

Community Standards of Deception: Deception Is Perceived to Be Ethical When It Prevents Unnecessary Harm

Emma E. Levine

Department of Behavioral Science, The University of Chicago Booth School of Business

We frequently claim that lying is wrong, despite modeling that it is often right. The present research sheds light on this tension by unearthing systematic cases in which people believe lying is ethical in everyday communication and by proposing and testing a theory to explain these cases. Using both inductive and experimental approaches, the present research finds that deception is perceived to be ethical and individuals want to be deceived when deception is perceived to prevent *unnecessary harm*. This research identifies eight community standards of deception: rules of deception that most people abide by and recognize once articulated, but have never previously been codified. These standards clarify systematic circumstances in which deception is perceived to prevent unnecessary harm, and therefore, circumstances in which deception is perceived to be ethical. This work also documents how perceptions of unnecessary harm influence the use and judgment of deception in everyday life, above and beyond other moral concerns. These findings provide insight into when and why people value honesty and paves the way for future research on when and why people embrace deception.

Keywords: deception, ethics, morality, communication

Supplemental materials: <https://doi.org/10.1037/xge0001081.supp>

Moral decency ensures for us the right to be deceived as surely as the right to truth: to extol the latter and deny the former is to misunderstand being human.

—David Nyberg, *The Varnished Truth*

This article was published Online First October 11, 2021.

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

Data and materials are available through the Open Science Framework: <https://osf.io/p6wch/>

This research is based on a chapter of the author's dissertation. The author is grateful for feedback and advice from the author's committee (Maurice Schweitzer, Katherine Milkman, Taya Cohen, Andrew Carton, and Adam Grant), as well as Daniel Bartels, Reid Hastie, Ed O'Brien, Jane Risen, Alex Shaw, and attendees of the Center for Decision Research brown-bag. Many thanks to Hussein Elkheshen, Solomon Lister, Mark Liam Murphy, Ilana Nathans, Michael White, and Ilan Wolf for their research assistance. This research was supported by the Operations, Information, and Decisions department, the Wharton Behavioral Lab, and the Wharton Risk Management and Decision Processes Center Russell Ackoff Fellowship at The Wharton School, University of Pennsylvania, as well as the Center for Decision Research, and the William S. Fishman and Charles E. Merrill Faculty Research Funds at the University of Chicago Booth School of Business. A Russell Sage Foundation Behavioral Economics small grant and a Qualtrics Behavioral Research grant also made this work possible.

Correspondence concerning this article should be addressed to Emma E. Levine, Department of Behavioral Science, The University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, United States. Email: Emma.Levine@chicagobooth.edu

We typically assume that deception is wrong and that individuals do not want to be deceived. Parents tell children that they should never lie (Talwar et al., 2007); most major religions take a categorical stance against deception (e.g., “Thou shalt not bear false witness,” The Old Testament, Exodus 20: 1–17), and recently, a number of public figures have revitalized the idea that complete honesty is essential for effective relationships and happiness (e.g., Blanton, 2005; Dalio, 2017; Harris, 2013; Scott, 2017). For example, psychotherapist Brad Blanton, whose book *Radical Honesty* has been a best-seller for decades, purports that “The best way to reduce stress, make life work, and heal the past is to tell the truth” (<https://www.radicalhonesty.com/>). Empirical research in psychology and economics largely echoes these claims. Numerous articles have documented the interpersonal costs of deception including anger, reduced liking, and distrust (Boles et al., 2000; Croson et al., 2003; Tyler et al., 2006).

Despite this view of deception, lying is ubiquitous in everyday life. People lie in roughly 20% of their social interactions (DePaulo & Kashy, 1998). They tell selfish lies to harm others and get ahead, but they also tell prosocial lies to make others feel more confident or to avoid hurting others' feelings (DePaulo & Bell, 1996; DePaulo & Kashy, 1998; DePaulo et al., 1996). In economic interactions, individuals may tell prosocial lies to help generate more money for a specific person, restore equality, or benefit their group (Erat & Gneezy, 2012; Gino & Pierce, 2009, 2010; Levine & Schweitzer, 2014, 2015; Wiltermuth, 2011).

The majority of existing research on judgments of deception has explored selfish lies, lies that benefit the liar at the expense of another party (e.g., Effron et al., 2015; Gino et al., 2009; Gneezy,

2005; Mazar et al., 2008; Shalvi et al., 2011; Shu et al., 2011; Yip & Schweitzer, 2016). Selfish lies are seen as unethical, and judgments of their unethicity increase as a function of the magnitude of harm they cause (Gneezy, 2005). Selfish lies also irreparably damage trust. In a repeated trust game, for example, Schweitzer et al. (2006) found that actors who betrayed someone's trust by telling selfish lies (e.g., breaking promises in order to exploit another person) were never able to fully restore trust. The insidious effects of selfish lies are unsurprising, given that they reflect violations of two of the most important moral values: honesty and benevolence (Haidt & Graham, 2007; Walker & Hennig, 2004).

But what happens when honesty and benevolence conflict? How do people make moral sense of prosocial lies? In the past decade, several articles have started to answer this question (Gino & Pierce, 2009; Hildreth et al., 2016; Levine & Schweitzer, 2014, 2015; Levine et al., 2018; Weisel & Shalvi, 2015; Wiltermuth, 2011). This body of research reveals that people are more likely to lie when they can justify it as benevolent (Weisel & Shalvi, 2015; Wiltermuth, 2011), and generally prioritize benevolence over honesty when judging others (Levine & Schweitzer, 2014, 2015). For example, lies that help others earn money are judged to be more ethical than truths that cause monetary damages (Levine & Schweitzer, 2014), and those who tell them are often judged to be trustworthy (Levine & Schweitzer, 2015). These findings are consistent with the growing body of research on the centrality of harm in moral judgment (e.g., Gray et al., 2012; Gray et al., 2014; Schein & Gray, 2015). Lies, like many other behaviors, are judged to be moral based on the degree to which they harm, or prevent harm, to others.

Though the primacy of benevolence over honesty has been established, the question of when, why, and which lies are seen as benevolent—and thus, which lies are seen as ethical by both communicators and targets—remains unanswered. Most existing work investigating reactions to prosocial lies employs economic games (Erat & Gneezy, 2012; Levine & Schweitzer, 2014, 2015) or studies deception as it relates to stealing and cheating (e.g., Wiltermuth, 2011). Economic games are useful for cleanly disentangling benevolence from honesty, and existing research examining cheating has yielded tremendous insight into how unethical behavior permeates group contexts. However, both of these bodies of research provide little insight into the complex moral judgments involved in everyday communication.

In everyday communication, people may be reluctant to attribute benevolent motives to deception, and therefore, may rarely—if ever—conclude that deception is more ethical than honesty. People may be reluctant to see deception as benevolent for (at least) three reasons. First, in routine conversations, it is difficult to disentangle benevolent and self-serving motives. Any act of false praise, comfort, or encouragement could be driven by a genuine desire to promote the welfare of the target *or* by the communicator's selfish desire to avoid conflict or improve their reputation (DePaulo et al., 1996). Second, in routine conversations, there is more ambiguity about what statements actually promote the welfare of a target. Whereas economic games allow for clear inferences about what is in a target's best interest (i.e., more money is better than less money), everyday conversations are not as clear. A communicator may believe an act of false praise promotes the target's welfare by protecting their feelings, whereas a target may feel that false praise undermines their welfare by limiting their

knowledge and autonomy (Harris, 2013; Kant, 1785/1959; Lupoli et al., 2018). Third, the long-term costs of deception are more salient in routine conversations between long-term relational partners than they are in one-shot economic games between strangers. Specifically, relational partners may be concerned that a single act of deception undermines trust in the communicator's future statements (Bok, 1978; Kant, 1949; Levine & Schweitzer, 2015), which could degrade the quality of a communicator-target relationship and ultimately harm the target. Given these dynamics, the circumstances in which deception credibly signals benevolent intent remain unclear.

The present investigation develops and tests a theoretical framework that unearths these circumstances and therefore clarifies when and why deception is rewarded in routine conversation. Deception is perceived to be ethical and individuals want to be deceived when honesty causes *unnecessary harm*. Perceptions of unnecessary harm are driven by two key factors: the degree to which deception will prevent harm to an individual at the moment of communication, and the instrumental value of truth (i.e., the degree to which honest information may yield meaningful learning, growth, or behavioral change). When honesty is perceived to be high in immediate harm and low in instrumental value, deception is perceived to be more ethical than honesty. Importantly, this framework can explain numerous and systematic circumstances in which individuals want to be deceived, and in which deception is judged to be ethical.

This is the first work to develop a moral folk theory of prosocial deception. As a result, this research makes important contributions to our understanding of both deception and moral psychology. First, this research advances our understanding of deception and communication by bridging research in these two domains, along with research in theology and philosophy, to elucidate the psychological principles that people draw upon when using and judging deception. In addition to providing an organizing framework for thinking about prosocial deception, this research establishes a set of community standards of deception: rules of deception that most people abide by and recognize once articulated, but have never previously been codified. Establishing these rules allows us to draw predictions about when communicators will use deception, and when targets will appreciate deception and penalize its opposite—honesty. Just as Kahneman et al.'s (1986a, 1986b) work on community standards of fairness overturned the assumption that individuals universally value self-interest, and demonstrated that concerns about fairness place systematic, rather than anomalous, constraints on market behavior, the present research challenges the assumption that people universally value full information (i.e., truth) and demonstrates that people have systematic preferences for deception.

In doing so, this work also challenges existing assumptions about how people fundamentally reason about deception. Prior work suggests that people engage in deception primarily due to cognitive blind spots (Bazerman & Tenbrunsel, 2012; Sezer et al., 2015), a lack of self-control (e.g., Gino et al., 2011), or the ability to rationalize small lies (Mazar et al., 2008). This past work suggests that lying is often a mistake, that people rarely justify their lies a-priori or would defend their lies to others. In contrast, the present work sheds light on the acts of deception that are well-reasoned, intentional, and seen as genuinely justified.

These findings also advance our understanding of moral psychology. Though honesty is one of the most important values in everyday morality, contemporary frameworks in moral psychology have largely ignored honesty (for example, moral foundations theory, [Graham et al., 2013](#); dyadic morality, [Gray et al., 2014](#); [Gray & Wegner, 2011](#); dual-process models of morality, [Greene et al., 2001](#); [Greene & Haidt, 2002](#); the social heuristics hypothesis, [Rand et al. 2014](#)). When people are asked open-ended questions about what they feel guilty about, honesty violations are the most common violation that individuals recall ([Iyer, 2010](#)), and acts of deception are among the most common immoral events that people witness in everyday life ([Hofmann et al., 2014](#)). Yet, most existing research has focused on how people understand more complex—and rare—immoral actions, such as torture ([Gray & Wegner, 2010](#)), murder ([Cushman et al., 2012](#)), sexual misconduct ([Haidt & Hersh, 2001](#); [Helzer & Pizarro, 2011](#)), dog-eating ([Haidt et al., 1993](#)), and of particular prominence, sacrificial harm (see [Bauman et al., 2014](#) for a review). More recent research has developed the psychology of utilitarianism in important ways but notes that “issues relating to honesty and promise-keeping are not central, or even particularly important, to utilitarianism” ([Kahane et al., 2018](#), p. 145). The present research develops and tests a theoretical framework that helps to explain the use and judgment of a vital and morally relevant human behavior: deception.

Finally, this work helps to bridge descriptive and normative ethics. Sissela Bok’s “Test of Publicity” ([Bok, 1978](#)), a widely-accepted normative standard for deciding whether deception is ethical, asks “which lies, if any, would survive the appeal for justification to reasonable persons” ([Bok, 1978](#), p. 93). To pass this test, a lie must be acceptable to all parties affected by a lie. Until now, the “Test of Publicity” has been a hypothetical, rhetorical tool that people can consider when deciding whether any particular lie is acceptable. The present research, however, provides an empirical answer to this test by identifying the set of lies that communicators, observers, and targets deem acceptable. Although this research does not attempt to provide any normative advice regarding whether people *should* lie in these circumstances, it does shed light on which lies people *believe* they and others should tell, even upon reflection. In doing so, this work contributes to public discourse on deception.

Scope and Overview of Research

The goal of the present research is to develop a framework for understanding when and why people believe that deception is ethical in everyday conversation. *Deception* is conceptualized as any act that intentionally misleads others ([Boles et al., 2000](#); [Gino & Shea, 2012](#); [Murnighan, 1991](#)). This includes both lies of commission and lies of omission.

This work asks: when, if ever, is deception seen as the right thing to do, relative to the alternative of honesty? There are many features of deception, such as how much a statement departs from the truth ([Rogers et al., 2017](#)) and the perceived likelihood that the deceptive statement could have been true ([Effron, 2018](#); [Shalvi et al., 2011](#)), that influence the perceived ethicality of deception and have been the focus of prior work. However, none of these tactics are seen as more ethical than telling the truth, particularly in the eyes of targets. The purpose of the present work is not to explore additional features that cause

lies to seem more or less justified, but rather to identify conditions under which lying is seen as objectively ethical (i.e., more ethical than honesty). We answer this question from the perspective of all people affected by deception, including communicators, observers, and targets of deception.

Finally, *everyday conversation* refers to routine, dyadic conversations between communicators and targets. Though people may also approve of deception in extreme circumstances (e.g., if there were a Nazi at your door and telling the truth would lead to the death of an innocent victim, [Varden, 2010](#)), these circumstances are outside of the scope of the present research. Many would and have argued that deception is acceptable in exceptional circumstances. However, the present research argues that deception is often not exceptional; people welcome and approve of deception in systematic, unexceptional ways that fundamentally affect our social functioning.

This research unfolds in two parts. It begins with an inductive study, in which participants answered open-ended questions about their preferences for, and moral judgments of, deception. This study revealed that the most common types of lies that people believe are ethical are *prosocial* lies, and in particular, lies that prevent harm to the target.

First, the study and the coding processed used to analyze the study are described. Second, the theoretical framework – the unnecessary harm framework – that was derived from this study, is introduced. This framework clarifies the dimensions of harm people consider when making moral judgments of prosocial lies. Third, several rules of deception that reflect situational antecedents of unnecessary harm are introduced.

The second part of this article presents four empirical tests of the unnecessary harm framework. Across all studies, stopping rules for data collection were decided in advance. Each study was run with multiple samples, which was decided a priori, to ensure robust results across populations. The exact stopping rules are described in the [Supplement 1](#) of the online supplementary materials. Data and materials are available through the Open Science Framework (<https://osf.io/p6wch/>). The inductive study, and Studies 2 and 3 were approved by the Institutional Review Board of the University of Pennsylvania. Studies 1 and 4 were approved by the Institutional Review Board of the University of Chicago.

Part 1: Developing the Unnecessary Harm Framework

Method

A survey with 304 adults was run (using both Amazon Mechanical Turk and a university laboratory pool; 55% female; $M_{\text{age}} = 29$ years) in which participants were randomly assigned to answer three questions about when they would want to be deceived (preferences condition; e.g., “Please come up with three concrete examples of instances in which you would want to be lied to”) or when they believed that deception was ethical (ethics condition; e.g., “Please come up with three concrete examples of instances in which it is ethical to lie”). An iterative coding procedure ([Strauss & Corbin, 1990](#)) was used, which entailed reading through participant responses to generate common categories, and refining these categories over a series of coding attempts. Each coding attempt

involved discussions with research assistant coders and consultation with existing literature. The coding process is described in detail in the [Supplement 2.2](#) of the online supplementary materials. [Table 1](#) briefly describes the coding categories for different justifications of deception and the frequency with which they appeared in participants' responses.

Overall Justifications and the Unnecessary Harm Framework

Only 5% of participants believed that deception was never acceptable, and fewer than 40% of participants judged deception based on its costs and benefits to parties other than the target (see [Table 1](#)). The vast majority of participants (91%) endorsed deception based on the degree to which it prevented harm to the target.

Importantly, participants did not suggest that deception that prevented *any* harm to the target was desirable. Rather, the coding process revealed that participants relied on a single construct—the prevention of *unnecessary* harm—to justify deception. Participants focused on two ways in which honesty could cause unnecessary harm to targets. First, honesty could lead to emotional pain or suffering, without having any meaningful impact on a target's future thinking or behavior (i.e., no instrumental value). A total of 69.4% of participants justified deception when honesty had no instrumental value. For example, many participants indicated that deception is ethical, and that they would want to be deceived when the truth would be painful to hear, but not actionable, or not important, as evidenced by the quotes below:

Participant 17: I would want someone to lie to me when the alternative of telling the truth would make me feel worse off and I would have no

control over what happens. For example, if my beloved dog died after being hit by a negligent driver, I'd much rather my parents or friends have told me the dog died peacefully in its sleep than to tell me the facts.

Participant 18: Another circumstance in which I would prefer someone lied to me in one in which I cannot do anything about the truth. For example, if I were out with my friends at a bar and I asked if I looked okay, I would prefer if my friends said yes because if I did not, there would be nothing I could do about it at the bar. This is very different than if I asked my friends the same question while at home, where I hope they would tell me the truth, so I could change whatever looked bad (as best as I could). In the example at the bar, if they told me the truth that I looked bad, my night would be ruined and I would have to stay at the bar knowing I looked bad, instead of blissfully being unaware that I looked bad.

Second, honesty could cause unnecessary harm if it were shared in a circumstance that would lead to elevated levels of harm (i.e., high immediate harm). A total of 70.8% of participants justified deception that prevented immediate harm. For example, many participants indicated that deception is *temporarily* acceptable when a target is particularly fragile, or when sharing the truth could distract a target from something important, as evidenced by the quote below:

Participant 4: If the truth about a situation will distract someone and affect his or her performance on an important task, I feel that lying is temporarily justified, as long as the truth is later revealed. To illustrate this, I will use an example. Last week, my friend had a midterm that she was very nervous and stressed about, and had been

Table 1
Coding Categories and Frequencies for Broad Justifications for Deception

Overall justification	Specific justification	Description for coders	Examples of participant responses	Total
Unnecessary harm avoidance	Immediate harm of honesty	These justifications include lies that are told to avoid harm to the target at the moment of communication. This type of harm is immediate and not long-lasting.	<ul style="list-style-type: none"> From my perspective, lying to someone else is the right thing to do when we can avoid hurting others or make others happy / comfortable Lying may be the right thing to do when telling that person the truth at that particular moment may be harmful to them. 	70.8%
	Instrumental value of honesty	These justifications focus on whether or not there are any potential long-term benefits of honesty. Specifically, is the honest information important, actionable, and objective? These responses suggest that lying is ok when honesty does not have the potential to affect future behavior or thinking in a meaningful way or bring about any other benefit.	<ul style="list-style-type: none"> As long as it isn't something that's incredibly important for them to know, why bother them with it when you can save them from the truth? I would want to be lied to under certain circumstances where I cannot change the result. 	69.4%
	TOTAL	This is a composite category reflecting the presence of either dimension above: immediate harm or instrumental value		91.0%
Utilitarian		These justifications incorporate costs and benefits to parties other than the target of the lie. Any responses that mention how a lie will affect the liar, society, or third parties are considered Utilitarian.	<ul style="list-style-type: none"> Lying to someone else is the right thing to do when it behooves both you and the other person to have them believe the lie. Lying may prevent conflicts. . . 	36.9%
Never (Deontological)		"Never" indicates that the participants included a statement expressing that lying is never acceptable. "Never" means that the person does not provide any justifications or examples of when/why lying is right.	<ul style="list-style-type: none"> There is no instance where lying to someone else is the right thing to do. I would never want someone to lie to me. 	5.0%

studying for a long time. The day of the test, I found out information about a boy that she was interested in. I knew that I had to tell her, however if I told her before her test, she may have been distracted, upset, and not able to fully focus. This could have compromised her performance on her test and may have caused her to get a significantly worse grade than she deserved. So, I lied to her when I told her what I did for lunch that day

In these cases, participants typically indicated that truth should be shared at a later time. Nonetheless, deception in that moment, was seen as acceptable.

These insights led to the development of a simple framework that clarifies when deception is seen as acceptable (see Figure 1). Individuals want to be deceived, and believe deception is ethical when it prevents unnecessary harm to the target. Judgments of unnecessary harm depend on two factors: the degree to which honesty has instrumental value (i.e., can lead to meaningful growth, understanding, or behavioral change in the long-run), and the degree to which honesty poses immediate harm.

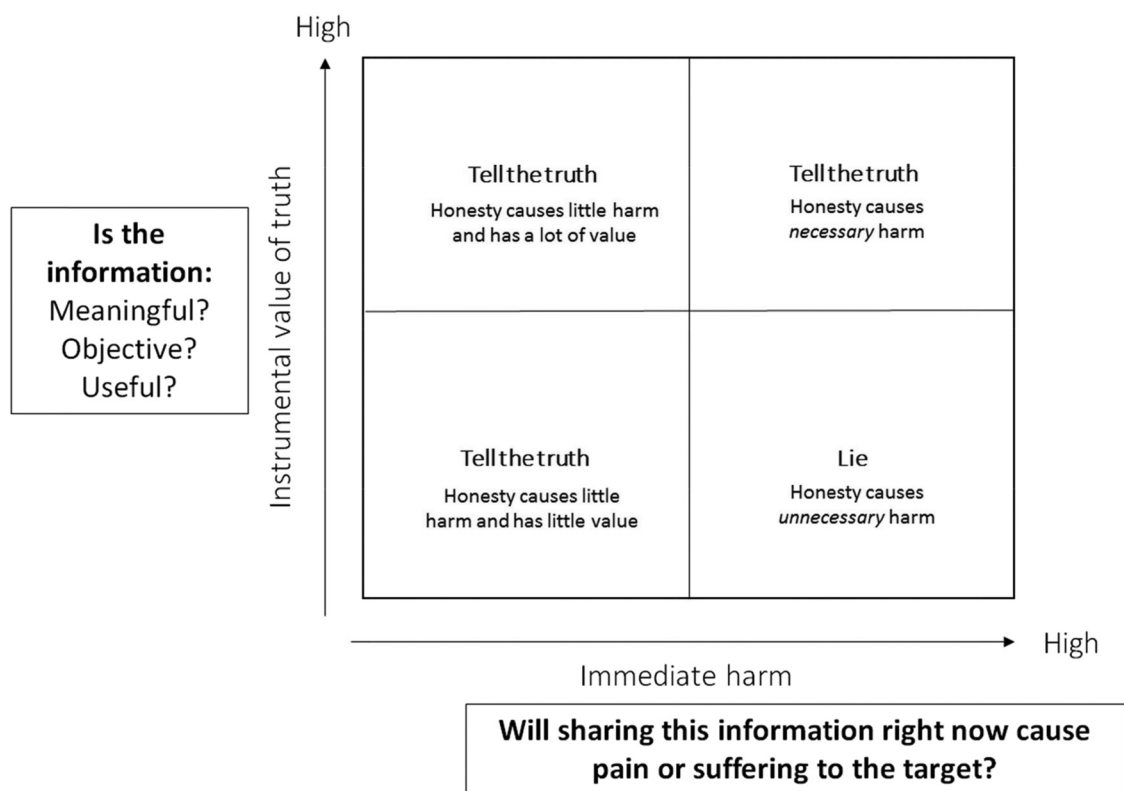
Figure 1 depicts the theoretical relationship between these two factors and the endorsement of deception. When honesty poses no immediate harm and provides no instrumental value (lower left quadrant), individuals are expected to endorse honesty. This

proposition is consistent with research demonstrating that people have preferences for truth, even when it has no costs or benefits (Holland et al. 1986; Levine & Schweitzer, 2014; Ortony & Gupta, 2018). Of course, when honesty does provide instrumental value, the preference for honesty is likely to be even stronger (upper left quadrant). When honesty does cause immediate harm, but is associated with instrumental benefits (upper right quadrant), individuals are also expected to endorse honesty, though perhaps not as strongly. In these circumstances, honesty causes *necessary* harm (Margolis & Molinsky, 2008; Molinsky & Margolis, 2005).

However, when honesty causes immediate harm and is not associated with instrumental value (lower right quadrant), honesty causes *unnecessary* harm. In these circumstances, there is expected to be the greatest consensus that deception is more ethical than honesty. In subsequent studies, this prediction is tested by examining (a) whether the endorsement of deception in the low instrumental value/high immediate harm quadrant is higher than all other quadrants, and (b) whether the percentage of participants who believe deception is more ethical than honesty is greater than 50% in the low instrumental value/high immediate harm quadrant.

Notably, this prediction—that deception is most likely to be seen as more ethical than honesty when immediate harm is high and

Figure 1
Deception and Unnecessary Harm Avoidance: A Framework



Note. A qualitative study revealed that participants rely on two key dimensions when judging the ethicality of deception: the degree to which honesty causes immediate harm (i.e., avoidable pain and suffering) and the degree to which honesty has instrumental value (i.e., will lead to long-term learning, growth, and behavioral change). When honesty is high in immediate harm and low in instrumental value, deception is perceived to prevent *unnecessary harm*, and therefore is seen as more ethical than truth-telling.

Table 2*Explanations of and Existing Support for the Community Standards of Deception*

Category of antecedent	Situational antecedent	Statement of community standard	Relationship to unnecessary harm	Other discussions of or evidence for the standard	Freq*
Attributes of target	1. Emotional fragility	It is acceptable to lie when targets are fragile	Telling the truth to someone who is vulnerable → increased immediate harm Telling the truth to someone who cannot process and react to the information rationally → decreased instrumental value	Psychology: People avoid information when they have low coping resources (Sweeny et al., 2010) Philosophy: Plato argues that falsehood is permissible when it prevents a mentally unstable person from harming himself (Jowett & Campbell, 1894)	4.7%
	2. Ability to understand	It is acceptable to lie when targets cannot understand the truth	Telling the truth to someone that will be confused → increased immediate harm Telling the truth to someone if they cannot process and react to the information rationally → decreased instrumental value	Psychology: People avoid information when they expect not to understand it (Sweeny et al., 2010) Applied ethics: Discussions in medical ethics question whether its ethical to lie to patients who lack cognitive capacity (i.e., therapeutic fibbing; Beach & Kramer, 1999; Richard et al., 2010)	15.3%
	3. Death bed	It is acceptable to lie when targets are near the end of their life	Telling the truth to someone who is vulnerable → increased immediate harm Telling the truth to someone who does not have time to act on it → decreased instrumental value	Philosophy: Aristotle argued that deception might conditionally be good when one's ability to evaluate a situation is detrimentally affected by illness (Zembaty, 1993)	7.3%
	4. Subjective versus objective	It is acceptable to lie about subjective information	Telling someone a subjective opinion, rather than an objective fact → decreased instrumental value	Psychology and communication: Misrepresenting one's subjective opinions is often part of politeness (Brown & Levinson, 1987; Goffman, 1955)	29.6%
	5. Trivial versus important	It is acceptable to lie about trivial information	Telling someone something that is not important → decreased instrumental value	Philosophy: People believe it is acceptable to lie about trivial things (traditional white lies; Bok, 1978) Communication: Politeness settings are those in which the truth is insignificant, but telling it is hurtful (Sweetser, 1987)	29.2%
	6. Controllable versus uncontrollable	It is acceptable to lie about things we cannot control or fix	Telling the truth to someone about something they cannot change or act upon → decreased instrumental value	Psychology: People avoid information when they lack control over the consequences of the information (Sweeny et al., 2010) Religion: In the Talmud, the House of Hillel argues that you should praise someone's purchase of an inferior good after it has been purchased, but not before it has been purchased (Telushkin, 1994)	19.6%
Attributes of context	7. Disruption to special moments and events	It is acceptable to lie to preserve special moments	Telling someone something that will distract them from something more important → increased immediate harm	Religion: in the Talmud, the House of Hillel argues that you should tell a bride she is beautiful on her wedding day, regardless of the truth (Telushkin, 1994)	6.0%
	8. The presence of others	It is acceptable to lie to help others save face	Telling someone something that will embarrass them → increased immediate harm	Psychology and communication: People believe it is acceptable (and in certain cultures, necessary) to modify behavior in public (sometimes by use of deception) to help others save face (Ho, 1976; Hofstede, 1980; Schwartz, 1992)	1.3%

Note. The qualitative study identified 8 community standards of deception, which specify conditions under which honesty is perceived to cause unnecessary harm. These standards are reviewed above. Details on the coding scheme used in the inductive study are available in the [Supplement 2](#) in the online supplemental materials.

*Frequency indicates the frequency with which each standard was mentioned in the inductive study ($N = 304$).

instrumental value is low—could be supported by two main effects or by an interaction between instrumental value and immediate harm. See [Supplement 7](#) of the online supplementary materials for greater explanation and visualization.

The Situational Antecedents of Unnecessary Harm (i.e., Some Community Standards of Deception)

In addition to revealing the abstract principles that justify deception, participants elucidated the specific situations in which those principles apply. These situations illustrate a number of community standards of deception, which pertain to the attributes of the target, the topic of honest information, and the context of the conversation. For example, participants endorsed lying to emotionally compromised targets (Standard 1) because they believed that honesty would cause the greatest immediate harm to fragile targets. [Table 2](#) describes these standards, explain how each standard relates to the proposed dimensions of unnecessary harm, and provide the frequency with which each standard appeared in participants' responses.

The inductive study generated eight community standards. When the truth causes harm, people believe it is ethical to lie to people who (Standard 1) are emotionally fragile; (Standard 2) lack the cognitive capacity to understand the truth; or (Standard 3) are at the end of their life; people believe it is ethical to lie about (Standard 4) subjective information; (Standard 5) trivial information; or (Standard 6) information about things that cannot be changed; and people believe it is ethical to lie when (Standard 7) true information would disrupt a sacred event; and (Standard 8) the conversation occurs in front of others. It is important to note that these eight situations are not an exhaustive list of community standards of deception. Rather, these eight situations reflect common and salient situations that lay people link with the construct of unnecessary harm.

Notably, the standards that correspond with the unnecessary harm framework are also supported by empirical research on interpersonal communication and information avoidance, as well as religious and philosophical discussions on the ethics of deception. For example, existing research has shown that people avoid information when they have no control over the consequences of the information ([Shiloh et al., 1999](#); [Yaniv et al. 2004](#)). The present research suggests that people also want to be deceived, and see deception as ethical, in these circumstances. [Table 2](#) summarizes convergent evidence for each community standard of deception.

Discussion

The unnecessary harm framework makes three central contributions to our psychological understanding of deception. First, it clarifies that people—including potential targets of deception—believe that it is ethical to lie when lying prevents harm to targets. Rather than considering the potential long-term costs of deception (i.e., to trust or individual autonomy), as many normative scholars have done ([Bok, 1978](#); [Kant, 1949](#)), communicators and targets seem to take a relatively myopic view. They consider whether lying at a particular moment would ultimately promote the target's welfare, consistent with recent work on prosocial lies ([Levine & Schweitzer, 2014, 2015](#)).

Importantly, this work extends research on prosocial lies by documenting the *dimensions* of harm that communicators, targets, and observers care about when judging deception in everyday life. People consider both the immediate harm caused by honesty and the instrumental value of honesty. This proposition is consistent with foundational work by [DePaulo and Bell \(1996\)](#), which also makes a distinction between the emotional and instrumental consequences of honesty. [DePaulo and Bell \(1996\)](#) found that people are more likely to give false praise about topics that are important to a target than topics that are unimportant to the target, suggesting this occurs because communicators are more attentive to emotional harm than instrumental value. Importantly, however, [DePaulo and colleagues](#) (see also [DePaulo & Kashy, 1998](#)) never examine whether individuals consciously assess these attributes, nor do they consider how these attributes influence moral judgment and target preferences.

Third, the unnecessary harm framework unearths the psychology underlying a number of seemingly unrelated community standards of deception. Lay people, scholars, and religious leaders have identified a number of “exceptions” to the norm of honesty (see [Table 2](#) for more details). These exceptions, however, are typically discussed as one-off anomalies, rather than systematic and justified deviations from honest communication. The unnecessary harm framework explains *why* these exceptions have been raised, thereby unifying findings and arguments across psychology, religion, and philosophy under a common theoretical framework.

Part 2: Empirical Tests

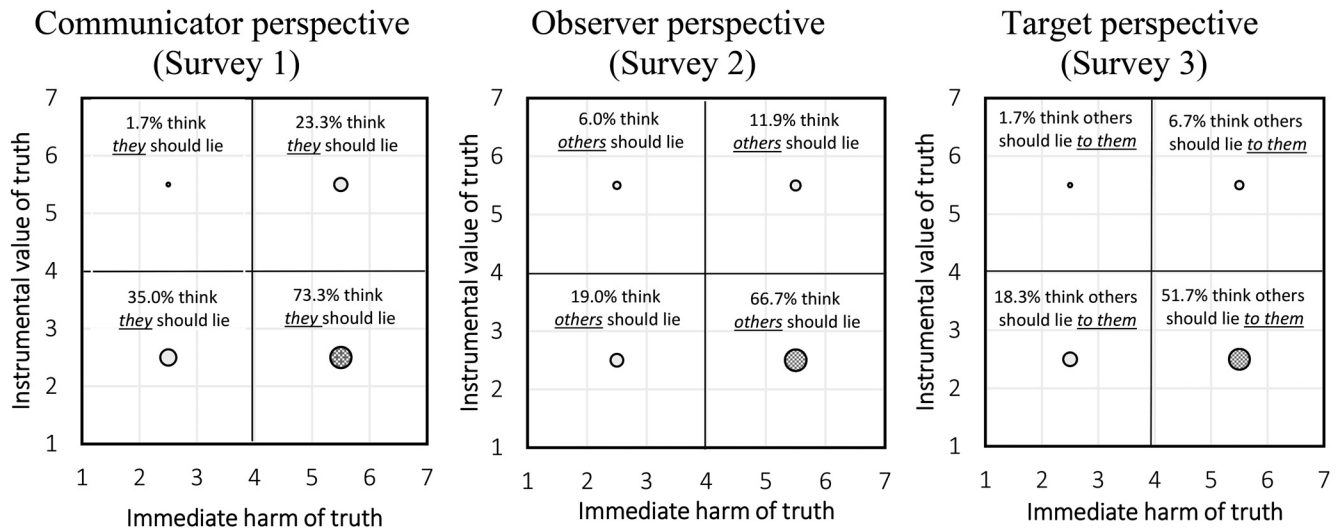
The next section of this article presents four sets of empirical tests of the unnecessary harm framework. Study 1 presents three experiments that test the causal effect of immediate harm and instrumental value on communicators', targets', and observers' endorsement of honesty. Study 2 presents three experiments that establish judgments of immediate harm and instrumental value as mechanisms linking the community standards of deception with the endorsement of deception. Study 3 rules out alternative mechanisms underlying these standards. Finally, Study 4 documents how perceptions of unnecessary harm influence the use and judgment of deception in everyday life.

Study 1: Directly Manipulating the Dimensions of Unnecessary Harm

Method

Study 1 consisted of three separate survey experiments. In each survey, participants were assigned to a condition from a 2 (immediate harm: low, high) \times 2 (instrumental value: low, high) within-subjects design. The first survey ($N = 60$ participants from a U.S. university laboratory; 47% female; $M_{\text{age}} = 26$ years) examined communicators' moral judgments of deception. The second survey ($N = 84$ participants on Amazon Mechanical Turk; 51% female; $M_{\text{age}} = 36$ years) examined third parties' moral judgments of deception. The third survey ($N = 60$ participants on Amazon Mechanical Turk; 50% female; $M_{\text{age}} = 36$ years) examined targets' preferences for deception.

Figure 2
Manipulating the Dimensions of Unnecessary Harm



Note. In Study 1, participants learned about a situation in which a person was considering sharing a truth that was either high or low in immediate harm and either was high or low in instrumental value. Participants in every perspective - Communicators, Observers, and Targets - believed that deception was the most ethical when the truth was high in immediate harm and low in instrumental value, consistent with the unnecessary harm framework.

In each survey, participants read four short scenarios (in a random order) that described a context in which honest information was (or was not) immediately harmful and had (or lacked) instrumental value. Each scenario corresponds with a different quadrant of the unnecessary harm framework (see Figure 1). For example, in the communicator survey (Survey 1) the scenarios read as follows:

Imagine that you have a negative opinion of another person that you are considering sharing.

*You know that sharing this opinion right now [low immediate harm: **will not be particularly hurtful** / high immediate harm: **will be incredibly hurtful**].*

*This information is [low instrumental value: **not very important** to know and **cannot influence** the person's behavior/ high instrumental value: **important** to know and **can influence** the person's future behavior.]*

In each of the four scenarios, participants were then asked, "Should you truthfully share this information with the person or lie to this person?" (choices: "Truthfully share this information" or "Lie to this person."). Participants then were asked to explain their answer and generate examples that fit the scenario.

The third-party survey (Survey 2) and the target survey (Survey 3) were nearly identical, except the scenarios were written from a third-party or target perspective. At the end of each survey, participants answered a few demographic questions.

Results and Discussion

First, the percentage of participants who believed lying was ethical in each of the four scenarios are examined. Then, the results of a logistic regression on the endorsement of deception, using *immediate harm*, *instrumental value*, and their interaction as independent variables are presented.

Consistent with the proposed theory, the majority of participants—across all perspectives—endorsed deception when information caused high immediate harm and had low instrumental value. The majority of communicators (73.3%, significantly greater than 50%; $p < .001$), third-party judges (66.7%, significantly greater than 50%; $p = .002$), and targets (51.7%, not significantly greater than 50%, $p = .796$) endorsed deception in this quadrant. Conversely, the majority of participants endorsed honesty in all other quadrants.

Planned contrasts confirmed that the frequency with which deception was endorsed was significantly greater in the low instrumental value/high immediate harm quadrant than all other quadrants (all $ps < .001$). Figure 2 depicts the frequency with which participants endorsed deception in each quadrant. The size of each data point is proportional to the percentage of people who endorsed deception in each condition, within each survey. Each data point is centered within its respective quadrant.

A series of mixed effects logit models (using the melogit function in Stata) on the endorsement of deception (1 = lying is endorsed, 0 = truth-telling is endorsed) including *immediate harm*, *instrumental value*, and their interaction as independent variables (see Table 3) was also run. A fixed-effects approach to control for vignette, and a random-effects approach to account for multiple observations per participant was used. Pseudo- R^2 was calculated using the method described in Tjur (2009).¹

¹ Logistic regressions in which standard errors were clustered at the participant-level to account for within-participant dependencies, were also run. These analyses yielded qualitatively identical results. This was the case in Study 1, as well as Studies 2 and 3. Code for these analyses are available on OSF (<https://osf.io/p6wch/>). During the review process the random-effects models were recommended, and LR tests confirmed that these were a better fit for the data in most cases. Therefore, random effects models are reported in the main manuscript.

Table 3*Effects of Immediate Harm and Instrumental Value on the Endorsement of Deception (Test 1)*

Independent variables	Survey 1 (Communicators)		Survey 2 (Third parties)		Survey 3 (Targets)		Pooled data	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Immediate harm of truth	2.56***	2.31***	2.45***	3.12***	2.11***	2.21***	2.40***	2.59***
Instrumental value of truth	-3.36***	-4.17***	-3.07***	-1.71**	-3.49***	-3.05**	-3.25***	-2.73***
Imm Harm \times InstrValue		1.08		-2.16**		-0.60		-0.76
Constant	-1.01**	-.90*	-1.69***	-2.10***	-2.05***	-2.12***	-0.95**	-1.05**
Participant random effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey (perspective) fixed effect							Yes	Yes
Observations	240	240	336	336	240	240	816	816
R ²	0.53	0.52	0.50	0.54	0.47	0.48	0.51	0.52

Note. Significant values bolded.

* $p < .05$. ** $p \leq .01$. *** $p < .001$.

There was not a significant Immediate Harm \times Instrumental interaction among Communicators or Targets. However, there was a significant Immediate Harm \times Instrumental Value interaction among observers (third parties). Among observers, immediate harm significantly influenced the endorsement of deception when instrumental value was low ($p < .001$), but not when it was high ($p = .10$). In other words, if the truth had instrumental value, observers endorsed truth telling, regardless of whether it caused immediate harm. Alternatively, if the truth lacked instrumental value, observers endorsed deception if the truth caused immediate harm, but endorsed truth telling otherwise. Cumulative evidence for a potential Immediate Harm \times Instrumental Value interaction is discussed in the General Discussion.

Study 2: Manipulating the Situational Antecedents of Unnecessary Harm (i.e., the Community Standards of Deception)

Study 2 extends the present investigation by manipulating each of the community standards of deception identified in my pilot study. This study tests how each standard influences the two proposed dimensions of unnecessary harm (immediate harm and instrumental value of truth), and consequently, the endorsement of deception. This study, therefore, provides a richer test of the unnecessary harm framework than Study 1, using common exemplars that are both evident in lay people's discussions of deception (revealed through participant responses) and philosophical, religious, and psychological scholarship (see Table 2).

Method

Procedure

Study 2 consists of three survey experiments with three different samples, each of which featured three different vignettes (total $N = 731$). Each vignette examined one of the community standards of deception identified in the inductive study. Each vignette featured a manipulation of whether or not the relevant standard was violated (between subjects). For example, Vignette 8 (the *Presence of Others* vignette) manipulated whether the target had the opportunity to receive negative feedback in public or private. Two

different vignettes were used to test Standard 6. Table 4 features the exact vignettes that tested each standard.

In each vignette, perspective was also manipulated; participants took the perspective of either an observer (observer condition) or the target (target condition) when judging deception. As in Study 1, the purpose of the perspective manipulation was to examine whether or not targets' preferences for deception converged with observers' moral judgments. In one survey, perspective was manipulated between-subjects and in two surveys, perspective was manipulated within subjects (see Table 4 for more details). Results did not differ based on whether perspective was a within or between subjects factor.

Each participant responded to three vignettes (in a random order), but never saw more than one version of the same vignette. Table 4 reports the vignettes that appeared in each survey, the sample details, demographics, and any design differences between the three surveys. Additional details about each survey are reported in the Supplement 3 of the online supplementary materials.

Dependent Variables

The main dependent variable in each vignette was a dichotomous choice: Participants chose whether truth-telling or lying was the preferred communication tactic in each vignette.² After participants selected the most ethical (their most preferred) response, participants answered a series of questions intended to examine perceptions of unnecessary harm (e.g., "To what extent would telling a lie protect the [individual]'s feelings?" and "To what extent would telling the truth in this scenario have the potential to influence the [individual]'s behavior?"). All items were measured using seven-point rating scales anchored at 1 = *not at all* and 7 = *extremely*. The items vary slightly in each survey, as the scales were refined. In each survey, an exploratory factor analysis (principal axis factoring, varimax rotation) on all the items was conducted. In every survey, the items loaded on to two distinct factors, which reflect the two dimensions of unnecessary harm:

²In the observer condition, participants answered the following question: "Which of the following options is the more ethical response?" In the target condition, participants answered the question, "Of the following options, how would you prefer that [the communicator] responds?" To answer these questions, participants chose between telling the truth and lying. The exact wording of the response options for each vignette appear in the Supplement 3.1 of the online supplementary materials.

Table 4*Manipulating the Community Standards (Study 2)*

Community standard (Vignette name)	Survey group	Scenario introduction	Control condition	Community standard violation condition
1. Emotional fragility	1	Imagine an employee who just turned in his weekly marketing report to his manager. Although the employee usually delivers good work, the manager - unbeknownst to the employee - does not think this report was done well. The employee has a meeting today with his manager.	The employee asks the manager what he thought of the report. 3.0% endorse deception	The employee's father was unexpectedly hospitalized this morning and the employee is very distressed. The manager knows this information. The employee asks the manager what he thought of the report. 19.5% endorse deception
$\chi^2 = 18.36, p < .001$				
2. Ability to understand	2	Imagine a doctor who realizes that her patient's cancer is terminal, meaning the cancer is not curable and the patient will likely die.	The doctor's patient is a 44-year-old adult. The patient can tell something is wrong and is very distressed. The patient asks the doctor if s/he is going to die. 7.7% endorse deception	The doctor's patient is a 4-year-old child. The patient can tell something is wrong and is very distressed. The patient asks the doctor if s/he is going to die. 33.2% endorse deception
$\chi^2 = 38.67, p < .001$				
3. Death bed	1	Imagine an individual who is seriously ill. During the individual's illness, his spouse cheated on him. The individual does not know this and still deeply loves his spouse.	Although the individual is still ill, he is very likely to recover. The individual is talking to his spouse about their relationship and asks if the spouse has ever been unfaithful. 31% endorse deception	The individual is still ill and is likely to die within the next 24 hours. The individual is talking to his spouse about their relationship and asks if the spouse has ever been unfaithful. 63.8% endorse deception
$\chi^2 = 28.67, p < .001$				
4. Subjective	3	Imagine an employee who must deliver an important presentation. She will pitch a new marketing plan to her manager and colleagues. She plans on wearing her favorite silk scarf during the presentation. She loves the scarf and thinks it brings her good luck. Imagine that the employee's colleague - unbeknownst to the employee - thinks the scarf is hideous.	The colleague also knows that many other employees share this opinion. The day of her presentation, the employee shows up in a suit and her silk scarf and asks how she looks in it. 39.4% endorse deception	The colleague also knows, however, that many other employees do not share this opinion. Many colleagues like the scarf. The day of her presentation, the employee shows up in a suit and her silk scarf and asks how she looks in it. 71.2% endorse deception
$\chi^2 = 54.98, p < .001$				
5. Trivial	2	Imagine an individual who is hosting a dinner party. The host serves soup, which one guest finds to be very salty. The host asks the guest what he thinks of the soup.	This individual, the host, cooks very often. The host is a professional chef and is hosting the party to try out new recipes for his/her restaurant. The host serves soup, which one guest finds to be very salty. The host asks the guest what he thinks of the soup. 18% endorse deception	This individual, the host, does not cook very often. The host has no professional cooking training and is hosting the party for fun. The host serves soup, which one guest finds to be very salty. The host asks the guest what he thinks of the soup. 37.8% endorse deception
$\chi^2 = 18.82, p < .001$				
6a. Uncontrollable (feature of person)	1	Imagine a summer intern who just delivered his end-of-internship presentation to his office. The intern stuttered quite a bit during the presentation. The intern's friend attended the presentation and believed that the intern's stutter notably decreased the quality of his presentation, compared to his fellow interns. Aside from the stutter, the presentation was pretty good.	The intern stuttered because he was nervous during this particular presentation. He can likely improve his ability to speak without a stutter. The intern's friend knows this information. The intern asks his friend what he thought of the presentation. 18.8% endorse deception	The intern stuttered because he has a diagnosed speech impediment. The intern cannot improve his ability to speak without a stutter. The intern's friend knows this information. The intern asks his friend what he thought of the presentation. 64.4% endorse deception
$\chi^2 = 93.31, p < .001$				
6b. Uncontrollable (no time to implement)	1	Imagine an employee who must deliver an important presentation. He will	The day before his presentation, the employee tells his	The day of his presentation, the employee shows up in his suit and he asks his (table continues)

Table 4 (continued)

Community standard (Vignette name)	Survey group	Scenario introduction	Control condition	Community standard violation condition
		pitch a new marketing plan to his manager and colleagues. He plans on wearing his favorite black suit during the presentation. Imagine that the employee's colleague – unbeknownst to the employee – thinks this suit is too tight and that the suit is inappropriate for the presentation.	colleague that he plans on wearing this suit and he asks the colleague how he looks in it. At this time, the employee has other suits available that he can wear. 7.6% endorse deception	colleague how he looks in it. At this time, the employee has no other suits available that he can wear. 64.4% endorse deception $\chi^2 = 93.31, p < .001$
7. Disruption to special moments and events	2	Imagine a manager who must fire 10% of his workforce. It is a Friday afternoon and top management has just given the manager a list of employees to lay off. It is the beginning of December and the manager has until January 1st to inform employees of their work status. After January 1st, employees will have 6 months – at full pay – to search for new jobs and finish their roles. Nothing about their work will change until that time. Imagine an employee who is on the layoff list. This employee has no idea that layoffs are coming, but the employee does know that the company is going through a reorganization.	The employee drops by the manager's office on his/her way out the door on Friday. The employee asks the manager if there's any news about the reorganization. 22.9% endorse deception	The employee is getting married this weekend – on Saturday – and s/he drops by the manager's office on his/her way out the door on Friday. The employee asks the manager if there's any news about the reorganization. 52% endorse deception $\chi^2 = 35.16, p < .001$
8. The presence of others	3	Imagine an employee who just turned in his weekly marketing report to his manager. Although the employee usually delivers good work, the manager – unbeknownst to the employee – does not think this report was well done.	The employee has a one-on-one meeting today with his manager. The employee enters the manager's office. The employee asks the manager what he thought of the report. 1.5% endorse deception	The employee is attending a company-wide networking event today. The employee walks into the event and begins talking to his manager and several other colleagues. In front of a group of colleagues, the employee asks the manager what he thought of the report. 38.3% endorse deception $\chi^2 = 115.90, p < .001$

Note. Table 4 depicts 8 tests of community standards of deception. In each vignette, violating a community standard significantly increased the endorsement of deception (significant values bolded). To conduct these tests, three separate surveys were run. Each survey (denoted by Survey group) featured three vignettes. In Survey Group 1: Mturk, $N = 267$; 47% female; $M_{\text{age}} = 35$, Perspective was manipulated between subjects. In Survey Group 2: U.S. university laboratory, $N = 195$; 52% female; $M_{\text{age}} = 25$, Perspective was manipulated within subjects. In Survey Group 3: Mturk, $N = 269$, 45% female; $M_{\text{age}} = 38$, Perspective was manipulated within subjects.

immediate harm, and the instrumental value of truth. Thus, composite scales for each dimension (in each survey, $\alpha s > .57$) were created. Appendix reports all scale items. After participants submitted their responses, they responded to a single item attention check (see Supplement 3.2 of the online supplementary materials) and answered demographic questions.

Results

Main Analyses

The purpose of the vignettes was to explore how each standard influenced perceptions of unnecessary harm and the endorsement of deception, rather than to examine the differences between these standards. Thus, results of each vignette are presented independently. Supplement 3.5 of the online supplementary materials also reports a set of analyses pooling all of the data, to further test the proposed theory.

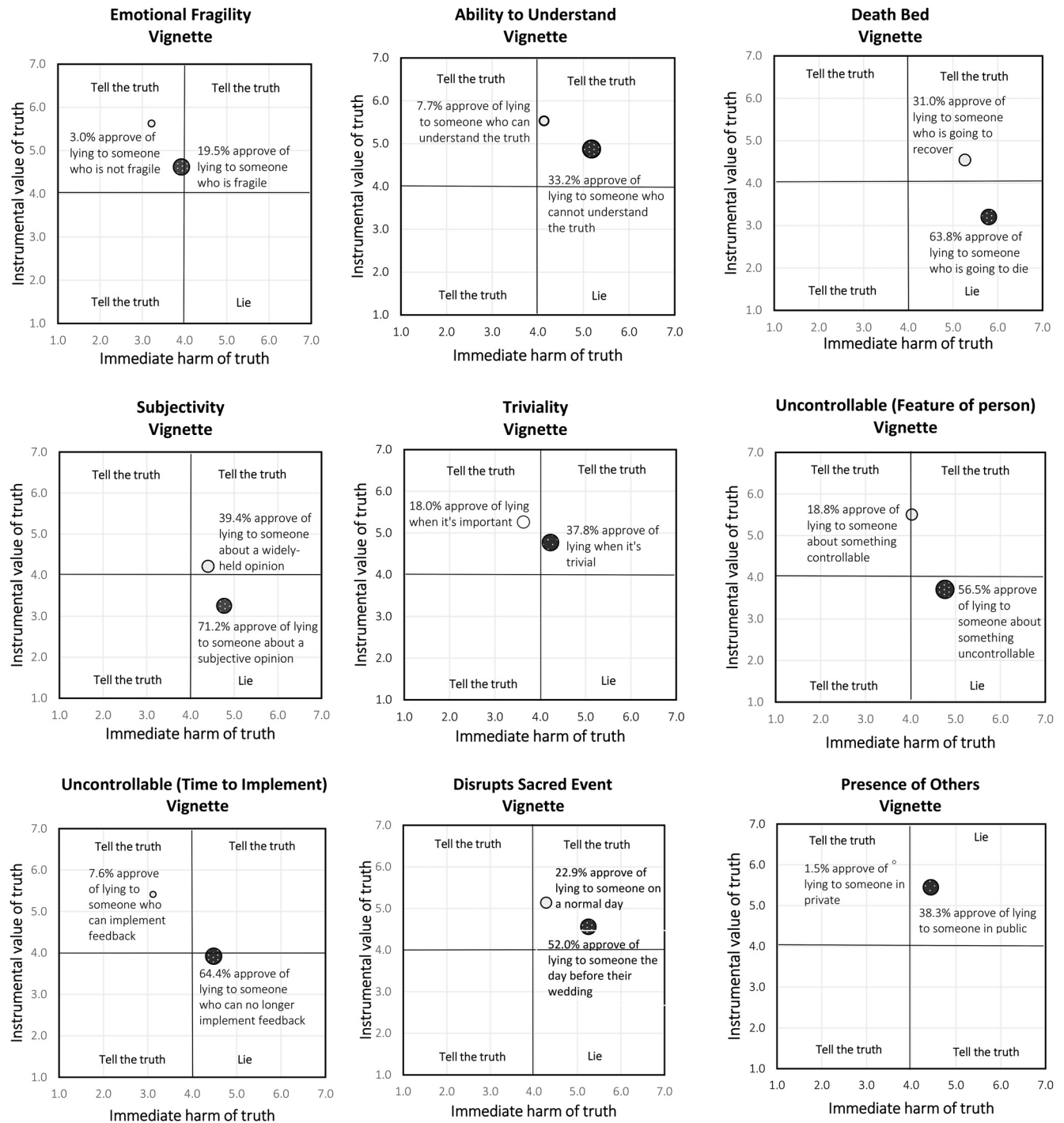
Overall, there was little evidence that targets and observers differed in their endorsement of deception. Across all vignettes, there were no main effects of perspective ($ps > .17$), and there was only an interaction between *perspective* and community standard

violation in one vignette.³ Thus, subsequent analyses are collapsed across perspective.

The main analyses feature chi-squared tests to compare the proportion of participants who endorsed deception when a community standard was or was not violated. Table 4 includes all proportions and statistical tests. There was a significant main effect of each standard in all nine vignettes (all $ps < .001$, see Table 4).

Importantly, each community standard was also mapped on to the proposed theoretical framework (see Figure 3). Specifically, for each vignette, the mean ratings of immediate harm and instrumental value of truth in the control condition and the mean ratings of immediate harm and instrumental value of truth in the community standard violation condition were plotted. The size of each data

³ Specifically, in Vignette 7 (disruption to special moments and events), there was a significant Perspective \times Community Standard Violation interaction on the endorsement of deception ($p = .027$). The interaction reveals that the effect of the *community standard violation* was greater in the preferences condition than in the ethics condition, but there was still a significant effect of the standard violation across both perspectives ($ps < .01$).

Figure 3*Mapping the Community Standards Onto the Unnecessary Harm Framework (Study 2)*

Note. For each vignette, the mean ratings of immediate harm (X axis) and instrumental value (Y axis) in the control condition (white dot) and in the community standard violation condition (dark gray, spotted dot) is plotted. The size of each data point is proportional to the percentage of people who endorsed deception in each condition. In every vignette, violating a community standard increased the endorsement of deception. Violating a community standard generally increased judgments of immediate harm and decreased judgments of the instrumental value of truth.

point is proportional to the percentage of people who endorsed deception in each condition. These graphs suggest that violating a community standard generally increases perceptions of immediate harm and lowers perceptions of the instrumental value of truth.

The four quadrants of the proposed theoretical framework also closely align with my empirical data. The majority of participants endorsed deception in vignettes that were judged to be in the high immediate harm-low instrumental value (lower right) quadrant of the theoretical framework, and endorsed honesty in all other quadrants. This pattern of results was also present when looking at participants' own ratings of immediate harm and instrumental value, across vignettes. Specifically, the majority of participants (84.3%; significantly greater than 50%; $p < .001$) endorsed deception, across scenarios, when they rated the truth as high in immediate harm (higher than 4, the midpoint of the scale) and low in instrumental value (lower than 4, the midpoint of the scale). In contrast, only 31.0% of participants endorsed deception when they rated the truth as low in immediate harm and low in instrumental value, 3.2% of participants endorsed deception when they rated the truth as low in immediate harm and high in instrumental value, and 33.1% of participants endorsed deception when they rated the truth as high in immediate harm and high in instrumental value (all proportions significantly lower than 50%; $ps < .001$).

Mediation Results

Mediation analyses, pooling all of the data from the nine vignettes were also conducted, to test the proposed theory more precisely. Across the nine vignettes, we expected perceptions of immediate harm and the instrumental value of truth to mediate the effects of community standard violations on the endorsement of deception.

A multilevel logistic mediation model using Stata's *gsem* function was performed. A random intercept was included in each model at the participant level to account for multiple observations per participant, and fixed effects were included to control for scenario and perspective.⁴ The mediation model included community standard violation as the independent variable, perceptions of immediate harm and instrumental value as simultaneous mediators, and the endorsement of deception as the dependent measure. There was significant evidence of mediation through both perceptions of immediate harm (indirect effect = 1.12, $SE = .08$, 95% CI [.96, 1.27]) and perceptions of the instrumental value of truth (indirect effect = .85, $SE = .07$, 95% CI [.72, .98]).

Discussion

Each of the community standards identified in the inductive study represents a situational antecedent of unnecessary harm. This study finds that each standard has a significant causal effect on targets' desire for and observers' moral judgments of deception, and that perceptions of immediate harm and instrumental value underlie these effects. When a community standard is violated—for example, when a target does not have time to implement feedback—honesty is perceived to be more harmful at the moment of communication and to yield less instrumental value. In these circumstances, deception is perceived to be more ethical than honesty and targets want to be deceived.

Study 3: Exploring Alternative Theories

Studies 1 and 2 provide evidence that perceptions of unnecessary harm influence attitudes toward deception, and may explain why people endorse several seemingly anomalous standards of deception. Study 3 tests the explanatory power of the unnecessary harm framework, relative to alternative theories of deception that have been proposed in psychology and philosophy.

If lay people think like philosophers, it is possible that they would focus on how deception can undermine trust in others' words (Bacon, 1872; Harris, 2013), or individual autonomy (e.g., Bok, 1978), rather than how deception can prevent harm to a target. Furthermore, it is possible that people would not consider the consequences of deception at all, but simply believe they have a categorical duty to be honest (Kant, 1785/1959). Although the inductive study suggests these possibilities are unlikely, the present experiment tests these alternative theories more directly to corroborate the inductive study results. Study 3 also addresses the possibility that communicators make egocentric moral judgments of deception, perceiving deception as more acceptable when it is in their self-interest (e.g., when they will benefit from lying and/or are unlikely to get caught; Bereby-Meyer & Shalvi, 2015; Shalvi et al., 2015).

Study 3 explores the relationship between these alternative theories and the endorsement of deception in two ways. First, we examine how these judgments predict the endorsement of deception generally, regardless of community standard violations. Second, we test whether any of these judgments underlie the relationship between community standard violations and the endorsement of deception, using mediation analyses. It is possible, for example, that lies that prevent unnecessary harm are seen as acceptable not because of the harm they prevent, per se, but rather because they spare the liar from discomfort and conflict (self-interest), they are not perceived as threats to a target's autonomy (because the target could not have acted on or understood the truthful information), or because they are less likely to be detected (because targets are motivated to avoid harmful truths). Study 3 tests these possibilities.

Method

Participants

Two separate samples were recruited to complete this study. The first sample consisted of 136 adults recruited via Amazon Mechanical Turk (43% female; $M_{\text{age}} = 32$ years). The second sample consisted of 142 adults from a U.S. university laboratory sample (61% female; $M_{\text{age}} = 23$ years). Both samples completed an identical survey. There were no effects of sample on the endorsement of deception. Thus, results are reported collapsed across samples (see Supplement 4.2 of the online supplementary materials for results split by sample).

⁴Two models were compared using AIC/BIC fit statistics: the lower the value (a deviation statistic), the better the model. The first model did not correlate the error terms for the mediating variables immediate harm and instrumental value. The second model correlated the error terms. The second model with the correlated errors had lowest AIC/BIC values and was therefore chosen.

Procedure

The survey consisted of three vignettes. In each vignette, a different standard of deception and whether participants judged the ethicality of deception from the perspective of an observer, target, or communicator was manipulated. Community standard violation and perspective were both between-subjects factors. As in the previous study, participants responded to multiple vignettes, but never saw more than one version of the same vignette. The order in which the vignettes were presented was randomized.

Vignettes

Study 3 featured new vignettes, each of which manipulated a different community standard. Although Study 3 only explored a subset of community standards, each standard pertains to a different type of situational antecedent to unnecessary harm. The first vignette manipulates an attribute of the target (i.e., the target's ability to understand the truth). The second vignette manipulates a feature of the information (i.e., whether the target could react to the information). The third vignette manipulates the context of the conversation (i.e., whether others were present).

Ability to Understand. In the first vignette, participants had to decide whether or not to inform a target that his daughter had died. The vignette manipulated whether or not the target suffered from dementia. This vignette corresponds with the standard: Lie to targets that cannot understand the truth (Standard 2, Table 2) and mirrors the medical ethics concept of "therapeutic fibbing" (Beach & Kramer, 1999). The exact vignette appears below.

Imagine a caregiver at a nursing home. The caregiver is responsible for Jeff, a 93-year-old man.

Control condition: *Jeff is in good physical and mental health.*

Violation condition: *Although Jeff is in good physical health, he suffers from severe dementia. This means that he often cannot make sense of his reality and is easily confused.*

The caregiver recently learned that Jeff's estranged daughter, who he has not heard from for over a decade, died 2 years ago.

One day, out of the blue, Jeff asks his caregiver if she has heard anything about his family.

Time to Implement Change. The second vignette depicted an individual who had made an error when writing a article. The vignette manipulated whether the mistake could be corrected. This vignette corresponds with the standard: Lie when honest feedback can no longer be implemented (Standard 6, Table 2):

Imagine a graduate student, Jeff, who is planning to submit a paper for publication. Jeff has poured months into his research and is very proud of the resulting manuscript.

Jeff's friend recently read Jeff's manuscript and noticed a few errors.

Control condition: *Jeff is submitting the final paper tomorrow—after he submits the manuscript he will no longer be able to implement changes.*

Violation condition: *Jeff submitted the paper yesterday—meaning he is no longer able to implement changes.*

Jeff asks his friend what he thought of the manuscript.

Presence of Others. The third vignette depicted an individual who had delivered a presentation poorly. The vignette manipulated whether the opportunity to give feedback occurred in public or private. This vignette corresponds with the standard: Lie when honesty would embarrass the target in front of others (Standard 8, Table 2):

Imagine a summer intern named Jeff, who just delivered his end-of-internship presentation to his office.

Jeff's PowerPoint slides were disorganized and he misspoke several times. Jeff's friend attended the presentation and believed that Jeff's presentation went very poorly. Jeff did not seem to realize that, and it is unclear what other audience members thought.

Control condition: *Immediately after the presentation, in a private space, Jeff asks his friend what he thought of the presentation.*

Violation condition: *Immediately after the presentation, in front of several remaining audience members, Jeff asks his friend what he thought of the presentation.*

Dependent Variables

Endorsement of Deception. After participants read each vignette, participants answered, "In the course of this conversation, which of the following options is the more ethical response?" Participants chose between: "Tell [the individual] the truth" and "Lie to [the individual]." The main dependent variable and the response options were identical across all perspectives. The response options were followed by short descriptions of the relevant truth or lie for each vignette. The exact wording of all response options appears in the Supplement 4.1 of the online supplementary materials.

After participants chose to endorse either deception or honesty, participants answered a series of questions intended to examine the proposed mechanisms and rule out alternatives. All items were measured using seven-point rating scales anchored at 1 = *not at all* and 7 = *extremely*.

Immediate Harm and Instrumental Value. Participants responded to four items about the immediate harm of honesty ($\alpha = .80$): "To what extent would honesty cause pain to [the target] at the moment of communication?," "To what extent would telling a lie protect [the target's] feelings at the moment of communication?," "To what extent would honesty cause harm to [the target] at the moment of communication?," and "To what extent would lying benefit [the target] at the moment of communication?"

Participants also responded to four items about the instrumental value of truth ($\alpha = .83$): "To what extent would telling the truth in this vignette have the potential to influence [the target's] future behavior?," "To what extent would telling the truth in this vignette be valuable for [the target's] long-term well-being?," "To what extent is the honest information necessary for [the target] to know?," and "To what extent is the honest information useful for

Table 5
Vignette-Level Analyses in Study 3

Independent variables	Dependent variable: Endorsement of deception; 1 = lie, 0 = tell the truth		
	Vignette		
	Ability to understand	Time to Implement	Presence of Others
Community standard violation ^a	.98*	2.07**	2.49***
Communicator ^b	-.12	-.72	1.40**
Target ^c	.36	-.72	.92
Communicator × Community Standard Violation	.18	.96	-1.57*
Target × Community Standard Violation	-.12	.60	-1.65**
Constant	-1.21***	-3.11***	-2.61***
Observations	278	278	278
R ²	0.05	0.16	0.09

Note. Significant values bolded.

^a Violation is coded as 1 = community standard violation, 0 = no standard violation. ^b Communicator is coded as 0 = target or observer perspective, 1 = communicator perspective. ^c Target is coded as 0 = liar or observer perspective, 1 = target perspective.

* $p \leq .05$. ** $p \leq .01$. *** $p < .001$.

[the target's] learning, growth or enlightenment?" These items were adapted from Study 2.

Alternative Mechanisms

Moral Duty. Participants responded to a single item about moral duty: "To what extent does [the potential liar] have a moral duty to tell the truth?"

Societal Harm. Participants responded to a single item about the degree to which lying could cause societal harm: "To what extent might telling this lie harm society as a whole?"

Perceived Autonomy Violation. Participants responded to two items about the degree to which lying violated the target's autonomy ($r = .46$): "To what extent does this lie infringe upon [the target's] autonomy?" and "To what extent does telling this lie prevent [the target] from making informed decisions?"⁵

Self-Interest. Participants responded to two items about the degree to which lying benefited the liar ($r = .59$): "To what extent is lying the easiest course of action for [the potential liar]?" and "To what extent does lying spare [the potential liar] from conflict?"

Probability of Detection. Finally, participants responded to two items about the degree to which lying could ever be discovered ($r = .26$): "To what extent is the honest information verifiable?" and "To what extent is it possible for [the target] to independently uncover the truth?"

Participants also answered a single-item manipulation check, which asked about the relevant community standard. There was a significant effect of community standard violation on the manipulation check in every vignette ($ps < .01$). After participants submitted their responses, they provided demographic information.

Results

Analytical Approach

As in Study 2, each vignette was first analyzed independently. For each vignette, a set of logistic regressions were conducted, to examine the effects of *Implicit Rule Violation* and *Perspective* on the endorsement of deception (1 = lying is endorsed, 0 = truth-telling is endorsed). In these regressions, *Implicit Rule Violation* was coded as 1 if the relevant rule was violated (e.g., if the target was

not able to understand the information and 0 otherwise). Two dummy variables for the Perspective conditions were also created (Target = 1 in the Target perspective condition and 0 otherwise; Communicator = 1 in the Communicator perspective and 0 otherwise; the Observer perspective served as the control). Then, analyses using the data from all scenarios were conducted, to test the proposed theory and examine alternative mechanisms.

Vignette-Level Results

The results of the vignette-level logistic regressions appear in Table 5. Figure 4 also depicts the proportion of participants who endorsed deception in each experimental condition in each vignette.

In each vignette, there was a significant effect of community standard violation ($ps < .05$). In the *ability to understand* and *time to implement* vignettes, there were no main or interaction effects of perspective ($ps > .44$). In other words, participants responded to community standard violations similarly if they considered the situation from the perspective of a liar, target, or an observer.

In the *presence of others* vignette, there was a significant, but unpredicted, perspective effects (see Table 5, Column 3). Communicators believed that lying was more ethical than observers did ($b = 1.40$, $p = .04$). There were also significant Communicator × Community Standard Violation ($b = -1.57$, $p = .05$) and Target × Community Standard Violation ($b = -1.65$, $p = .05$) interactions; the community standard violation had a weaker effect on communicators and targets than observers. These results suggest that observers may be more attuned to the value of deception in public contexts.

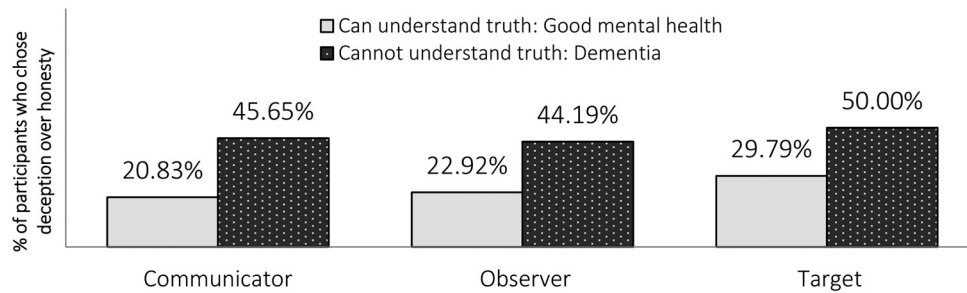
Pooled Results

Next, an analysis using the data from all scenarios was conducted, to test the proposed theory and rule out alternative mechanisms. A series of mixed effects logit models were run (using the *melogit* function in Stata) on the endorsement of deception (1 = lying is endorsed, 0 = truth-telling is endorsed) including community

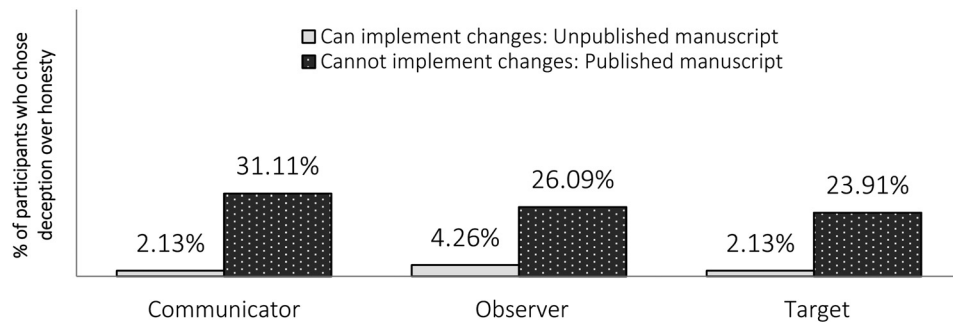
⁵ Although perceptions of autonomy are closely related to perceptions of the instrumental value of truth, these two constructs loaded on distinct factors in an exploratory factor analysis (principal axis factoring, varimax rotation).

Figure 4
Community Standard Violations Across Perspectives (Study 3)

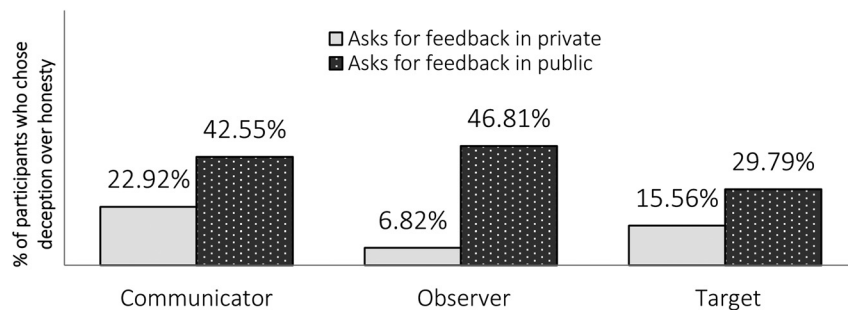
Panel 1. Dementia and a daughter's death (Ability to understand vignette)



Panel 2. Errors in a published versus unpublished manuscript (Time to implement vignette)



Panel 3. Public versus private feedback on a presentation (The presence of others vignette)



Note. In Study 3, participants in all perspectives (Communicators, Observers, and Targets) were more likely to endorse deception when a community standard was violated.

standard violation, perspective, gender, age, and the seven mechanism measures as independent variables (see Table 6). As in Studies 1 and 2, a fixed-effects approach was used to control for vignette, and a random-effects approach to account for multiple observations per participant. Pseudo- R^2 was calculated using the method described in Tjur (2009).⁶

Community standard violations powerfully influence the endorsement of deception (all $bs > .84$, $ps < .001$, Models 2–7), whereas perspective matters much less. As hypothesized, perceptions of immediate harm and instrumental value also significantly influence the endorsement of deception (particularly when the

Immediate Harm \times Instrumental Value interaction is not included, $ps < .001$, Models 5–6, and 9–10). There was not a significant interaction between immediate harm and instrumental value (see Models 7–8). Of the alternative mechanisms examined, only perceptions of moral duty significantly impacted the endorsement of

⁶ If the LR test from the multilevel model was significant ($p < .05$), the random (multilevel) model was used to account for multiple observations per participant. This was the case for all the logit models in Study 3, except for one, in which case a clustered standard error approach was used, rather than the random model. See Table 6.

Table 6
Pooled Results Across Vignettes in Study 3

Independent variables	Dependent variable = Endorsement of deception; 1 = lie, 0 = tell the truth									
	Model number									
	1	2	3	4	5	6	7	8	9	10
Community standard violation ^a		1.47***	1.48***	1.80***		.86***		.84**		1.02***
Communicator ^b			0.16	0.39						
Target ^c			<0.01	0.40						
Liar × Community Standard Violation				−0.35						
Target × Community Standard Violation				−0.62						
Immediate harm of truth					1.05***	1.03**	0.48	0.54	1.06***	1.02**
Instrumental value of truth					−1.14***	−1.08***	−1.87***	−1.71***	−.74***	−.67**
Imm Harm × InstrValue							0.13 ⁺	0.11		
Self-interest									0.12	0.12
Autonomy									−0.11	−.10
Probability of detection									−0.10	−.06
Societal harm									0.01	−0.01
Moral duty									−.74***	−.77***
Gender ^d	0.50**									
Age	−0.02*									
Constant	−1.60***	−2.78***	−2.84***	−3.06***	−1.00	−1.82*	2.14	0.90	0.67	−.30
Vignette fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Participant random effect	No ^e	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	834	834	834	834	834	834	834	834	834	834
R ²	0.05	0.18	0.18	0.18	0.46	0.48	0.47	0.49	0.55	0.55

Note. Significant values bolded.

^a Violation is coded as 1 = community standard violation, 0 = no standard violation. ^b Communicator is coded as 0 = target or observer perspective, 1 = communicator perspective. ^c Target is coded as 0 = communicator or observer perspective, 1 = target perspective. ^d Gender is coded as 1 = female, 0 = male. ^e If the LR test from the multilevel model is significant ($p < .05$), this suggests that the random effect is significant and a random model should be used, in lieu of clustering standard errors. This was the case for all the logit models in study 3, except model 1. Therefore, in model 1, I clustered standard errors at the participant level, rather than using a random-effects approach.

⁺ $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p < .001$.

deception ($ps < .001$, Models 9–10), controlling for all other mechanisms. In other words, the degree to which lying could be easily detected, served the liar's interests, violated the target's autonomy, or could lead to societal harm, did not influence the endorsement of deception after controlling for community standard violations and perceptions of unnecessary harm.

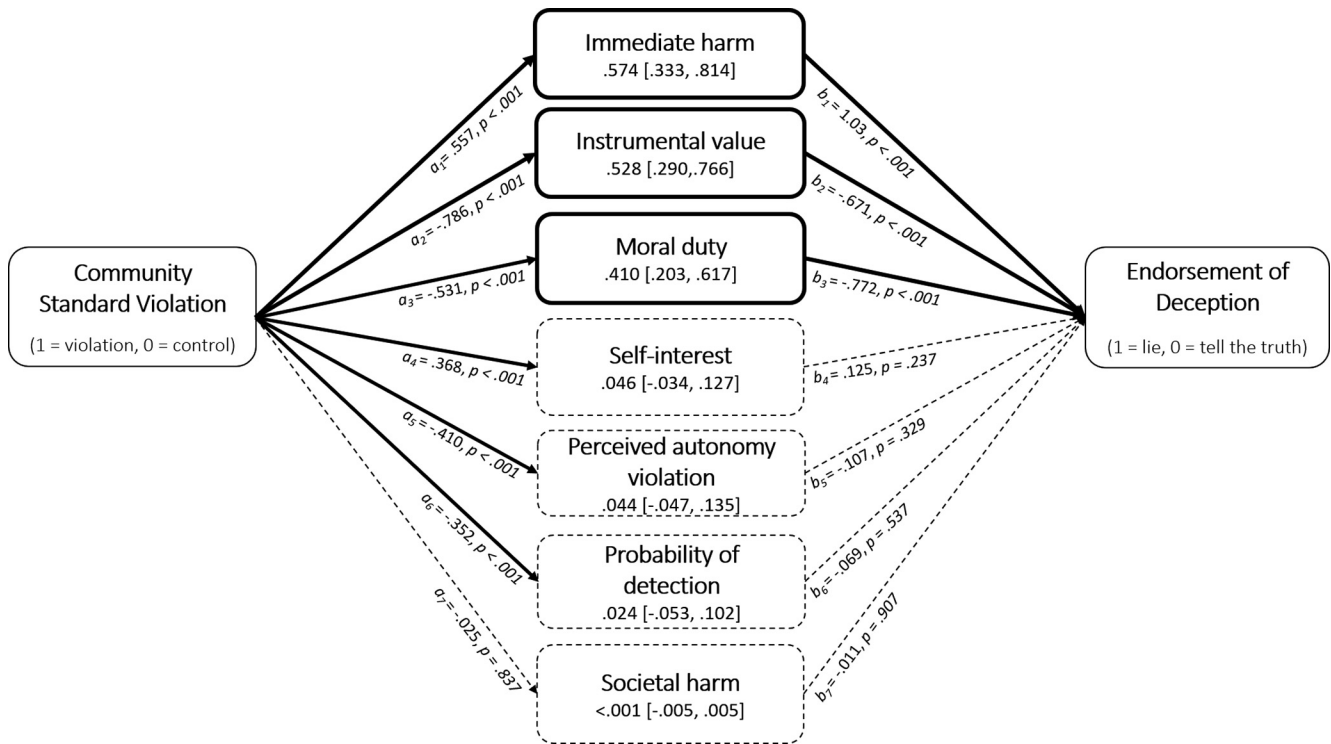
Although the majority of participants did not endorse deception when a particular community standard was violated in this study (see Figure 4), the majority of participants did endorse deception when they rated the truth as high in immediate harm and low in instrumental value, consistent with the unnecessary harm framework. Specifically, 75.5% of participants (significantly greater than 50%; $p < .001$) endorsed deception, across scenarios, when they rated the truth as high in immediate harm (higher than 4, the midpoint of the scale) and low in instrumental value (lower than 4, the midpoint of the scale). In contrast, only 26.1% of participants endorsed deception when they rated the truth as low in immediate harm and low in instrumental value, 2.3% of participants endorsed deception when they rated the truth as low in immediate harm and high in instrumental value, and 23.5% of participants endorsed deception when they rated the truth as high in immediate harm and high in instrumental value (all proportions significantly lower than 50%; $ps < .022$).

Mediation Analyses

In addition to the logistic regressions, mediation analyses were conducted using Stata's gsem function. The mediation model included community standard violation as the independent variable, the seven potential mechanisms as simultaneous mediators, and the endorsement of deception as the dependent measure. A random intercept was included in each equation at the participant level to account for multiple observations per participant, and fixed effects were included to control for scenario and perspective. There was significant evidence of mediation through both proposed mechanisms: immediate harm and instrumental value. There was also significant mediation through perceptions of moral duty.⁷ However, the direction of the effect does not echo philosophical assumptions about one's duty to tell the truth (e.g., Kant, 1785/1959). Lay people do not believe that they have a categorical

⁷ One potential concern with this analysis is that the moral duty item is too conceptually similar to the preference for deception. Therefore, the mediation analysis was rerun without this item. Although the significance of some of the effects did change, perceptions of immediate harm and instrumental value continued to have significantly larger indirect effects than any other potential mechanisms. See the Supplement 4.3 of the online supplementary materials for details.

Figure 5
Mediation Analysis in Study 3



Note. Numbers for a pathways denote the β coefficients from regressions on each mechanism, using community standard violation as a predictor. Numbers for b pathways denote the β coefficients from regressions on the endorsement of deception, using community standard violation and all mechanisms as predictors. Numbers in the center boxes reflect the point estimate for the indirect effect and the 95% confidence interval around the indirect effect for each proposed mechanism. Bold, solid lines signify significant pathways; dotted lines signify insignificant pathways. Consistent with the unnecessary harm framework, judgments of the immediate harm and instrumental value of truth mediated the effect of community standards violations on the endorsement of deception.

imperative to tell the truth. Rather, they believe that when a community standard is violated, they have *less* duty to tell the truth, leading them to endorse deception. Figure 5 depicts the theoretical mediation model, the coefficients on all a and b pathways, as well as the indirect effects.

Discussion

Study 3 suggests that perceptions of the immediate harm and the instrumental value of truth, the two hypothesized dimensions of unnecessary harm, underlie the effects of community standard violations on the endorsement of deception. Perceptions of autonomy, societal harm, self-interest, and the probability of detection do not.

It is important to note, however, that some of the alternative mechanisms investigated here are influenced by community standard violations (see a and b pathways in Figure 5). For example, lying when an community standard is violated is perceived to be less of an autonomy violation ($a_5 = -.409$, $p < .001$), harder to detect ($a_6 = -.356$, $p < .001$) and more in the communicator's interest ($a_4 = .370$, $p = .001$) than lying when an community standard is not violated. However, these judgments do not influence the endorsement of deception, particularly after controlling for perceptions of unnecessary harm (see b pathways, all $ps > .187$).

Alternatively, perceptions of unnecessary harm to the target do independently influence the endorsement of deception, above and beyond the effects of all other mechanisms investigated.

Study 4: Perceptions of Unnecessary Harm in Everyday Life

Studies 4a and 4b examine how perceptions of unnecessary harm influence the lies people tell in their everyday conversations using representative U.S. samples. Study 4a examines communicator judgments and Study 4b examines third-party judgments. These studies also explore how the dimensions of unnecessary harm influence the use and judgment of deception, above and beyond alternative mechanisms, as in study 3.

Method

Study 4a and 4b were nearly identical, except that participants in Study 4a reported lies (or truths) they recently told, and participants in Study 4b judged the lies (or truths) reported by Study 4a participants. For each study, 300 adults from a U.S. representative sample were recruited via Prolific Academic (<https://www.prolific.co/>). We ended up with a final sample of 296 participants

who completed Study 4a (50% female; $M_{\text{age}} = 45$ years; $M_{\text{work experience}} = 23$ years), and 285 participants who completed Study 4b (50% female; $M_{\text{age}} = 44$ years; $M_{\text{work experience}} = 21$ years). Study 4b was preregistered at aspredicted.org (<https://aspredicted.org/yw6z2.pdf>).

Study 4a

Participants in Study 4a first read a passage explaining that people often face situations in which they are tempted to lie. Then, participants were asked to, "Think of the last time you struggled with the decision to tell the truth. That is, think of the most recent time that you considered lying or were tempted to lie, and either ended up lying or telling the truth." Participants then described the situation using a free response text box.

After describing the situation, participants wrote the name of the person they had considered lying to and indicated the nature of their relationship with this person (choices: friend, family member, coworker, stranger, significant other, roommate, other).⁸ Then, participants indicated whether they ultimately lied or told the truth in the situation they described. Participants then responded to a series of scale measures.

Study 4b

Each participant in Study 4b read 10 randomly-selected situations described by participants in Study 4a (i.e., the text response describing a situation in which the participant was tempted to lie).⁹ Then participants rated the situation on the same measures as participants in Study 4a (see below).

Dependent Variables

Moral Judgment of Deception. Participants indicated how ethical it was to lie in the situation they [the participant in study 4a] described (1 = *completely unethical*, 7 = *completely ethical*).

Immediate Harm and Instrumental Value. Next, participants responded to potential mechanism items. Participants responded to the same four items about the immediate harm of honesty used in study 3, plus an additional item, "To what extent would telling the truth in this situation cause unnecessary harm to [the target]" (1 = *not at all*, 7 = *extremely*; $\alpha = .89$). Participants also responded to the same four items about the instrumental value of truth used in Study 3 ($\alpha = .75$).

Alternative Mechanisms

Self-Interest. Participants responded to the same two items about the degree to which lying benefited the liar as they had in Study 3, plus the item "To what extent would lying in this situation benefit you?" (1 = *not at all*, 7 = *extremely*; $\alpha = .68$).

Societal Harm. Participants responded to the same item about the degree to which lying could cause societal harm as they had in Study 3, plus an additional item: "If everyone lied in situations like this, to what extent would this harm societal trust?" (1 = *not at all*, 7 = *extremely*; $r = .485$, $p < .001$). This latter item better reflects philosophical arguments about the deleterious effects of any lie on society and trust in language (i.e., the Principle of Veracity, Bok, 1978).

Perceived Autonomy Violation. Participants responded to a single item about the degree to which lying violated the target's

autonomy: "To what extent does this lie infringe upon [the target's] autonomy?" (1 = *not at all*, 7 = *extremely*).

Probability of Detection. Finally, participants responded to a single item about the degree to which lying could ever be discovered: "How likely is it that you would be caught lying in this situation?" (1 = *not at all*, 7 = *extremely*).

After participants submitted their responses, they provided demographic information.¹⁰

Results

In Study 4a, a total of 56.4% of participants described situations in which they ended up lying. Table 7 depicts examples of lies that were perceived to be ethical (above the midpoint of 4), and that were rated to be high in immediate harm (above the midpoint of 4) and low in instrumental value (below the midpoint of 4). The descriptive statistics of all scale measures split by the decision to lie are presented in the Supplement 5.2 of the online supplementary materials.

Analytical Approach

The purpose of this study was to explore how perceptions of unnecessary harm influence the decision to lie (Study 4a only) and the moral judgment of deception (Studies 4a and 4b), above and beyond other variables associated with the ethics of deception. Therefore, a series of regressions (logistic regressions predicting the decision to lie, OLS regressions predicting moral judgments) using perceptions of unnecessary harm, perceptions of self-interest, societal harm, autonomy violation, and probability of detection as independent variables were run. Tables 8 and 9 depict the results.

Decision to Lie (Study 4a)

In Study 4a, perceptions of immediate harm and instrumental value significantly predicted the decision to lie (see Model 1 in Table 8; $ps < .001$). There was not a significant Immediate Harm \times Instrumental Value interaction; however, including the interaction in the model did reduce the predictive value of immediate harm, as it did in Studies 2 and 3 (see Models 2 and 4 in Table 8).

Self-interest and perceptions of societal harm also influenced the decision to lie in Study 4a. However, after controlling for these judgments, perceptions of immediate harm and instrumental value still significantly predicted the decision to lie (see Model 3), though only when the Immediate Harm \times Instrumental Value interaction was not included in the model (see Model 4).

⁸ See the Supplement 5.3 of the online supplementary materials for analyses split by relationship-type and a discussion of how deception differed across relationships.

⁹ The preregistration had a typo, indicating that each participant would rate 20, rather than 10, situations described by participants in Study 4a.

¹⁰ Participants in Study 4a also answered questions about the degree to which they believed the target would be angry and would believe the liar had good intentions if they discovered the lie, as well as three exploratory items about their communication medium and the degree to which they planned their lie in advance. See the Supplement 5.1 of the online supplementary materials for a discussion of these measures and their results.

Table 7*Examples of Lies Told in Everyday Life That Are Perceived to Be High in Unnecessary Harm, Study 4a*

Participant response	Participant ratings		
	Immediate harm	Instrumental value	Ethical
I was depressed and told my parents about it because I was doing bad in school and I ended up telling them I was depressed, hence my bad grades and I regret it because I didn't want anybody to know I was depressed, even my parents. My mom ended up getting depressed from it too, thus I regret my actions further. I was taking pills to make myself feel better but I didn't want to rely on pills and I told my mom I was feeling better, and I genuinely did feel better. Months later I was feeling depressed again and my mother asked me how I was doing, because she knows that depressive thoughts can come back and I told her I was happy, which was a lie because around that time I was feeling depressed again.	5.00	1.00	7.00
My cousin and I were talking last night about our families and how much we missed our fathers that had passed on. In her eyes, her father was the greatest person she had ever known. She never knew the things that I knew about him and some of the things that I personally knew he had done. That would break her heart. Some of those things were to me. There was a part of me that wanted to let her know who he really was. But I thought, Why should I destroy the image she had of him. It would change nothing. She still believes he was the perfect father!! . . .	5.80	2.50	7.00
My sister in law looked horrible in her wedding dress. She was obviously happy with her choice and when she asked me what I thought of the dress I figured she was just asking for reassurance and a self-esteem boost. I told her she looked amazing. She felt amazing and that was all that really mattered.	5.80	1.00	7.00
My sister is an animal lover and cares very much for every species. She rescued some orphaned possums from a hazardous situation and released them into an area where they would better survive (an orchard behind our neighborhood where there are other possums, raccoons, etc.). One of them left the orchard and went into our neighbor's yard where it was killed by a dog. The neighbor told me about it but we agreed not to tell my sister as she would have been devastated and felt responsible for having released the possums in the vicinity. By withholding that information, I wasn't telling a lie per se but by not apprising her of the truth, it felt like one.	6.80	2.50	6.00

Note. Participants in Study 4a recalled instances in which they struggled to tell the truth and ultimately lied or told the truth. Table 7 depicts examples in which participants lied and believed doing so was ethical. Participants' lies are accompanied by their own ratings of immediate harm, instrumental value, and the ethicality of lying in the situation described.

Moral Judgment of Deception (Studies 4a and 4b)

Moral judgments of deception showed similar patterns. Among communicators (Study 4a) and third parties (Study 4b), perceptions of immediate harm and instrumental value significantly predicted the belief that deception was ethical (see Models A1 and B1 in Table 9; p s < .001). I found evidence for a significant Immediate Harm \times Instrumental Value interaction among communicators (Study 4a; see Models A2 and A4 in Table 9), but not third parties (Study 4b; see Models B2 and B4 in Table 9). In Study 4a, I found that including the interaction in the model reduced the predictive value of immediate harm on the judgment of deception, as was the case for the decision to lie (see Models A2 and A4 in Table 9). This was not the case in Study 4b.

Perceptions of societal harm also influenced the moral judgment of deception among communicators (Study 4a) and third parties (Study 4b). Surprisingly, third parties, but not communicators, also judged lies to be more ethical when they promoted the communicator's self-interest. Importantly, however, perceptions of unnecessary harm still predicted the moral judgment of deception, controlling for judgments of societal harm, though the nature of these relationships (i.e., whether immediate harm or instrumental value remained significant) varied across models.

Mean Ratings Among Communicators (Study 4a) Versus Third Parties (Study 4b)

Next, we examine whether communicators' judgments were miscalibrated, relative to third-party judgments. *T*-tests were used to compare

the mean level of ratings provided by third parties and communicators (see Table 10). Third-party raters in Study 4b judged the truth as having greater instrumental value (p < .001), and perceived deception as causing greater societal harm (p < .001), as representing a greater autonomy violation (p < .001), as being more likely to be detected (p < .001), and as being less in the communicator's self-interest (p = .025), relative to communicators in Study 4a. These results suggest that communicators may systematically underestimate the instrumental value of truth and the degree to which their deception will be detected, as well as the societal harm and the autonomy damages caused by deception. However, it is not clear whether these underestimations bias communicators' judgments of ethicality. Third parties' and communicators' moral judgments did not differ, and perceptions of immediate harm, instrumental value and societal harm generally influenced moral judgments in the same manner across these samples (see Tables 8 and 9).

Discussion

Study 4 provides evidence that perceptions of unnecessary harm influence the judgment and use of deception in everyday life. Notably, the types of conversations participants evaluated were not constrained in this study. Participants considered a broad range of lies, including lies of convenience, self-interest, and prosociality (which has been the focal context thus far). Perhaps not surprisingly, then, perceptions of unnecessary harm were not always the strongest predictors of communicators' deception, relative to alternative mechanisms (particularly self-interest and societal harm). Nonetheless,

Table 8*Logistic Regression on Choice to Lie (1 = Lie, 0 = Truth), Study 4a*

Independent variables	Model number			
	1	2	3	4
Immediate harm	0.366***	0.338*	0.299***	0.259
Instrumental value	-0.490***	-0.514**	-0.362***	-0.397*
Immediate Harm × Instrumental Value		0.007		0.011
Self-interest			0.255**	0.256**
Societal harm			-0.330**	-0.329**
Perceived autonomy violation			0.123	0.122
Probability of detection			-0.121	-0.119
Constant	0.902*	0.986	0.309	0.421
Observations	296	296	296	296
R ²	0.108	0.108	0.159	0.159

Note. Significant values bolded.

+ $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p < .001$.

perceptions of unnecessary harm continued to influence communicators' moral judgment of deception in everyday life, even after controlling for these alternative mechanisms (Models A3 and B3 in Table 9). Third-party judgments of immediate harm and instrumental value also influenced judgments of ethicality in these settings,

though third-party judges generally believed the truth had more instrumental value, and that deception caused greater societal harm, than did communicators. These results suggest that unnecessary harm plays an important role in real-world deception, and that communicators may misjudge how valuable their truths could be.

Table 9*OLS Regression on Moral Judgment of Deception in Studies 4a and 4b*

Panel A: Study 4A, Communicators				
Independent variables	Model number			
	A1	A2	B3	A4
Immediate harm	0.330***	-0.032	0.297***	-0.063
Instrumental value	-0.279***	-0.630***	-0.095	-0.446**
Immediate Harm × Instrumental Value		0.102**		0.102**
Self-interest			0.070	0.076
Societal harm			-0.337***	-0.335***
Perceived autonomy violation			0.010	-0.002
Probability of detection			-0.084	-0.070
Constant	3.742***	4.937***	3.912***	5.054***
Observations	296	296	296	296
R ²	0.133	0.156	0.203	0.226
Panel B: Study 4B, Third parties				
Independent variables	Model number			
	B1	B2	B3	B4
Immediate harm	0.346***	0.258***	0.334***	0.234**
Instrumental value	-0.447***	-0.509***	-0.281***	-0.351***
Immediate Harm × Instrumental Value		0.020		0.023
Self-interest			0.113***	0.112***
Societal harm			-0.231***	-0.232***
Perceived autonomy violation			-0.027	-0.028
Probability of detection			-0.042	-0.041
Constant	4.637***	4.890***	4.481***	4.771***
Stimuli (Lie from Study 4a) Fixed effect	Yes	Yes	Yes	Yes
Participant random effect	Yes	Yes	Yes	Yes
Observations	2851			
R ²	0.384	0.385	0.420	0.421

Note. Significant values bolded.

** $p \leq .01$. *** $p \leq .001$.

Table 10
Mean Judgments in Studies 4a and 4b

Dependent variables	Study 4a: Communicators		Study 4b: Third parties		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Moral judgment of lie	3.81	1.77	3.69	0.99	1.03	.304
Immediate harm	3.32	1.73	3.47	1.02	1.74	.084
Instrumental value	3.70	1.49	4.30	0.88	7.94	<.001
Self-interest	4.99	1.53	4.80	0.79	2.25	.025
Societal harm	2.54	1.42	3.40	0.85	10.55	<.001
Autonomy violation	2.30	1.61	3.20	0.83	9.45	<.001
Probability of detection	3.04	1.76	3.86	1.03	8.51	<.001

General Discussion

Despite the centrality of honesty and deception to conceptualizations of morality and immorality, little empirical research has carefully investigated when and why lying is judged to be ethical. One inductive study, three choice experiments, 12 vignette experiments, and two correlational studies, unearth community standards of deception: the conditions under which people systematically endorse deception. Deception is perceived to be more ethical than honesty and individuals want to be deceived when honesty is perceived to cause *unnecessary harm*. This research deepens our understanding of deception and morality in important ways.

Contributions to Deception Research

First, this work illuminates how people fundamentally think about deception. Specifically, this work identifies systematic circumstances in which deception is seen as more ethical than honesty, and it provides an organizing framework for understanding these circumstances. A large body of research identifies features of lies that make them seem more or less justifiable and therefore, that lead people to tell greater or fewer lies (e.g., Effron, 2018; Rogers et al., 2017; Shalvi et al., 2011). However, little research addresses whether people, upon introspection, ever actually believe it is *right* to tell lies; that is, whether lying is ever a morally superior strategy to truth-telling. The present research finds that people believe lying is the right thing to do when it prevents unnecessary harm. Notably, this finding reveals that lay people seem to have a relatively pragmatic view of deception and honesty. Rather than believing deception is a categorical vice—for example, because it damages social trust (Bok, 1978; Kant, 1949) or undermines autonomy (Bacon, 1872; Harris, 2013; Kant, 1785/1959)—people seem to conceptualize deception as a tactic that can and should be used to regulate another vice: harm.

Although this view of deception runs counter to prevailing normative claims and much of the existing scholarship in psychology and economics, it is important to note that this idea—that deception is and *should* be used pragmatically—is not novel. In fact, many of the standards of deception identified in the present research are alluded to in other philosophical, religious, and practical discussions of deception (see Table 2 for a review). Until now, however, these ideas have been siloed in disparate literatures, and behavioral scientists have lacked a parsimonious framework for understanding why individuals endorse deception in various circumstances. The present research identifies a common psychology that explains a number of seemingly unrelated “exceptions” to the

norm of honesty, thereby unifying findings and arguments across psychology, religion, and philosophy under a common theoretical framework.

Furthermore, the unnecessary harm framework identifies *the dimensions of harm* that communicators, targets, and observers care about when judging deception. Existing research on deception has focused almost entirely on the “immediate harm” dimension, for example, by demonstrating that people frequently lie to be polite (Brown & Levinson, 1987), to protect others’ feelings (DePaulo & Bell, 1996), to boost others’ egos (Roberts et al., 2020), and to prevent conflicts (DePaulo et al., 1996). However, as the present work demonstrates, most people do not believe that protecting a person from immediate harm is a sufficient justification for lying. Lying when honesty both causes immediate harm and has instrumental value is seen as paternalistic and is not seen as ethical (Lupoli et al., 2018). In contrast, lying when honesty causes immediate harm and does *not* have instrumental value is seen as ethical.

Prior scholarship on deception has not considered whether there are circumstances in which the truth itself is of little value. The present research clarifies that when the truth is not actionable, understandable, or important (i.e., the truth lacks instrumental value), people value truth-telling less, and more readily endorse lying. Although this insight seems obvious in retrospect and is consistent with work on the antecedents of information avoidance (Sweeny et al., 2010), it is not reflected in psychological or philosophical research on deception. Importantly, accounting for the instrumental value of truth helps to explain the prevalence and endorsement of both minor and major lies. Lying about significant events, such as infidelity or death, is perceived to be ethical when the truth is no longer useful or understandable.

Thus, this research also contributes a growing body of research on prosocial and paternalistic lies. Existing work on attitudes toward prosocial lies has primarily operationalized prosocial deception using economic games (e.g., Erat & Gneezy, 2012; Levine & Schweitzer, 2014, 2015). While this methodology has the strength of experimental control, it does not shed light on how people judge lies in everyday life (Levitt & List, 2007). The present research suggests that the prosocial lies examined in past work, which have featured lies about trivial information (e.g., a coin flip) that had clear monetary benefits (e.g., Levine & Schweitzer, 2014), are seen as ethical precisely because the information had no instrumental value, and the lie prevented objective harm. In other words, the present work clarifies *why* the lies studied in past work were rated positively, and explores the boundaries of these effects.

Contributions to Moral Psychology

The present research also contributes to and extends a number of recent developments in moral psychology. Specifically, the finding that moral judgments of deception hinge on beliefs about harm is consistent with a large body of research in moral psychology suggesting that harm is the essence of morality (Gray et al., 2012). Notably, however, existing research on harm and morality has not explicitly explored deception, nor has it distinguished between necessary and unnecessary harm.

In addition, the finding that moral judgments of deception align with preferences for deception is consistent with an emerging body of research on preferences for moral contractualism, which suggests that many moral judgments are driven by the desire to treat others as they wish to be treated (Everett et al., 2016). Consistent with this idea, I find that observers generally believe it is ethical to lie in the same circumstances they would want to be deceived as targets.

This finding bears striking resemblance to Bok's normative standard for deciding whether a lie is ethical, the "Test of Publicity." To pass this test, a lie must be acceptable to the deceived party. Although Bok does concede that some lies may pass this test (e.g., lies of trivial importance, or lies in extreme circumstances), she largely assumes that people rarely—if ever—would consent to being deceived. Thus, in Bok's view, deception is rarely assumed to be ethical. In contrast, this work suggests that many, systematic lies meet the "Test of Publicity" and are therefore perceived to be ethical.

More broadly, this research sheds light on the degree to which different normative theories are useful for understand everyday moral intuitions. The unnecessary harm framework is neither deontological nor strictly utilitarian. In my studies, few individuals held a deontological view of deception, believing that deception is categorically wrong (Table 1, and the Supplement 5.1 of the online supplementary materials for additional evidence). Furthermore, although the proposed framework is notably consequentialist—individuals weigh the short-term harm against the long-term benefits of truth—it is not broadly utilitarian for two reasons.¹¹ First, most individuals focus narrowly on the consequences of lying for the target, rather than the consequences for the liar or society writ large. Second, lying is still seen as unethical when honesty is not harmful or instrumental (i.e., in the absence of consequences for the target). This suggests that the deontology-utilitarian dichotomy may not be the most useful framework to understand the morality of everyday social interaction, despite its popularity in empirical research (Bauman et al., 2014). In contrast, contractualism seems to align with lay beliefs about deception quite well: People believe that the use of both honesty and deception should be constrained by the particular desires of the target of communication.

Limitations and Future Directions

Further Understanding Community Standards of Deception

Of course, this research is not without limitations. Although the inductive study provides an initial set of community standards, future research is needed to establish a complete set. For example, the present theory focuses on deception that is motivated by the

desire to *prevent* harm. However, there may be additional lies that are endorsed based on the benefit they create (e.g., lies that create a surprise). Furthermore, the present research focuses on how people think about deception when communicating negative information that affects others (e.g., when delivering negative news or critical feedback). People may think differently about deception when communicating personal information about themselves (e.g., when keeping secrets engaging in self-disclosure). Future research should explore lay theories of deception as it pertains to personal secrecy (Slepian et al., 2017).

Future research should also explore how the documented standards relate to one another. Notably, the community standards of deception established here are not necessarily independent. Many interesting cases reflect a violation of several standards. For example, falsely telling a bride she is beautiful on her wedding day (see e.g., Table 2 and Table 7) may reflect the desire to avoid disrupting a sacred event (Standard 7), the belief that a person is particularly fragile on their wedding day (Standard 1), and the belief that a person cannot change their wedding dress at the last minute (Standard 6). People may more readily endorse deception in such cases, when several standards are violated. Future research should test this directly.

Further Understanding the Unnecessary Harm Framework

The Relationship Between Immediate Harm and Instrumental Value. The key prediction of the unnecessary harm model is that deception is most likely to be seen as more ethical than honesty when immediate harm is high and instrumental value is low. Although there was consistent support for this prediction across studies, there was not consistent evidence for the nature of this effect. Across studies, there was sometimes a significant Immediate Harm \times Instrumental Value Interaction, and sometimes found two main effects. Furthermore, the nature of the Immediate Harm \times Instrumental Value interaction varied from study to study (when it was significant). For example, among observers in Study 1, immediate harm influenced the endorsement of deception more when instrumental value was low than when it was high. However, in Study 2 (pooled analyses, see the Supplement 3.5 of the online supplementary materials) and in Study 4a, immediate harm influenced the endorsement of deception more when instrumental value was high than when it was low. However, only Study 1 included all levels of immediate harm and instrumental value. The remainder of the studies relied on participants' perceptions of these dimensions when faced with difficult conversations, and therefore judgments were biased toward high levels. More research is needed to systematically test the Immediate Harm \times Instrumental Value relationship.

Disentangling Perceptions of Harm From Perceptions of Intent. Future research should also disentangle perceptions of

¹¹ Notably, in the present research there is little evidence to suggest that people use utilitarian reasoning to support a deontological stance on deception. A priori, it would have reasonable to predict that people believe (or are motivated to believe) that deception carries large costs (e.g., to society and the autonomy of the target), and that these costs are never outweighed by the benefit of any single lie, leading to a deontological stance. Such beliefs would be consistent with recent research on moral intuitionism (Ditto & Liu, 2011) and could explain why deception is seen as unethical even in the absence of concrete consequences (Gray et al., 2014). However, given that people focused on the consequences of lying for the target when deciding its ethicality, it seems unlikely that such reasoning plays a central role in explaining specific acts of deception.

harm from perceptions of intent. A growing body of research on the importance of mind perception (e.g., Gray et al., 2012) suggests that people's approval of lies that prevent unnecessary harm may be driven by their beliefs about the character and intentions of a communicator who would tell such a lie, rather than the harm the lie prevents, per se. Consistent with this possibility, a supplementary study (see S3 in the Supplement 6.3 of the online supplementary materials) finds that perceptions of unnecessary harm also influence targets' beliefs about a liar's intentions, and that perceptions of the liar's intentions are significantly correlated with moral judgments ($r = .481, p < .01$).

However, experiments that orthogonally manipulate a liar's intentions and the degree to which the lie successfully prevents unnecessary harm are needed to precisely test whether intentions or harm are driving moral judgment. Imagine a communicator who employs the logic of unnecessary harm, but nevertheless ends up telling a lie that harms the target. Will this person be seen as ethical? In the case of everyday deception, good intentions may not be sufficient for garnering moral credit. The degree to which the liar's intentions align with the target's beliefs are also likely to influence judgments of deception (Lupoli et al., 2018). Paternalistic lies that are motivated by good intentions—even the intention to prevent unnecessary harm—may backfire if the cues a liar uses to make judgments of harm are seen as biased. For example, a well-intend person who deceives a woman because he believes women are fragile, is unlikely to be lauded (at least by women).

The existence of mutually agreed upon principles does not imply the correct use of principles. People (communicators, targets, and third parties) believe that deception prevents unnecessary harm when the target lacks cognitive capacity, is emotionally fragile, or cannot implement feedback. However, communicators and targets may not necessarily make identical assessments about the presence or absence of these circumstances. A communicator, for instance, may be motivated to believe that a target is less competent or more fragile than he or she really is, as in the example above. This may be particularly problematic for populations that, due to interpersonal biases, are perceived as especially fragile or incompetent (e.g., Jampol & Zayas, 2020). Results from Studies 4a and 4b corroborate this possibility: Communicators underestimated the instrumental value of truth, relative to third-party judges, suggesting that communicators may have been motivated to see the truth as less valuable.

Additional Predictions of the Unnecessary Harm Framework. The present research also points to a number of interesting—and testable—predictions for future research on honesty and deception. Specifically, the two dimensions of unnecessary harm (immediate harm and instrumental value) differ in important ways, which might lead communicators and targets to prioritize them at different points in time. Immediate harm is characterized by an immediate emotional response to a specific situation (e.g., how will this person feel as a result of hearing this information right now?), whereas instrumental benefit is characterized by a cognitive consideration of more long-term and abstract consequences (e.g., how and what will a person learn from this information?). Emotional responses loom large when individuals react to situations in the moment whereas cognitive calculations loom large when a person is deliberating about their future actions (e.g., Loewenstein & Small, 2007; Monin et al., 2007). As a result, individuals may overweight the immediate harm of honesty in the heat of conversation (e.g., Levine & Cohen, 2018), but when thinking about a potential

conversation from a distance, individuals may be more attuned to the potential long-term benefits of honesty. Thus, communicators may intend to be honest (and targets may expect to appreciate honesty) when they consider having an unpleasant conversation, but when the moment to inflict (or experience) pain actually comes, they may prefer deception. Future research should build on the proposed two-dimensional model to more deeply understand how people make trade-offs when communicating, and the consequences thereof.

Conclusion

Deception is typically characterized as unethical, and existing research assumes that individuals would rarely consent to being deceived. In contrast to these views, the present research demonstrates that individuals frequently consent to and morally justify deception, and they do so in systematic ways. Deception is seen as ethical when it prevents unnecessary harm.

References

- Bacon, F. (1872). *The letters and the life of Francis Bacon* Longmans, Green and Company.
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, 8(9), 536–554. <https://doi.org/10.1111/spc3.12131>
- Bazerman, M. H., & Tenbrunsel, A. E. (2012). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton University Press.
- Beach, D. L., & Kramer, B. J. (1999). Communicating with the Alzheimer's resident: Perceptions of care providers in a residential facility. *Journal of Gerontological Social Work*, 32(3), 5–26. https://doi.org/10.1300/J083v32n03_02
- Bereby-Meyer, Y., & Shalvi, S. (2015). Deliberate honesty. *Current Opinion in Psychology*, 6, 195–198. <https://doi.org/10.1016/j.copsyc.2015.09.004>
- Blanton, B. (2005). *Radical honesty: How to transform your life by telling the truth*. Sparrowhawk Publishing.
- Bok, S. (1978). *Lying: Moral choices in public and private life*. Pantheon.
- Boles, T. L., Croson, R. T., & Murnighan, J. K. (2000). Deception and retribution in repeated ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, 83(2), 235–259. <https://doi.org/10.1006/obhd.2000.2908>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge University Press.
- Croson, R., Boles, T., & Murnighan, J. K. (2003). Cheap talk in bargaining experiments: Lying and threats in ultimatum games. *Journal of Economic Behavior & Organization*, 51(2), 143–159. [https://doi.org/10.1016/S0167-2681\(02\)00092-6](https://doi.org/10.1016/S0167-2681(02)00092-6)
- Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, 12(1), 2–7. <https://doi.org/10.1037/a0025071>
- Dalio, R. (2017). *Principles: Life and work*. Simon and Schuster.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70(5), 979–995. <https://doi.org/10.1037/0022-3514.70.5.979>
- DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told to those who care. *Journal of Personality and Social Psychology*, 71(4), 703–716. <https://doi.org/10.1037/0022-3514.71.4.703>
- DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, 74(1), 63–79. <https://doi.org/10.1037/0022-3514.74.1.63>
- Ditto, P. H., & Liu, B. (2011). Deontological dissonance and the consequentialist crutch. In M. Mikulincer & P. R. Shaver (Eds.), *The social*

- psychology of morality: *Exploring the causes of good and evil* (pp. 51–70). APA Press.
- Effron, D. A. (2018). It could have been true: How counterfactual thoughts reduce condemnation of falsehoods and increase political polarization. *Personality and Social Psychology Bulletin*, 44(5), 729–745. <https://doi.org/10.1177/0146167217746152>
- Effron, D. A., Bryan, C. J., & Murnighan, J. K. (2015). Cheating at the end to avoid regret. *Journal of Personality and Social Psychology*, 109(3), 395–414. <https://doi.org/10.1037/pspa0000026>
- Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, 58(4), 723–733. <https://doi.org/10.1287/mnsc.1110.1449>
- Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, 145(6), 772–787. <https://doi.org/10.1037/xge0000165>
- Gino, F., Ayala, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20(3), 393–398. <https://doi.org/10.1111/j.1467-9280.2009.02306.x>
- Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, 115(2), 191–203. <https://doi.org/10.1016/j.obhdp.2011.03.001>
- Gino, F., & Pierce, L. (2009). Dishonesty in the name of equity. *Psychological Science*, 20(9), 1153–1160. <https://doi.org/10.1111/j.1467-9280.2009.02421.x>
- Gino, F., & Pierce, L. (2010). Lying to level the playing field: Why people may dishonestly help or hurt others to create equity. *Journal of Business Ethics*, 95(S1), 89–103. <https://doi.org/10.1007/s10551-011-0792-2>
- Gino, F., & Shea, C. (2012). Deception in negotiations. *The Oxford Handbook of Economic Conflict Resolution*, 2012, 47–60. <https://doi.org/10.1093/oxfordhb/9780199730858.013.0005>
- Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1), 384–394. <https://doi.org/10.1257/0002828053828662>
- Goffman, E. (1955). On face-work; an analysis of ritual elements in social interaction. *Psychiatry*, 18(3), 213–231. <https://doi.org/10.1080/00332747.1955.11023008>
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55–130.
- Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, 143(4), 1600–1615. <https://doi.org/10.1037/a0036149>
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124. <https://doi.org/10.1080/1047840X.2012.651387>
- Gray, K., & Wegner, D. M. (2010). Torture and judgments of guilt. *Journal of Experimental Social Psychology*, 46(1), 233–235. <https://doi.org/10.1016/j.jesp.2009.10.003>
- Gray, K., & Wegner, D. M. (2011). Morality takes two: Dyadic morality and mind perception. In M. Mikulincer & P. R. Shaver (Eds.), *Herzliya series on personality and social psychology. The social psychology of morality: Exploring the causes of good and evil* (pp. 109–127). American Psychological Association. <https://doi.org/10.1037/13091-006>
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108. <https://doi.org/10.1126/science.1062872>
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6(12), 517–523. [https://doi.org/10.1016/S1364-6613\(02\)02011-9](https://doi.org/10.1016/S1364-6613(02)02011-9)
- Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65(4), 613–628. <https://doi.org/10.1037/0022-3514.65.4.613>
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20(1), 98–116. <https://doi.org/10.1007/s11211-007-0034-z>
- Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, 31(1), 191–221. <https://doi.org/10.1111/j.1559-1816.2001.tb02489.x>
- Harris, S. (2013). *Lying*. Four Elephants Press.
- Helzer, E. G., & Pizarro, D. A. (2011). Dirty liberals! Reminders of physical cleanliness influence moral and political attitudes. *Psychological Science*, 22(4), 517–522. <https://doi.org/10.1177/0956797611402514>
- Hildreth, J. A. D., Gino, F., & Bazerman, M. (2016). Blind loyalty? When group loyalty makes us see evil or engage in it. *Organizational Behavior and Human Decision Processes*, 132(1), 16–36. <https://doi.org/10.1016/j.obhdp.2015.10.001>
- Ho, D. Y. F. (1976). On the concept of face. *American Journal of Sociology*, 81(4), 867–884. <https://doi.org/10.1086/226145>
- Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science*, 345(6202), 1340–1343. <https://doi.org/10.1126/science.1251560>
- Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, 10(4), 15–41. <https://doi.org/10.1080/00208825.1980.11656300>
- Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. MIT Press.
- Iyer, R. (2010). *The case for honesty as a moral foundation*. <http://www.yourmorals.org/blog/2010/12/the-case-for-honesty-as-a-moral-foundation/>
- Jampol, L., & Zayas, V. (2020). Gendered white lies: Women are given inflated performance feedback compared with men. *Personality and Social Psychology Bulletin*, 47(1), 57–69.
- Jowett, B., & Campbell, L. (Eds.). (1894). *Plato's Republic: The Greek text* (Vol. 2). Clarendon Press.
- Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, 125(2), 131–164. <https://doi.org/10.1037/rev0000093>
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986a). Fairness and the assumptions of economics. *The Journal of Business*, 59(4), S285–S300. <https://doi.org/10.1086/296367>
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1986b). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 76(4), 728–741.
- Kant, I. (1949). *On a supposed right to lie from altruistic motives*. Chicago University Press.
- Kant, I. (1959). *Foundations of the metaphysics of morals* (L. W. Beck, Trans.). Bobbs-Merrill. (Original work published 1785)
- Levine, E. E., & Cohen, T. R. (2018). You can handle the truth: Mispredicting the consequences of honest communication. *Journal of Experimental Psychology: General*, 147(9), 1400–1429. <https://doi.org/10.1037/xge0000488>
- Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology*, 53, 107–117. <https://doi.org/10.1016/j.jesp.2014.03.005>
- Levine, E., Hart, J., Moore, K., Rubin, E., Yadav, K., & Halpern, S. (2018). The surprising costs of silence: Asymmetric preferences for prosocial lies of commission and omission. *Journal of Personality and Social Psychology*, 114(1), 29–51. <https://doi.org/10.1037/pspa0000101>
- Levine, E., & Schweitzer, M. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, 126, 88–106. <https://doi.org/10.1016/j.obhdp.2014.10.007>
- Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives*, 21(2), 153–174. <https://doi.org/10.1257/jep.21.2.153>

- Loewenstein, G., & Small, D. A. (2007). The Scarecrow and the Tin Man: The vicissitudes of human sympathy and caring. *Review of General Psychology*, 11(2), 112–126. <https://doi.org/10.1037/1089-2680.11.2.112>
- Lupoli, M. J., Levine, E. E., & Greenberg, A. E. (2018). Paternalistic lies. *Organizational Behavior and Human Decision Processes*, 146, 31–50. <https://doi.org/10.1016/j.obhdp.2018.01.001>
- Margolis, J. D., & Molinsky, A. (2008). Navigating the bind of necessary evils: Psychological engagement and the production of interpersonally sensitive behavior. *Academy of Management Journal*, 51(5), 847–872. <https://doi.org/10.5465/amj.2008.34789639>
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), 633–644. <https://doi.org/10.1509/jmkr.45.6.633>
- Molinsky, A., & Margolis, J. (2005). Necessary evils and interpersonal sensitivity in organizations. *Academy of Management Review*, 30(2), 245–268. <https://doi.org/10.5465/amr.2005.16387884>
- Monin, B., Pizarro, D. A., & Beer, J. S. (2007). Deciding versus reacting: Conceptions of moral judgment and the reason-affect debate. *Review of General Psychology*, 11(2), 99–111. <https://doi.org/10.1037/1089-2680.11.2.99>
- Murnighan, J. K. (1991). *The dynamics of bargaining games*. Prentice Hall.
- Ortony, A. J., & Gupta, S. (2018). Lying and deception. In J. Meibauer (Ed.), *The Oxford handbook of lying* (pp. 149–169). Oxford University Press.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677. <https://doi.org/10.1038/ncomms4677>
- Richard, C., Lajeunesse, Y., & Lussier, M. T. (2010). Therapeutic privilege: Between the ethics of lying and the practice of truth. *Journal of Medical Ethics*, 36(6), 353–357. <https://doi.org/10.1136/jme.2009.033340>
- Roberts, A. R., Levine, E. E., & Sezer, O. (2020). Hiding success. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspi0000322>
- Rogers, T., Zeckhauser, R., Gino, F., Norton, M. I., & Schweitzer, M. E. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others. *Journal of Personality and Social Psychology*, 112(3), 456–473. <https://doi.org/10.1037/pspi0000081>
- Schein, C., & Gray, K. (2015). The unifying moral dyad: Liberals and conservatives share the same harm-based moral template. *Personality and Social Psychology Bulletin*, 41(8), 1147–1163. <https://doi.org/10.1177/0146167215591501>
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1–65.
- Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, 101(1), 1–19. <https://doi.org/10.1016/j.obhdp.2006.05.005>
- Scott, K. (2017). *Radical candor: How to get what you want by saying what you mean*. Pan Macmillan.
- Sezer, O., Gino, F., & Bazerman, M. H. (2015). Ethical blind spots: Explaining unintentional unethical behavior. *Current Opinion in Psychology*, 6, 77–81. <https://doi.org/10.1016/j.copsyc.2015.03.030>
- Shalvi, S., Dana, J., Handgraaf, M. J., & De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, 115(2), 181–190. <https://doi.org/10.1016/j.obhdp.2011.02.001>
- Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science*, 24(2), 125–130. <https://doi.org/10.1177/0963721414553264>
- Shiloh, S., Ben-Sinai, R., & Keinan, G. (1999). Effects of controllability, predictability, and information-seeking style on interest in predictive genetic testing. *Personality and Social Psychology Bulletin*, 25(10), 1187–1195. <https://doi.org/10.1177/0146167299258001>
- Shu, L. L., Gino, F., & Bazerman, M. H. (2011). Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated forgetting. *Personality and Social Psychology Bulletin*, 37(3), 330–349. <https://doi.org/10.1177/0146167211398138>
- Slepian, M. L., Chun, J. S., & Mason, M. F. (2017). The experience of secrecy. *Journal of Personality and Social Psychology*, 113(1), 1–33. <https://doi.org/10.1037/pspa0000085>
- Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage, Inc.
- Sweeny, K., Melnyk, D., Miller, W., & Shepperd, J. A. (2010). Information avoidance: Who, what, when, and why. *Review of General Psychology*, 14(4), 340–353. <https://doi.org/10.1037/a0021288>
- Sweetser, E. E. (1987). The definition of “lie”: An examination of the folk models underlying a semantic prototype. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 43–66). Cambridge University Press. <https://doi.org/10.1017/CBO9780511607660.003>
- Talwar, V., Murphy, S. M., & Lee, K. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development*, 31(1), 1–11. <https://doi.org/10.1177/0165025406073530>
- Telushkin, J. (1994). *Jewish wisdom: Ethical, spiritual, and historical lessons from the great works and thinkers*. Morrow.
- Tjur, T. (2009). Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination. *The American Statistician*, 63(4), 366–372. <https://doi.org/10.1198/tast.2009.08210>
- Tyler, J. M., Feldman, R. S., & Reichert, A. (2006). The price of deceptive behavior: Disliking and lying to people who lie to us. *Journal of Experimental Social Psychology*, 42(1), 69–77. <https://doi.org/10.1016/j.jesp.2005.02.003>
- Varden, H. (2010). Kant and lying to the murderer at the door. . . One more time: Kant’s legal philosophy and lies to murderers and Nazis. *Journal of Social Philosophy*, 41(4), 403–421. <https://doi.org/10.1111/j.1467-9833.2010.01507.x>
- Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, 86(4), 629–647. <https://doi.org/10.1037/0022-3514.86.4.629>
- Weisel, O., & Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences of the United States of America*, 112(34), 10651–10656. <https://doi.org/10.1073/pnas.1423035112>
- Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes*, 115(2), 157–168. <https://doi.org/10.1016/j.obhdp.2010.10.001>
- Yaniv, I., Benador, D., & Sagi, M. (2004). On not wanting to know and not wanting to inform others: Choices regarding predictive genetic testing. *Risk Decision and Policy*, 9(4), 317–336. <https://doi.org/10.1080/14664530490896573>
- Yip, J. A., & Schweitzer, M. E. (2016). Mad and misleading: Incidental anger promotes deception. *Organizational Behavior and Human Decision Processes*, 137, 207–217. <https://doi.org/10.1016/j.obhdp.2016.09.006>
- Zembyat, J. S. (1993). Aristotle on lying. *Journal of the History of Philosophy*, 31(1), 7–29. <https://doi.org/10.1353/hph.1993.0008>

(Appendix follows)

Appendix

Mechanism Questions in Community Standard Experiments (Study 2)

Immediate harm

- To what extent would telling the truth in this scenario cause unnecessary harm? (Survey Groups 1–2)
- To what extent would honesty cause pain to the [individual]? (Survey Groups 1–3)
- To what extent would telling a lie protect the [individual]'s feelings? (Survey Groups 1–3)
- To what extent would honesty cause harm to the [individual]? (Survey Group 3 only)
- To what extent would lying benefit the [individual]? (Survey Group 3 only)

Instrumental value

- To what extent would telling the truth in this scenario have the potential to influence the [individual]'s behavior? (Survey Groups 1–3)

- To what extent would telling the truth in this scenario be valuable to the [individual]'s improvement* overall well-being? (Survey Groups 1–3)
- To what extent is the honest information necessary for the [individual] to know? (Survey Group 3 only)
- To what extent is the honest information useful for the [individual]'s growth or enlightenment? (Survey Group 3 only)

Note. *The word improvement was removed from the *Death bed* vignette and from Survey Group 3 to eliminate confusion.

Received August 15, 2019

Revision received March 22, 2021

Accepted March 23, 2021 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!