investigate anyone who was suspected of illegality (corruption, more specifically, in Adut's case), they could escape suspicions of ulterior motives for their enforcement, similar to the enforcer in our *Mandated and Enforce* conditions. Accordingly, Adut (2004:547) observed that the French investigating magistrates in the 1990s were able to seize their "opportunities for public displays of moral rectitude and courage" and enhance their status by accusing then-high-status political actors of being corrupt. Their institutional role gave them a mandate, which helped the sincere motive inference become more viable (also see Erikson 1966; Evans-Pritchard 1976; Reilly 2016).<sup>19</sup> Yet the case of French magistrates seems to be a rather special case. It is difficult to come up with other examples in which enforcers are so well protected from suspicions of ulterior motive by a mandate, as even enforcers occupying such institutional roles are rarely given a mandate to respond to specific norm violations. It seems much more common for enforcement to require some degree of entrepreneurial initiative, as is highlighted by Becker's (1963) classic insight, and as shown in our results, suspicions may be aroused if an enforcer is viewed as straying beyond the mandate.

It is also worth noting that enforcement is most attractive as a strategy for signaling sincere conformity in situations of high accusability, and therefore, it is most pertinent to investigate what inference audiences draw in such situations. Insofar as the individual motivation for enforcement is about signaling sincere conformity, enforcement has little value in situations of low accusability: why incur the personal cost and enforce when there is little reason to portray one's conformity as sincere (see also Centola et al. 2005:1031)?<sup>20</sup> Consequently, investigating the perception of enforcers is most appropriate in situations of high accusability where enforcement is most attractive as a strategy, and our results demonstrate that the most common form of norm enforcement (i.e., entrepreneurial enforcement) elicits the suspicion of deviance in such situations. The larger implication of our theory is that insofar as the enforcement of a norm requires entrepreneurial initiative, the norm may be underenforced—despite its potential value as a commitment-enhancing strategy—and ultimately fail to govern social action.

At the same time, enforcement does not necessarily cease under such conditions. After all, the empirical examples we discussed earlier—e.g., purges such as the Cultural Revolution as well as moral panics (Goode and Ben Yehuda 1994) and "the outbreaks of enforcement" that mark many scandals (Adut 2008)—involve widespread enforcement even though such enforcement occasions suspicion.<sup>21</sup> But if the context is one of high accusability and enforcement requires entrepreneurial initiative, why might enforcement still continue? Addressing this question is a fruitful avenue for future research: although this article's implication is that norms will be underenforced among strategic actors behaving rationally, with full information on the costs of enforcement, such restrictive assumptions of the intent, rationality, and knowledge of the actors might not apply in certain situations.<sup>22</sup>

## Conclusion

Previous research has suggested that contributions to a public good can be driven by selective reputational incentives, and contributions, such as prosocial behaviors, are more likely to occur because of them (e.g., Willer 2009). However, more recent research suggests that an overt pursuit of the reputational benefits may lower the perception of commitment (e.g., Hahl and Zuckerman 2014; Simpson and Willer 2008), and this article identified conditions under which a contributor to a public good (i.e., an enforcer) can plausibly deny suspicions of an ulterior motive. This article thus suggests that situational features may be able to heighten the perception that prosocially oriented actors are seeking reputational or material benefits rather than the public good, and this in turn might make prosocial behavior less likely even in the presence of selective incentives for prosocial behavior. Future research should investigate (1) how such suspicions of ulterior motives affect different prosocial behaviors and (2) why some prosocial behaviors persist despite suspicions of ulterior motives. By addressing these questions, we will be able to more fully account for conditions under which prosocial behaviors persist and conditions under which they do not.

## Notes

- 1 Willer et al. (2009) focus on the case of a norm from which individuals privately dissent (an “unpopular norm”). But insofar as the norm is perceived to be privately endorsed by individuals who publicly conform (which is the case in Willer et al. 2009), audiences have the same inferential problem as they would have with enforcement of a norm that is actually privately endorsed by members of the group because they, in either case, think that the norm (whether popular or unpopular) is privately endorsed by other individuals (Prentice and Miller 1993; cf. Kim 2017). We discuss this issue further in the Methodological Appendix in the online supplement.
- 2 If it is situationally impossible for the enforcer to be the deviant (e.g., victims of sexual assault disputes in which victims cannot be the perpetrator of the same sexual assault case), enforcement may not attract suspicions of deviance. Even in such cases, however, the enforcer may encounter a reputational backlash (1) because they appear to profit from public enforcement or (2) because the publicity of their enforcement is distasteful (Adut 2008).
- 3 See the Methodological Appendix in the online supplement for more details.
- 4 Such terms as “greenwashing” (encouraging the adoption of environment-friendly policies and activities to cover their own environmental misdeeds) and “pinkwashing” (advertising and encouraging the adoption of LGBT-friendly policies to cover their own misdeeds) reflect such suspicion.
- 5 Perhaps the more illustrative examples are various sayings in many cultures that imply a recognition of insecurity that stimulates norm enforcement. One of many English examples is, “whoever smelt it dealt it.” Similar sayings abound in other cultures: A Chinese example is, “one who yells out ‘catch the thief’ is the thief”; a Korean version is, “the dog that scolds another dog for being dirty is actually dirtier”; and a Hebrew version is, “all who accuse others of a disqualification suffer from that very disqualification.”

- 6 At the same time, these accounts document that enforcement occurred despite the fact that it breeds suspicion (cf. McDonnell 2015). We further discuss this point in the General Discussion section.
- 7 Research on prosocial behavior and costly punishment in evolutionary psychology and behavioral economics is broadly relevant here. The question in those lines of work is whether actors are perceived as prosocial when they actively sanction others who fail to engage in prosocial behavior (e.g., if they fail to allocate money to others in a dictator game). Such sanctioning can be motivated either by sincere commitment to the group or by vengeance against the norm violator (Brañas-Garza et al. 2014; Fehr and Gächter 2002; Henrich et al. 2001; Raihani and Bshary 2015; cf. Jordan et al. 2016). A key implication is that enforcers are more likely to be perceived as committed when they are not the primary victim of deviance and thus less likely to be motivated by vengeance (Feldman Hall et al. 2014; Raihani and Bshary 2015:101). This implication is consistent with the approach taken here in that the key question is under what situational conditions the interpretation of ulterior motive is salient. However, this line of work does not speak directly to the contexts we discuss in our article because this work focuses on cases for which there is no debate about what is and is not prosocial behavior (cf. Herrmann, Thöni, and Gächter 2008). Therefore, there is no way for these evolutionary models to account for the possibility of (hidden) deviance from a norm motivated by principled objection to that norm on the basis of the endorsement of an alternative norm. Our framework encompasses such a possibility.
- 8 Sunstein (1996) suggests that norm entrepreneurs are those who enforce new norms and thus are often “interested in changing social norms.” However, because many norms are rarely enforced (and thus are “fragile”), norm entrepreneurship can be more broadly defined as the enforcement of norms that are rarely enforced or not explicitly established (see his discussion on p. 909). As elaborated on previously, our empirical tests examine the effect of the enforcement of such a fragile norm.
- 9 We limited the subject pool to Amazon Mechanical Turk participants who are located in the United States, experienced in the platform (i.e., completed more than 100 online surveys in the platform before), and reliable (i.e., had their previous surveys approved—as opposed to unapproved for omission or inattention—by surveyors more than 90 percent of the time).
- 10 The Methodological Appendix in the online supplement details how our studies compare with study 3 of Willer et al. (2009), who suggested that enforcement always elicits a higher perception of commitment than that of a bystander.
- 11 In all studies in this article, all subjects who completed the experiment were included in the analysis. Responses to attention checks (seven or eight attention questions total) show that subjects understood the vignette generally well (more than 90 percent of the subjects answered all attention questions correctly). Analysis without subjects who failed one or more attention checks provides substantively the same results.
- 12 In reality, no such discussion format exists. We adapted this format from Study 2 in Willer et al. (2009), in which subjects are in the shoes of one of the students participating in the discussion.
- 13 A sample attention-checking question is, “[w]hich of the following best describes the students in the orientation session?” with the correct answer being, “A group of six incoming freshmen participated in the orientation session. They had met each other only briefly before, and they were about to spend the entire year together in the same dorm.”
- 14 No demographic factors except for gender affected subjects’ doubts about student 6’s commitment: men were in general more suspicious than women that student 6 changed



his opinion. However, this effect did not substantively change the main effects of our manipulations, so we do not discuss this further.

- 15 Although the t-test is efficient and easy to interpret, it also assumes equal variance between two populations of comparison. For a robustness check, we also use the Mann–Whitney test that does not assume equal variance, and all results remain substantively the same (Fay and Proschan 2010; Mann and Whitney 1947).
- 16 All statistical tests in this article were done in R. All ANOVA tests in this article are type II unbalanced ANOVA because each wording subcondition does not always have an equal number of subjects. Unbalanced ANOVA was run by using the “car” package.
- 17 We further discuss this point in the Methodological Appendix in the online supplement.
- 18 The significant interaction effect from Table 5 in the online supplement suggests that the main effect of entrepreneurial enforcement is significantly weaker for the Entrepreneurial–Simple subcondition. This is in line with our intuition that entrepreneurially enforcing in an implausible manner not invoking a normative principle does not enhance one’s perception of commitment.
- 19 Erikson’s analysis of the Salem witch trials (1966; also, Evans–Pritchard [1976]) also documents how adolescent girls who were already identified as “witch doctors” were regarded as reliable when they identified other girls as witches. The designated role provided them with a mandate to assess others’ professed sincerity in upholding the norm (of not being a witch), and they could escape suspicions of ulterior motives for hiding their own deviance and protect their status in the community. See also Reilly (2016:25–8) for an intriguing contrast between entrepreneurial enforcement and enforcement in response to a mandate in a community of stand-up comedians. He describes how Joe Rogan—who entrepreneurially accused Carlos Mencia of joke theft—was often criticized for his “opportunistic and self-promotional grandstanding,” whereas Marc Maron—who asked Mencia to address the issue on his interview podcast—was able to escape such criticism because he had little to gain and the topic was “seemingly unavoidable.”
- 20 Obviously, we are not barring exceptional cases in which individuals might be personally invested in the issue to incur a personal cost and enforce the norm regardless of public perception, although those individuals might still refrain from entrepreneurially enforcing lest they be regarded as opportunists.
- 21 Another example comes from the scandal surrounding Russia’s meddling in the 2016 U.S. presidential election, in which Donald J. Trump has made accusations of deviance in order to defend himself (Phillips 2017).
- 22 We thank the editor for this suggestion.

## References

- Adut, Ari. 2004. “Scandal as Norm Entrepreneurship Strategy: Corruption and the French Investigating Magistrates.” *Theory and Society* 33(5):529–78. <https://doi.org/10.1023/B:RYS0.0000045718.42431.7d>
- Adut, Ari. 2005. “A Theory of Scandals: Victorians, Homosexuality, and the Fall of Oscar Wilde.” *American Journal of Sociology* 111(1):213–48. <https://doi.org/10.1086/428816>
- Adut, Ari. 2008. *On Scandal: Moral Disturbances in Society, Politics, and Art*. Cambridge: Cambridge University Press.
- Adut, Ari. 2012. “A Theory of the Public Sphere.” *Sociological Theory* 30(4):238–62. <https://doi.org/10.1177/0735275112467012>

- Becker, Howard S. 1960. "Notes on the Concept of Commitment." *American Journal of Sociology* 66(1):32–40. <https://doi.org/10.1086/222820>
- Becker, Howard S. 1963. *Outsiders: Studies in The Sociology Of Deviance*. New York: Free Press.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–68. <https://doi.org/10.1093/pan/mpr057>
- Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press.
- Brañas-Garza, Pablo, Antonio M. Espín, Filippos Exadaktylos, and Benedikt Herrmann. 2014. "Fair and Unfair Punishers Coexist in the Ultimatum Game." *Scientific Reports* 4:6025.
- Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6(1):3–5. <https://doi.org/10.1177/1745691610393980>
- Carlos, W. Chad, and Ben W. Lewis. 2017. "Strategic Silence: Withholding Certification Status as a Hypocrisy Avoidance Tactic." *Administrative Science Quarterly* 1–40. <https://doi.org/10.1177/0001839217695089>
- Centola, Damon, Robb Willer, and Michael Macy. 2005. "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms." *American Journal of Sociology* 110(4):1009–40. <https://doi.org/10.1086/427321>
- Cheng, Nien. 1986. *Life and Death in Shanghai*. London: Grafton Books.
- Coleman, James. 1990. *Foundations of Social Theory*. Belnap Press.
- Erikson, Kai. 1966. *Wayward Puritans: A Study in the Sociology of Deviance*. Boston, MA: Prentice Hall.
- Evans-Pritchard, Edward E. 1976. *Witchcraft, Oracles and Magic among the Azande*. Oxford: Oxford University Press.
- Fay, Michael P., and Michael A. Proschan. 2010. "Wilcoxon-Mann-Whitney or T-Test? On Assumptions for Hypothesis Tests and Multiple Interpretations of Decision Rules." *Statistics Surveys* 4:1–39. <https://doi.org/10.1214/09-SS051>
- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415(6868):137–40. <https://doi.org/10.1038/415137a>
- Feldman Hall, Oriël, Peter Sokol-Hessner, Jay J. Van Bavel, and Elizabeth A. Phelps. 2014. "Fairness Violations Elicit Greater Punishment on Behalf of Another than for Oneself." *Nature Communications* 5(5306):5306.
- Fine, Gary A. 1996. "Reputational Entrepreneurs and the Memory of Incompetence: Melting Supporters, Partisan Warriors, and Images of President Harding." *American Journal of Sociology* 101(5):1159–93. <https://doi.org/10.1086/230820>
- Gilbert, Daniel T., and Patrick S. Malone. 1995. "The Correspondence Bias." *Psychological Bulletin* 117(1):21–38. <https://doi.org/10.1037/0033-2909.117.1.21>
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. Anchor Books.
- Goode, Erich, and Nachman Ben-Yehuda. 1994. "Moral Panics: Culture, Politics, and Social Construction." *Annual Review of Sociology* 20:149–71. <https://doi.org/10.1146/annurev.so.20.080194.001053>
- Hahl, Oliver, and Ezra W. Zuckerman. 2014. "The Denigration of Heroes? How the Status Attainment Process Shapes Attributions of Considerateness and Authenticity." *American Journal of Sociology* 120(2):504–54. <https://doi.org/10.1086/678304>
- Head, John G. 1974. *Public Goods and Public Welfare*. Durham, NC: Duke University Press.

- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *The American Economic Review* 91(2):73–78. <https://doi.org/10.1257/aer.91.2.73>
- Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Antisocial Punishment Across Societies." *Science* 319(5868):1362–67. <https://doi.org/10.1126/science.1153808>
- Homans, George C. 1950. *The Human Group*. Brace & Company, New York.
- Horne, Christine. 2001. "The Enforcement of Norms: Group Cohesion and Meta-Norms." *Social Psychology Quarterly* 64(3):253–66. <https://doi.org/10.2307/3090115>
- Horne, Christine. 2007. "Explaining Norm Enforcement." *Rationality and Society* 19:139–70. <https://doi.org/10.1177/1043463107077386>
- Humphreys, Laud. 1970. *Tearoom Trade: Impersonal Sex in Public Places*. Duckworth.
- Ingram, Paul, Lori Qingyuan Yue, and Hayagreeva Rao. 2010. "Trouble in Store: Probes, Protests, and Store Openings by Wal-Mart, 1998–2007." *American Journal of Sociology* 116(1):53–92. <https://doi.org/10.1086/653596>
- Jones, Edward E., and Victor A. Harris. 1967. "The Attribution of Attitudes." *Journal of Experimental Social Psychology* 3(1):1–24. [https://doi.org/10.1016/0022-1031\(67\)90034-0](https://doi.org/10.1016/0022-1031(67)90034-0)
- Jordan, Jillian J., Moshe Hoffman, Paul Bloom, and David G. Rand. 2016. "Third-Party Punishment as a Costly Signal of Trustworthiness." *Nature* 530(7591):473–76. <https://doi.org/10.1038/nature16981>
- Kim, Minjae. 2017. "A Man is Known by His Cup: Signaling Commitment via Costly Conformity." Working Paper, Massachusetts Institute of Technology.
- King, Brayden G., and Nicholas A. Pearce. 2010. "The Contentiousness of Markets: Politics, Social Movements, and Institutional Change in Markets." *Annual Review of Sociology* 36(1):249–67. <https://doi.org/10.1146/annurev.soc.012809.102606>
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Mann, Henry B., and Douglas R. Whitney. 1947. "On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other." *The Annals of Mathematical Statistics* 18(1):50–60. <https://doi.org/10.1214/aoms/1177730491>
- Mason, Winter, and Siddharth Suri. 2011. "Conducting Behavioral Research on Amazon's Mechanical Turk." *Behavior Research Methods* 44(1):1–23. <https://doi.org/10.3758/s13428-011-0124-6>
- McDonnell, Mary-Hunter, and Brayden King. 2013. "Keeping up Appearances: Reputational Threat and Impression Management after Social Movement Boycotts." *Administrative Science Quarterly* 58(3):387–419. <https://doi.org/10.1177/0001839213500032>
- McDonnell, Mary-Hunter. 2015. "Radical Repertoires: The Incidence and Impact of Corporate-Sponsored Social Activism." *Organization Science* 27(1):53–71. <https://doi.org/10.1287/orsc.2015.1017>
- Miller, Dale T., and Cathy McFarland. 1991. "When Social Comparison Goes Awry: The Case of Pluralistic Ignorance." Pp. 287–313 in *Social Comparison: Contemporary Theory and Research*, edited by J. Suls and T. A. Wills. Hillsdale, NJ: Erlbaum.
- Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 85(6):1356–75. <https://doi.org/10.1086/227168>

- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Parigi, Paolo, Jessica J. Santana, and Karen S. Cook. 2017. "Online Field Experiments: Studying Social Interactions in Context." *Social Psychology Quarterly* 80(1):1–19. <https://doi.org/10.1177/0190272516680842>
- Parsons, Talcott. 1937. *The Structure of Social Action*. The Free Press.
- Phillips, Amber J. 2017. "Trump's Defense against Comey Has Fallen into a Predictable Pattern: Make a Baseless Accusation." *Washington Post*. Retrieved June 12, 2017. <https://www.washingtonpost.com/news/the-fix/wp/2017/06/12/trumps-defense-against-comey-has-fallen-into-a-predictable-pattern-make-a-baseless-accusation/>.
- Prentice, Deborah A., and Dale T. Miller. 1993. "Pluralistic Ignorance and Alcohol Use on Campus: Some Consequences of Misperceiving the Social Norm." *Journal of Personality and Social Psychology* 64(2):243–56. <https://doi.org/10.1037/0022-3514.64.2.243>
- Raihani, Nichola J., and Redouan Bshary. 2015. "The Reputation of Punishers." *Trends in Ecology & Evolution* 30(2):98–103. <https://doi.org/10.1016/j.tree.2014.12.003>
- Reilly, Pat. 2016. "No Laughter Among Thieves: Informal Intellectual Property Rights and Status Dynamics in Stand-Up Comedy." Working Paper, University of California, Los Angeles.
- Rowley, Charles K., and Alan T. Peacock. 1975. *Welfare Economics: A Liberal Restatement*. First Edition. London: Martin Robertson.
- Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *The Review of Economics and Statistics* 36(4):387–9. <https://doi.org/10.2307/1925895>
- Schelling, Thomas. *The Strategy of Conflict*. 1980. Cambridge, MA: Harvard University Press.
- Simpson, Brent, and Robb Willer. 2008. "Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior." *Social Psychology Quarterly* 71(1):37–52. <https://doi.org/10.1177/019027250807100106>
- Simpson, Brent, Ashley Harrell, and Robb Willer. 2013. "Hidden Paths from Morality to Cooperation: Moral Judgments Promote Trust and Trustworthiness." *Social Forces* 91(4):1529–48. <https://doi.org/10.1093/sf/sot015>
- Simpson, Brent, and Robb Willer. 2015. "Beyond Altruism: Sociological Foundations of Cooperation and Prosocial Behavior." *Annual Review of Sociology* 41(1):43–63. <https://doi.org/10.1146/annurev-soc-073014-112242>
- Sunstein, Cass R. 1996. "Social Norms and Social Roles." *Columbia Law Review* 96(4):903–68. <https://doi.org/10.2307/1123430>
- Walder, Andrew G. 2006. "Ambiguity and Choice in Political Movements: The Origins of Beijing Red Guard Factionalism." *American Journal of Sociology* 112(3):710–50. <https://doi.org/10.1086/507854>
- Walder, Andrew G. 2009. *Fractured Rebellion: The Beijing Red Guard Movement*. Cambridge, MA; London: Harvard University Press.
- Willer, Robb. 2009. "Groups Reward Individual Sacrifice: The Status Solution to the Collective Action Problem." *American Sociological Review* 74(1):23–43. <https://doi.org/10.1177/000312240907400102>
- Willer, Robb, Ko Kuwabara, and Michael W. Macy. 2009. "The False Enforcement of Unpopular Norms." *American Journal of Sociology* 115(2):451–90. <https://doi.org/10.1086/599250>

- Willer, Robb, Christabel L. Rogalin, Bridget Conlon, and Michael T. Wojnowicz. 2013. "Overdoing Gender: A Test of the Masculine Overcompensation Thesis." *American Journal of Sociology* 118(4):980–1022. <https://doi.org/10.1086/668417>
- Yoon, Yeosun, Zeynep Gürhan-Canli, and Norbert Schwarz. 2006. "The Effect of Corporate Social Responsibility (CSR) Activities on Companies with Bad Reputations." *Journal of Consumer Psychology* 16(4):377–90. [https://doi.org/10.1207/s15327663jcp1604\\_9](https://doi.org/10.1207/s15327663jcp1604_9)

**Acknowledgements:** We thank Ari Adut, Hannah Birnbaum, Ronald Burt, Vanessa Conzon, Daniel DellaPosta, Roberto Fernandez, Jae-Kyung Ha, Oliver Hahl, Kate Kellogg, Minkyung Kim, Josh Krieger, Aruna Ranganathan, Dawn Robinson, and Robb Willer; audiences in the 2016 Annual Meeting of the American Sociological Association session on attitudes, norms, and behaviors; and the Economic Sociology Working Group and Behavioral Lab at the Massachusetts Institute of Technology Sloan School of Management for their helpful comments and feedback. All the usual disclaimers apply.

**Minjae Kim:** Sloan School of Management, Massachusetts Institute of Technology.  
E-mail: minjae@mit.edu.

**Ezra W. Zuckerman Sivan:** Sloan School of Management, Massachusetts Institute of Technology. E-mail: ewzucker@mit.edu.