Original Manuscript

Policy Insights from the
Behavioral and Brain Sciences
1–9
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23727322251408075
journals.sagepub.com/home/bbs

Harnessing the Benefits of Honesty for Individuals, Relationships, and Society

Bonnie M. Le[1,*] and William J. Chopik[2,*]

Abstract
Honesty is highly valued across societies. Communicating in truthful ways is widely seen as moral and desired across interpersonal relationships. Despite the value societies place on honesty, there are personal and social factors that often inhibit truthful communication and its potential benefits. In the current paper, we describe the implications of honesty for policy in three key domains: in interpersonal and group relationships, within healthcare settings between patients and providers, and in broader online settings. We highlight honesty as an interpersonal communication process focused on sharing the truth and fostering its understanding in others. We further draw on insights on the personal and social factors that motivate people to communicate honestly, or not, and the associated costs and benefits of doing so. Anchoring on key theories and findings related to honesty, we highlight how truthful communication may have important policy implications for individuals, relationships, and society.

Keywords
honesty, interpersonal relationships, groups relationships health, online communication, well-being

Social Media Post

Being honest is valued but can be difficult to do. How are people to navigate the potential costs, but desired benefits, of honesty? In the current paper, we highlight research that can inform policy in ways that harness the benefits, and mitigate the costs, of honesty for individuals, relationships, and broader society.

Key Points

- In the current paper, we describe the promise and pitfalls of honesty in three key domains: within interpersonal and intergroup relations, between patients and practitioners, and within broader online interactions.
- Within close relationships and between groups, emphasizing that honesty is largely beneficial for the self and other can maximize communication and uptake of truthful information. While there are times when white lies can serve a purpose, sharing the truth is generally less harmful than we believe it will be.
- In healthcare, honesty as an interpersonal communication process between practitioners and patients is essential for effective treatment. Practitioner honesty facilitates patient decision-making; patient honesty promotes informed treatment from practitioners. Honest lines of communication are rooted in building a foundation of trust in managing health.
- Honesty and truth are becoming increasingly difficult to discern online. Misinformation is rampant, and deception and fraud have become normative, posing serious challenges to governments and societies. Policy recommendations for such emerging and pervasive dishonesty are challenging to recommend beyond encouraging regulatory guardrails in online spaces. To the extent possible, identifying and penalizing the spread of false information can protect the integrity of truthful information online.

Honesty—or sharing truthful information, while abstaining from lying, deception, and omissions of information (Miller, 2021; Reynolds et al., 2025; Shalvi et al., 2025)—is highly valued in our relationships and across societies (Anderson, 1968; Chandler, 2018). Honesty can be expressed consistently across contexts at the enduring trait level and also as behaviors that can vary based on situations (Reynolds et al., 2025). Honesty can be seen as an interpersonal communication process in which one person shares the truth while also working to foster an understanding of

[1]University of Rochester, Rochester, USA
[2]Michigan State University, East Lansing, USA

[*]Both authors contributed equally to this work.

Corresponding Author:
Bonnie Le, University of Rochester, Rochester, USA.
Email: bonnie.le@rochester.edu

the truth in others (Cooper et al., 2023; Fritz, 2020). In the current paper, after detailing how honest communication can be promoted, we describe the policy implications of honesty in three domains: in interpersonal and intergroup relations, in healthcare settings, and in broader online interactions, including those between individuals and artificial intelligence (AI).

Research has shown that people are generally honest, with deviations from honesty generally being relatively benign or motivated by ostensibly good intentions (e.g., sparing feelings) (DePaulo et al., 1996; Shalvi et al., 2025). While people are generally honest, there are several factors that may encourage or discourage honesty in their daily life, relationships, and in online spaces. One primary personal concern people consider when being honest is how sharing the truth may shape others' perceptions of them. People are often most inclined to be honest when doing so maintains or promotes a positive self-image (Shalvi et al., 2025). Similarly, experimental work has indicated that people are more inclined to be honest when doing so is self-serving or personally beneficial (Levine & Munguia Gomez, 2020).

In addition to egocentric concerns, people often consider whether honesty will help or harm others or society. For example, when people lie in their relationships, they often do so for prosocial reasons, such as to avoid hurting others' feelings (DePaulo et al., 1996; Levine et al., 2020). This raises a trade-off individuals often consider between the harm honest information may incur relative to its benefits, particularly when sharing negative information (Levine et al., 2020). Across several studies, people were generally cautious about sharing truthful negative information, overestimating how much recipients of honesty will be hurt by the truth, which in turn can decrease willingness to be honest (Levine & Cohen, 2018). For instance, people are reluctant to share negative news or feedback, including with close others and in organizational settings. Despite this, sharing honest messages, especially when doing so in ways that facilitate receptivity from recipients, can promote social connection and recipient growth (Fulham et al., 2022; Levine & Cohen, 2018; Levine et al., 2020). Thus, people are most motivated to be truthful when doing so preserves or promotes their relationships.

The outcomes of being honest can either promote or hinder personal and relationship outcomes (Le et al., 2022). While empirical work on the outcomes of being honest is still burgeoning, much of the existing findings have pointed to the benefits of being honest. For instance, honesty has been linked to greater personal well-being and health (Allan & Alba-Fisch, 2015; Weziak-Bialowolska et al., 2021). Of course, being honest is not always beneficial. In times when being truthful has personal costs, such as when sharing information that others do not want shared (e.g., whistleblowing), people may experience stress and strain relationships with those who do not want truthful information to be shared (Johnson, 2003; Park & Lewis, 2018).

Relationally, honesty has also been found to largely be associated with benefits. Honesty can build trust and intimacy with others (LaFollette & Graham, 1986). This is true not only when people express themselves honestly, but also when people see their friends and partners as honest (Le et al., 2025; Thielmann et al., 2017). Although people are hesitant to be truthful about negative information, research has consistently found that sharing the truth—even when it may hurt—is received more positively than expected and strengthens relational ties (Le et al., 2025; Levine & Cohen, 2018; Levine et al., 2020). For example, honesty in daily social interactions, such as having difficult conversations with close relationship partners and providing negative feedback in organizational contexts—especially in ways that facilitate recipient receptivity—increases social connection, relationship satisfaction, and recipient growth (Fulham et al., 2022; Le et al., 2025; Levine & Cohen, 2018; Levine et al., 2020). However, honesty isn't overwhelmingly or homogenously good. When honesty causes unnecessary harm to others (i.e., telling someone you do not like the gift they got you or that their cooking is bad), this can undermine relationships and well-being; here, prosocial lies may have value and function in preserving relationships (Levine & Lupoli, 2022; Levine et al., 2020).

The varied contexts in which honesty is expressed versus not may make it challenging to discern who is being honest and what information is truthful. Research has indicated that there is little agreement among friends about who is honest (Yang et al., 2025). This is perhaps unsurprising given consistent evidence indicating the difficulty of detecting lies (Brennen & Magnussen, 2023). Further compounding the challenge of discerning honesty, expressions of honesty can entail truthful sharing of information that is collectively agreed upon (i.e., scientific facts) or relatively idiosyncratic (i.e., personal feelings) (Le et al., 2025). While accurately discerning honesty is challenging, expressions and perceptions of honesty—regardless of whether these perceptions are grounded in reality—can support satisfying relationships (Le et al., 2025). Thus, while accurate perceptions of honesty are difficult to cultivate, when expressed under appropriate contexts, honesty still holds promise for personal and relationship well-being.

Altogether, we can think of honesty as an interpersonal communication process that emphasizes sharing the truth and fostering truthful understanding in others (Cooper et al., 2023). Being honest can be hard—there are personal image concerns involved in being honest as well as social concerns of hurting others when sharing undesirable information (Levine et al., 2020; Shalvi et al., 2025). Research has largely indicated that honesty benefits personal and interpersonal health and well-being (Le et al., 2022). However, there are important contexts in which honesty can backfire, such as when sharing private or unnecessarily hurtful information (Levine et al., 2020; Park & Lewis, 2018). Despite its costs, honesty is generally valued, although discerning the truth in others is challenging (Yang et al., 2025).

Although much psychological research has focused on the individual and relational considerations of honesty, there are also several imperatives concerning honesty that have implications for the functioning of organizations and societies. These considerations emerge either through the establishment of norms around honesty, the emergence of technologies that shape communication and opportunities to share the truth, or even the ways that individual and relational processes trickle up to shape social and public institutions. More broadly, integrating across current findings about honesty, several insights arise that may help us harness the benefits, and mitigate the costs, of honesty at the individual, relational, and societal level. In the sections that follow, we focus on three domains: interpersonal and intergroup relationships, healthcare settings between patients and practitioners, and online spaces between individuals and AI.

Honesty in Interpersonal and Group Relationships

Some of the clearest implications of honesty lie in how truth-telling affects other people. Capitalizing on the ways honesty can be an interpersonal communication strategy (Cooper et al., 2023; Fritz, 2020), people weigh several considerations to determine whether they ought to be honest with others. As previously discussed, people want to see themselves and be seen by others in a positive light (Swann et al., 1992). In this way, honesty is one of the most valued characteristics in choosing friends and romantic partners (Anderson, 1968; Chandler, 2018). Even a modest admission that lying is occasionally acceptable is met with greater punishment than the consistent statement that honesty is always the best policy, even if one does not abide by the saying (Huppert et al., 2023). Yet at the same time, another perspective is that strategic acts of deception are okay to use to avoid conflict—and that these acts promote relationship stability without constituting a violation of moral norms (Balliet & Van Lange, 2013; Shalvi et al., 2025).

This perspective presents an uncomfortable possibility—that navigating social relationships in occasionally dishonest ways might benefit relationships, if it is motivated by ostensibly "right" reasons (e.g., sparing others' feelings, avoiding conflict). Although a tempting possibility, being dishonest, even with good intent, is misaligned with what people say they want from relationship partners, with how relationship partners actually respond to uncomfortable truths, *and* the extent to which both people benefit from honesty. For example, given that people value honesty in others and want to be seen as honest (Chandler, 2018), the tendency to lie to avoid uncomfortable situations or to smooth over conflict might be misplaced. In a series of laboratory and field studies, Levine and Cohen (2018) show that people directed to be honest underestimate how pleasurable, socially connecting, and even benign being honest feels (compared to being kind). Indeed, the fear of aversive conversations and the desire to avoid hurting other people's feelings are some of the main reasons why people don't even start these difficult conversations or lie when these conversations do happen (Levine et al., 2025). Thus, it may not be necessary to avoid the truth or assume that the truth will hurt, because it might not (Levine & Cohen, 2018), and close others want the truth anyway (Reynolds et al., 2025).

That is not to say that telling the truth in relationships and groups is not difficult or that there are never negative consequences (e.g., for whistleblowers; Johnson, 2003). But the benefits for relationships and groups may still outweigh the costs. Honesty can shine a light on unspoken truths or injustices, correct a wayward path to encourage success, build trust and intimacy between people, and correct misconceptions that lead to disastrous consequences. But how does one do this in the context of relationships and groups? A recent model proposed by Levine and colleagues (2025) suggests that focusing on making honesty salient and normal, encouraging reflection to avoid self-interested biases, instilling accountability and transparency, and privileging integrity so as to not be compromised by a myopic focus on performance are important places to start. These efforts and the honest communication they foster encourage personal growth and improvement for everyone involved. It helps people maintain the integrity they so often seek, build trust, reduce stress, and motivate further prosociality (Bellucci & Park, 2020; Levine et al., 2025; ten Brinke et al., 2015; Weziak-Bialowolska et al., 2021). Thus, honesty may generate unpleasant feelings. However, therein also lies the solution—that honesty and authenticity in the short run might resolve more conflict than it creates in the long run.

Being dishonest in interpersonal contexts has many policy and practical implications—it can compromise people's relationships, sow discord in groups and organizations, and cultivate a cynical culture in society. Unlike other settings (see below), few actionable policy recommendations can be implemented at a societal level. Passing a law saying you cannot lie to friends or having governments or organizations incentivize honesty with spouses and coworkers doesn't make sense and is unlikely to be unsuccessful. But what can individuals do to encourage honesty in interpersonal and group settings? Increasing honesty in interpersonal relationships might be one of the more manageable contexts for interventions. There is a substantial literature on individual-level interventions to encourage honesty that provides some actionable recommendations people can implement in the relationships and groups in which they find themselves (Hertwig & Mazar, 2022). These interventions include nudges (subtle behavioral interventions to drive decisions), economic frameworks (evaluating cost-benefit calculations), and targeting internal rewards (focusing on the psychological or moral costs and benefits). Nudge interventions might be something like making the honest option a default or auto-verify information. Economic interventions might be

something like levying a financial penalty for dishonesty (making consequences of lying more transparent). Internal reward interventions might be to have people make honesty pledges or invoking identity considerations (e.g., framing actions about being an honest *person* rather than merely telling people that dishonesty is wrong).

But what does this look like in interpersonal and group contexts? For nudges, honesty cues (e.g., "Can we try to be honest and open with each other?") and setting norms (e.g., "Team members tend to give straightforward feedback because it helps everyone connect better, sets expectations, and keeps us on the same page.") might disabuse people of the misperception that honesty is something to be avoided lest people's feelings are hurt. Given the social transmission of honesty norms (Mann et al., 2014), intervening at the relationship-level might be more effective than appealing to solely individual-level motivations. For economic interventions, making salient the social costs of dishonesty would also likely help, such as when a business partner links honesty to trustworthiness and closeness or how when people's social networks overlap to a greater degree, lies are easily traceable and policed (Brass et al., 1998). Appealing to internal rewards in an interpersonal context might also be effective. For example, close others often share an identity and increasingly consider themselves to be a shared unit with similar values (Emery et al., 2025). If people in relationships have a sense of who they are (i.e., self-concept clarity), defining their relational identity on a foundation of honesty, making commitments to be honest, and having shared relational cues toward honesty would likely result in increased honesty with both each other and people outside the relationship.

Although these are just examples of potential recommendations for increasing honesty interpersonally, these and other efforts should target the reasons why people tend to be dishonest. Open disclosure through explicit agreements, reminding others of the benefits and preferences for honesty, institutionalizing honesty (e.g., establishing a default norm of honesty), and reducing barriers to honesty (e.g., by making it seem less evaluative or harmful) could collectively address the interpersonal barriers to being honest. Ultimately, in all of these recommendations, the onus typically falls on individuals to enact honest behavior, set honesty norms, or to make explicit the consequences of dishonesty. However, given the interdependence and social transmission of honest behavior, it might also be particularly effective at encouraging honest behavior.

Honesty in Healthcare Settings

In modern medicine, honesty is seen as a character virtue that is central to being a good and responsible physician, rooted in the key principle of "doing no harm" (Gold, 2004; Pellegrino, 2002). While most physicians endorse honesty to their patients, being fully honest is often difficult to do. Consistent with the psychological literature, egocentric concerns influence practitioner honesty. For example, speaking up about the truth may compromise the perceived or felt competence of a practitioner, such as when disclosing a mistake to one's medical team. Being honest can also create tension in medical teams working under high pressure, such as when calling out another person's error (Szymczak, 2016). Nevertheless, honesty about medical errors can guard against legal actions such as malpractice (Jotkowitz et al., 2006). There are also practical constraints to being honest. Practitioners are often under immense time pressure—inhibiting their ability to get into nuanced and detailed information about the truth (Mohammed, 2005). Further, medical assessments are not always definitive. This ambiguity can lead to avoidance of honesty or endorsement of benevolent deception (Hart, 2022; Mohammed, 2005).

While there are some personal (i.e., competence) and practical (i.e., time) concerns with being honest, practitioners primarily weigh patient welfare. Such concerns align with the idea that prosocial motivations often underlie dishonesty. In their focus on preventing harm for their patients, practitioners may at times engage in "benevolent deception," or withholding or obscuring truthful information from patients (Farber, 2020; Hart, 2022). Withholding truth may psychologically (but not ethically) absolve practitioners of causing patient distress, suffering, or extinguishing hope when sharing a negative or terminal diagnosis or prognosis (Gold, 2004; Huddle, 2012; Jotkowitz et al., 2006). Despite concerns that the truth may harm patients, little work suggests this is the case (Jotkowitz et al., 2006). While research is limited, sharing important health information has been found to promote some benefits, including greater trust from patients and in the healthcare system and foster more positive relationships among patients, relatives, and staff (Centeno-Cortés & Núñez-Olarte, 1994). This is consistent with aforementioned work that honestly sharing negative information is not as harmful as people think (Levine & Cohen, 2018; Levine et al., 2020). While this research was not conducted in health settings, it dovetails with arguments that sharing the truth with patients can support their autonomy to make better decisions, thereby supporting their health (Gold, 2004; Hart, 2022). Thus, while practitioners have historically been concerned about harming patients by sharing truthful information, evidence suggests that the benefits may outweigh the risks of doing so.

Nonetheless, dishonesty can persist in the healthcare domain, especially in situations where the utility of health information is ambiguous. One complex domain in which prosocial lies have a place concerns patients whose cognitive functioning is compromised. In these cases, honest communication may not be comprehensible, ineffective, or might even make matters worse if there is no actionable guidance. Dishonesty with such patients might include telling white lies, withholding the truth, subtle guidance or misdirection, and going along with a patient's misperceptions. Some

caregivers characterize these lies as sometimes being positive or at least permissible given the challenges they face (Elvish et al., 2010). There is evidence that therapeutic lying may reduce symptoms and the use of restraints related to dementia, but more research is needed to clarify whether dishonesty in this context is indeed more beneficial than harmful (Jotkowitz et al., 2006). Thus, while honesty by practitioners has been increasingly endorsed as a best practice, there are contexts in which dishonesty can, debatably, be acceptable.

Honesty about health information works both ways. In medical assessments, practitioners often rely on patients to share important information about their habits, behaviors, and symptoms. Despite the importance of accurate information for practitioners, individuals may be reluctant to share information due to personal concerns (e.g., embarrassment, non-compliance with prior recommendations). While patients tend to be more honest than not, upwards of 90% of patients have reported being dishonest at some point with their practitioner (Farber, 2020), such as about information they feel ashamed to share (e.g., substance abuse) or out of concern about reactions or consequences (e.g., for suicidal thoughts). While patients believe these lies may help them feel better, it often prompts them to feel worse over time (Farber, 2020). Additionally, patient dishonesty may compromise a practitioner's ability to optimally support patient health. In one study, it has been found that clinicians detected hidden clients feelings less than half of the time and only detected patient lies 9% of the time (Hill et al., 2001). Thus, administering treatment with selective or incorrect information is a challenge faced by practitioners.

Facilitating honesty interpersonally in the medical field requires sensitivity and compassion from practitioners and directly consulting patients on how much of the truth they want to know (Jotkowitz et al., 2006). This can be especially important when considering the cultural diversity of patients. North Americans prefer honesty from their practitioners to make autonomous decisions. However, in China, both physicians and patients prefer family members be the primary decision-makers about health (Jotkowitz et al., 2006). Thus, practitioners should aim for truth-telling to be a *dialogue*. Doing so allows practitioners to maintain honesty while being sensitive to patients' unique preferences and the cultural context. Specifically, through this dialogue, practitioners can align the information they share with patients' individualized desires for how much of the truth they seek to know and feel they can manage. Ultimately, these efforts may promote patients' autonomy and therapeutic efficacy (Gold, 2004; Jotkowitz et al., 2006).

Given the interplay between practitioners, patients, and the broader medical, cultural, and legal environment, policies surrounding honesty in the medical setting should consider all of these factors. In the United States, there are legal requirements in many states requiring full disclosure from practitioners, which ultimately facilitates practitioner honesty and penalizes dishonesty (i.e., malpractice) (Jotkowitz et al., 2006). These policies align with the broader movement in the medical field to promote full disclosure. Nonetheless, policies concerning honesty in the medical setting could benefit from weighing patient perspectives: patient desire to know the truth, the patient's capacity to comprehend the truth, and cultural preferences for truth disclosure. In other words, there are important nuances that can and should shape practitioners' full disclosure of patient health information, and policies surrounding practitioner honesty should weigh these nuances in guiding recommendations for when practitioners share the truth, how much of the truth they share, and over what time frame they share the truth.

Policy considerations have focused primarily on practitioners, but of course, patient honesty can be crucial for practitioners to support optimal health outcomes. But much like in the realm of interpersonal relationships more broadly, enforcing honesty on behalf of patients may be most effective through individual nudges (i.e., emphasizing honesty as normative and often more beneficial than costly) versus through the use of stringent policy recommendations. Given the sensitive and private nature of health histories and behaviors, intervening in ways that encourage patients to willfully and consensually provide important health information could be a promising step in encouraging patient honesty. Finally, creating trusting relationships between practitioners and patients is also essential (Farber, 2020), and this can include practitioners being clear about their privacy policies and how patient information will be used in their treatment.

Honesty in Online Communication and with Artificial Intelligent Agents

Online contexts somewhat resemble interpersonal settings but introduce many additional considerations less relevant to those contexts. Here we discuss two important characteristics of online contexts: (1) online communication introduces more opportunities and motives to behave dishonestly and (2) the integration of AI in online communication and information consumption may also lead people astray by providing incomplete information or obsequious answers, while also occasionally providing several benefits to combat misinformation.

Establishing a culture and norm of honesty is difficult in offline settings, even when people know each other, are invested in relationships, when they want to be honest, and with clear accountability and punishment of dishonesty (Huppert et al., 2023). One might assume that, because online interactions are less psychologically rich than in-person interactions, people may be more likely to disclose truths and vulnerabilities (Carbone & Loewenstein, 2023). But some evidence suggests that online contexts may reduce inhibition and enable dishonest behavior (Hancock & Toma, 2009; Suler, 2004). Although there are contexts in which honesty can be either enhanced or diminished

online, it is likely the case that the *consequences* of dishonesty may be occasionally worse online than in face-to-face interactions given the scope and reach of large online ecosystems. For example, approximately 20–30% of U.S. adults have lost some amount of money in online scams or fallen victim to online financial fraud or cybercrime (Global Anti-Scam Alliance, 2025; Newall, 2023) and 73% of U.S. adults have experienced some kind of online scam or attack (Gottfried et al., 2025). Nearly 60% of adults worldwide have experienced some type of online scam (Global Anti-Scam Alliance, 2025). Among the nearly $13 billion lost to fraud in 2024, imposter scams (where people lie about their identity and intentions) accounted for nearly $3 billion (Federal Trade Commission, 2025). All told, the costs of dishonesty in online settings can be devastating for people and society.

The introduction of AI helps make dishonest actors more effective, particularly when it comes to personalization of communication, selective information presented, or even suppressing factual and truthful information. For example, AI-generated phishing emails have effectiveness rates that are comparable to human-generated emails, although the scale of AI-generated capabilities is much larger (Heiding et al., 2024). Some of the most glaring examples of AI mimicking some of the aforementioned interpersonal processes (e.g., lying to maintain relational harmony) are scenarios in which AI chatbots and large language models (LLMs) provide obsequious answers that compliment users to encourage them to continue to communicate at the expense of the truth (Bo et al., 2025). The authoritative nature of LLM responses also provides a sense of truth-telling given their leveraging of large amounts of information, and people rely on conversational agents as though they were human relational partners with benevolent motives (Glikson & Woolley, 2020; Logg et al., 2019). Further, exposure to AI-generated material also leads to distrust in traditional media sources given that the former could challenge misinformation (Vaccari & Chadwick, 2020). Without clear regulatory guardrails or insights into how LLMs verify the accuracy of content, ensuring that people are provided with truthful information will prove challenging. In guiding policy-level interventions, it is useful to focus on how people use and perceive AI. For example, most trust LLM-based chatbots with practical tasks but not health and political information (Presiado et al., 2024). This likely stems how easily LLMs to produce misinformation—one explanation for the occasional low accuracy rates (Su et al., 2025). Lawmakers have only recently evaluated AI and misinformation, motivating efforts to police accuracy and protect consumers (U.S. House Committee on Financial Services, 2025; Zakrzewski, 2023). These discussions recommend reminding people to question AI-generated content, making transparent how LLM chatbots are trained, building user- or system-generated ways to correct information, requiring benchmarks for accuracy and truthfulness, and conducting audits of these systems. Admittedly, it is unclear which policies are most effective in this space. Although models and studies provide potential guidance to encourage honesty among humans (and guidance about being critical consumers of online content) (Levine et al., 2025; Presiado et al., 2024), emerging work in the decades to come will hopefully provide actionable solutions to navigate online spaces with AI agents and capabilities that are still being developed.

In many ways, encouraging honesty in online settings, in terms of discouraging lying by perpetrators (and increasing skepticism among potential victims), requires different intervention techniques and recommendations. The recommendations for evaluating the truthfulness of LLM chatbots and AI-generated content, although largely untested, provide some guidance for how countries and organizations can calibrate their interventions to the mode of interpersonal communication at the system-level (in the case of AI) and at the individual or interpersonal-level (in the case of human interactions). What does policy look like in online spaces more generally? It can be difficult to discourage lying among people willing to orchestrate online scams that target vulnerable people. However, potential strategies to discourage online lying include reducing psychological distance (by reminding scammers of the toll and suffering their efforts have on real people) to humanize potential victims (Levine et al., 2025), making norms more salient through transforming online spaces to punish dishonesty (Tenbrunsel & Messick, 2004), and implementing systems that enhance transparency and accountability (Federal Bureau of Investigation, 2024). Such systematic interventions have shown promising results in reducing scams in the context of mobile dating apps (Kelly & Lung, 2025).

Many recommendations also exist for helping people *avoid* being duped by dishonesty found online, but they are beyond the scope of this article (e.g., teaching people about manipulation tactics, increasing situational awareness, seeking out disconfirming information) (Button & Cross, 2017; Wineburg & McGrew, 2019). Although practical and policy recommendations for navigating interpersonal and healthcare settings have been discussed and studied for many years, there is comparatively less work on online and AI-related spaces. These spaces may prove to be more challenging to intervene in to increase honesty, but there are likely a few strategies that could mitigate the risk and consequences of online dishonesty.

Conclusion

The challenge of honesty has major implications for people's and society's health and happiness in the immediate and far-reaching future. Moving forward, policy targeted at the personal, relational, organizational, and virtual levels will benefit from promoting behaviors and systems that encourage truthful information while also confronting the challenging contexts in which deviations of honesty may be

appropriate. In the latter case, clearly defining when honesty may cause real harm will be important to clarify and build into policy. Taken together, these steps may help us harness the benefits and mitigate the costs of sharing and receiving the truth.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Bonnie M. Le https://orcid.org/0000-0002-1857-6716
William J. Chopik https://orcid.org/0000-0003-1748-8738

References

Allan, B. A., & Alba-Fisch, M. (2015). Balance among character strengths and meaning in life. *Journal of Happiness Studies*, *72*(5), 1247–1261. https://doi.org/10.1007/s10902-014-9557-9

Anderson, N. H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*(3), 272–279. https://doi.org/10.1037/h0025907

Balliet, D., & Van Lange, P. A. M. (2013). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin*, *139*(5), 1090–1112. https://doi.org/10.1037/a0030939

Bellucci, G., & Park, S. Q. (2020). Honesty biases trustworthiness impressions. *Journal of Experimental Psychology. General*, *149*(8), 1567–1586. https://doi.org/10.1037/xge0000730

Bo, J. Y., Kazemitabaar, M., Deng, M., Inzlicht, M., & Anderson, A. (2025). *Invisible saboteurs: Sycophantic LLMs mislead novices in problem-solving tasks* (No. arXiv:2510.03667). arXiv.

Brass, D. J., Butterfield, K. D., & Skaggs, B. C. (1998). Relationships and unethical behavior: A social network perspective. *Academy of Management Review*, *23*(1), 14–31. https://doi.org/10.2307/259097

Brennen, T., & Magnussen, S. (2023). Lie detection: What works? *Current Directions in Psychological Science*, *32*(5), 395–401. https://doi.org/10.1177/09637214231173095

Button, M., & Cross, C. (2017). *Cyber frauds, scams and their victims*. Routledge.

Carbone, E., & Loewenstein, G. (2023). Privacy preferences and the drive to disclose. *Current Directions in Psychological Science*, *32*(6), 508–514. https://doi.org/10.1177/09637214231196097

Centeno-Cortés, C., & Núñez-Olarte, J. M. (1994). Questioning diagnosis disclosure in terminal cancer patients: A prospective study evaluating patients' responses. *Palliative Medicine*, *8*(1), 39–44. https://doi.org/10.1177/026921639400800107

Chandler, J. (2018). Likeableness and meaningfulness ratings of 555 (+487) person-descriptive words. *Journal of Research in Personality*, *72*, 50–57. https://doi.org/10.1016/j.jrp.2016.07.005

Cooper, B., Cohen, T. R., Huppert, E., Levine, E. E., & Fleeson, W. (2023). Honest behavior: Truth-seeking, belief-speaking, and fostering understanding of the truth in others. *Academy of Management Annals*, *17*(2), 655–683. https://doi.org/10.5465/annals.2021.0209

DePaulo, B. M., Kirkendol, S. E., Kashy, D. A., 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

Elvish, R., James, I., & Milne, D. (2010). Lying in dementia care: An example of a culture that deceives in people's best interests. *Aging & Mental Health*, *14*(3), 255–262. https://doi.org/10.1080/13607861003587610

Emery, L. F., Hughes, E. K., & Muise, A. (2025). Self-expansion theory: Origins, current evidence, and future horizons. *Social and Personality Psychology Compass*, *19*(9), e70082. https://doi.org/10.1111/spc3.70082

Farber, B. A. (2020). Disclosure, concealment, and dishonesty in psychotherapy: A clinically focused review. *Journal of Clinical Psychology*, *76*(2), 251–257. https://doi.org/10.1002/jclp.22891

Federal Bureau of Investigation. (2024). *Criminals use generative artificial intelligence to facilitate financial fraud*. https://www.ic3.gov/PSA/2024/PSA241203

Federal Trade Commission. (2025). *New FTC data show a big jump in reported losses to fraud to $12.5 Billion in* 2024. https://www.ftc.gov/news-events/news/press-releases/2025/03/new-ftc-data-show-big-jump-reported-losses-fraud-125-billion-2024

Fritz, J. H. (2020). Honesty as ethical communicative practice: A framework for analysis. In *Integrity, honesty, and truth seeking* (pp. 127–152). Oxford University Press.

Fulham, N. M., Krueger, K. L., & Cohen, T. R. (2022). Honest feedback: Barriers to receptivity and discerning the truth in feedback. *Current Opinion in Psychology*, *46*. https://doi.org/10.1016/j.copsyc.2022.101405

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence. *Review of Empirical Research. Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Global Anti-Scam Alliance. (2025, October 7). *Global scams on the rise: Over half of adults worldwide report scam encounters, 23% lost money*. GASA. https://www.gasa.org/post/global-scams-on-the-rise-over-half-of-adults-worldwide-report-scam-encounters

Gold, M. (2004). Is honesty always the best policy? Ethical aspects of truth telling. *Internal Medicine Journal*, *34*(9–10), 578–580. https://doi.org/10.1111/j.1445-5994.2004.00673.x

Gottfried, J., Park, E., & Anderson, M. (2025, July 31). *Online scams and attacks in America today*. Pew Research Center. https://www.pewresearch.org/internet/2025/07/31/online-scams-and-attacks-in-america-today/

Hancock, J. T., & Toma, C. L. (2009). Putting your best face forward: The accuracy of online dating photographs. *Journal of Communication*, *59*(2), 367–386. https://doi.org/10.1111/j.1460-2466.2009.01420.x

Hart, J. L. (2022). Deception, honesty, and professionalism: A persistent challenge in modern medicine. *Current Opinion in Psychology*, *47*, 101434. https://doi.org/10.1016/j.copsyc.2022.101434

Heiding, F., Schneier, B., Vishwanath, A., Bernstein, J., & Park, P. S. (2024). Devising and detecting phishing emails using large language models. *IEEE Access*, *12*, 42131–42146. https://doi.org/10.1109/ACCESS.2024.3375882

Hertwig, R., & Mazar, N. (2022). Toward a taxonomy and review of honesty interventions. *Current Opinion in Psychology*, *47*, 101410. https://doi.org/10.1016/j.copsyc.2022.101410

Hill, C. E., Thompson, B. J., Cogar, M. C., & Denman, D. W., III (2001). *Beneath the surface of long-term therapy: Therapist and client report of their own and each other's covert processes* (p. 167). American Psychological Association.

Huddle, T. S. (2012). Honesty is an internal norm of medical practice and the best policy. *The American Journal of Bioethics*, *12*(3), 15–17. https://doi.org/10.1080/15265161.2012.656812

Huppert, E., Herzog, N., Landy, J. F., & Levine, E. (2023). On being honest about dishonesty: The social costs of taking nuanced (but realistic) moral stances. *Journal of Personality and Social Psychology*, *125*(2), 259–283. https://doi.org/10.1037/pspa0000340

Johnson, R. A. (2003). *Whistleblowing: When it works–and why*. Lynne Rienner Publishers.

Jotkowitz, A., Glick, S., & Gezundheit, B. (2006). Truth-telling in a culturally diverse world. *Cancer Investigation*, *24*(8), 786–789. https://doi.org/10.1080/07357900601063972

Kelly, S. M., & Lung, N. (2025, October 22). *Tinder mandates face verification for new users to fight spam, fraud*. Dallas News. https://www.dallasnews.com/business/technology/2025/10/22/tinder-mandates-face-verification-for-new-users-to-fight-spam-fraud/

LaFollette, H., & Graham, G. (1986). Honesty and intimacy. *Journal of Social and Personal Relationships*, *3*(1), 3–18. https://doi.org/10.1177/0265407586031001

Le, B. M., Chee, P. X., Shimshock, C. J., & Le, J. D. V. (2025). Expressed and perceived honesty benefits relationships even when couples are not accurate. *Social Psychological and Personality Science*, *16*(8), 907–918. https://doi.org/10.1177/19485506241312876

Le, B. M., Chopik, W. J., Shimshock, C. J., & Chee, P. X. (2022). When the truth helps and when it hurts: How honesty shapes well-being. *Current Opinion in Psychology*, *46*, 101397. https://doi.org/10.1016/j.copsyc.2022.101397

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., & Lupoli, M. J. (2022). Prosocial lies: Causes and consequences. *Current Opinion in Psychology*, *43*, 335–340. https://doi.org/10.1016/j.copsyc.2021.08.006

Levine, E. E., & Munguia Gomez, D. (2020). I'm just being honest." when and why honesty enables help versus harm. *Journal of Personality and Social Psychology*, *120*(1), 33. https://doi.org/10.1037/pspi0000242

Levine, E. E., Roberts, A. R., & Cohen, T. R. (2020). Difficult conversations: Navigating the tension between honesty and benevolence. *Current Opinion in Psychology*, *31*, 38–43. https://doi.org/10.1016/j.copsyc.2019.07.034

Levine, E. E., Wallace, L., Kim, Y., & Perry, S. (2025). *The experience-outcome model of honest conversation: A theory of difficult conversations across situations, people, and time*. https://osf.io/5n7h4_v2

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Mann, H., Garcia-Rada, X., Houser, D., & Ariely, D. (2014). Everybody else is doing it: Exploring social transmission of lying behavior. *PLOS ONE*, *9*(10), e109591. https://doi.org/10.1371/journal.pone.0109591

Miller, C. B. (2021). *Honesty: The philosophy and psychology of a neglected virtue*. Oxford University Press.

Mohammed, F. (2005). Honesty is the best policy. *BMJ*, *331*(7529), 1406. https://doi.org/10.1136/bmj.331.7529.1406

Newall, M. (2023, January 12). *Nearly 1 in 3 Americans report being a victim of online financial fraud or cybercrime*. Ipsos (on Behalf of Wells Fargo. https://www.ipsos.com/en-us/nearly-1-3-americans-report-being-victim-online-financial-fraud-or-cybercrime

Park, H., & Lewis, D. (2018). The negative health effects of external whistleblowing: A study of some key factors. *Social Science Journal*, *55*(4), 387–395. https://doi.org/10.1016/j.soscij.2018.04.002

Pellegrino, E. D. (2002). Professionalism, profession and the virtues of the good physician. *The Mount Sinai Journal of Medicine, New York*, *69*(6), 378–384.

Presiado, M., Montero, A., Lopes, L., & Hamel, L. (2024). *KFF Health misinformation tracking poll: Artificial intelligence and health information*. KFF. https://www.kff.org/public-opinion/kff-health-misinformation-tracking-poll-artificial-intelligence-and-health-information/

Reynolds, C. J., Jayawickreme, E., Wheat, R., Stokes, E., Santos, C., Fleeson, W., & Furr, R. M. (2025). Honesty as truthfulness: New trait and state measures of truthfulness to advance research and theorizing on when and why people are honest. *European Journal of Personality*. Advanced online publication. https://doi.org/10.1177/08902070251338199

Shalvi, S., Levine, E. E., Thielmann, I., Jayawickreme, E., Van Rooij, B., Teodorescu, K., Schurr, A., Furr, R. M., Aglioti, S. M., Zettler, I., Cohen, T. R., Pittarello, A., Barkan, R., Köbis, N., Leib, M., Mitkidis, P., Schulz, J., Dimant, E., van Kleef, G. A., & Ritov, I. (2025). Chapter five - the science of honesty: A review and research agenda. In B. Gawronski (Ed.), *Advances in experimental social psychology* (Vol. 72, pp. 241–327). Academic Press.

Su, Z., Zhou, X., Rangreji, S., Kabra, A., Mendelsohn, J., Brahman, F., & Sap, M. (2025). AI-LieDar: Examine the trade-off between utility and truthfulness in LLM agents. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies: Vol. 1: Long Papers* (pp. 11867–11894).

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Swann, W. B., Stein-Seroussi, A., & Giesler, R. B. (1992). Why people self-verify. *Journal of Personality and Social Psychology*, *62*(3), 392–401. https://doi.org/10.1037/0022-3514.62.3.392

Szymczak, J. E. (2016). Infections and interaction rituals in the organisation: Clinician accounts of speaking up or remaining silent in the face of threats to patient safety. *Sociology of Health & Illness*, *38*(2), 325–339. https://doi.org/10.1111/1467-9566.12371

ten Brinke, L., Lee, J. J., & Carney, D. R. (2015). The physiology of (dis) honesty: Does it impact health? *Current Opinion in Psychology*, *6*, 177–182. https://doi.org/10.1016/j.copsyc.2015.08.004

Tenbrunsel, A. E., & Messick, D. M. (2004). Ethical fading: The role of self-deception in unethical behavior. *Social Justice Research*, *17*(2), 223–236. https://doi.org/10.1023/B:SORE.0000027411.35832.53

Thielmann, I., Zimmermann, J., Leising, D., Hilbig, B. E., & Back, M. (2017). Seeing is knowing: On the predictive accuracy of self– and informant reports for prosocial and moral behaviours. *European Journal of Personality*, *31*(4), 404–418. https://doi.org/10.1002/per.2112

U.S. House Committee on Financial Services. (2025). *Digital assets, financial technology, and artificial intelligence Subcommittee analyzes how financial services regulators and firms are using AI*. https://financialservices.house.gov/news/documentsingle.aspx?DocumentID=410883

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, *6*(1), 2056305120903408. https://doi.org/10.1177/2056305120903408

Weziak-Bialowolska, D., Bialowolski, P., & Niemiec, R. M. (2021). Being good, doing good: The role of honesty and integrity for health. *Social Science and Medicine*, *291*(August), 114494. https://doi.org/10.1016/j.socscimed.2021.114494

Wineburg, S., & McGrew, S. (2019). Lateral Reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, *121*(11), 1–40. https://doi.org/10.1177/016146811912101102

Yang, H., Low, A. E., & Chopik, W. J. (2025). Fibbing friends: Self and friend perceptions of honesty and honesty-adjacent characteristics. *Journal of Research in Personality*, *116*, 104612. https://doi.org/10.1016/j.jrp.2025.104612

Zakrzewski, C. (2023). *FTC investigates OpenAI over data leak and ChatGPT's inaccuracy*. The Washington Post. https://www.washingtonpost.com/technology/2023/07/13/ftc-openai-chatgpt-sam-altman-lina-khan/