

It Might Become True: How Prefactual Thinking Licenses Dishonesty

Beth Anne Helgason¹

Daniel A. Effron¹

¹ Department of Organisational Behaviour, London Business School

Article in Press at Journal of Personality and Social Psychology

July 12th, 2022

Corresponding author: Beth Anne Helgason (bhelgason@london.edu)

Address: London Business School, Sussex Place, Regent's Park, London NW1 4SA, England

© 2022, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/pspa0000308

Abstract

In our “post-truth” era, misinformation spreads not only because people believe falsehoods, but also because people sometimes give dishonesty a moral pass. The present research examines how the moral judgments that people form about dishonesty depend not only on what they know to be true, but also on what they imagine might *become* true. In six studies ($N = 3,607$), people judged a falsehood as less unethical to tell in the present when we randomly assigned them to entertain *prefactual thoughts* about how it might become true in the future. This effect emerged with participants from 59 nations judging falsehoods about consumer products, professional skills, and controversial political issues – and the effect was particularly pronounced when participants were inclined to accept that the falsehood might become true. Moreover, thinking prefactually about how a falsehood might become true made people more inclined to share the falsehood on social media. We theorized that, even when people recognize a falsehood as factually incorrect, these prefactual thoughts reduce how unethical the falsehood seems by making the broader meaning that the statement communicates, its *gist*, seem truer. Mediation evidence was consistent with this theorizing. We argue that prefactual thinking offers people a degree of freedom they can use to excuse lies, and we discuss implications for theories of mental simulation and moral judgment.

KEYWORDS: lying, misinformation, mental simulation, prefactual thinking, moral judgment

Abstract: 218 words

It Might Become True: How Prefactual Thinking Licenses Dishonesty

Theranos founder Elizabeth Holmes raised over \$700 million for her company based on a lie. The technology that she claimed could run hundreds of medical tests with only a drop or two of blood never existed (Carreyrou, 2018). Yet Holmes may have felt certain that it *would* exist eventually if she kept trying. Entrepreneurship, she said, requires the attitude, “We will fail over a thousand times till we get this thing to work, but we will get it on the 1001st time” (Parloff, 2014). The possibility that her technology *might exist in the future* does not make her lie that it *currently* exists any truer. Yet one wonders whether reflecting on this possibility helped Holmes and others perceive the lie as less unethical.

Leaders and organizations frequently make claims that are verifiably false in the present, but that could conceivably become true in the future—at least with a bit of imagination. Consider Donald Trump’s false statement that COVID-19 testing was available to all Americans in March of 2020 (Trump, 2020). Despite learning that this claim was false, some Americans might have imagined that if the government’s response to the coronavirus crisis is executed successfully, then COVID-19 testing will eventually become universally available.

The present research examines whether imagining how a falsehood might become true in the future leads people to judge it as less unethical to tell in the present. In other words, we examine how moral judgments are shaped by *prefactual thinking* (Epstude et al., 2016). We propose that when people imagine how a falsehood might become true, they perceive the falsehood’s broader meaning—its *gist*—as truer, even though they recognize that its specific details are factually inaccurate. For example, a person might recognize that Trump’s specific claim about the availability of COVID-19 tests is false, but—after imagining how this claim might become true in the future—perceive the claim’s broader meaning as more truthful (e.g.,

the government is making progress in the fight against COVID-19). As a result, they may think the falsehood deserves less moral condemnation.

These moral judgments of misinformation are important. The more morally acceptable people think it is to tell a particular falsehood, the less accountable they will hold the people who tell it, and the more inclined they will be to spread the falsehood themselves (Effron, 2018; Effron & Raj, 2020). More broadly, commentators worry that letting people off the hook for dishonesty contributes to a “post-truth” era that undermines trust in society and institutions (Keyes, 2004; The Economist, 2016).

Because nobody knows what will come to pass, people have the flexibility to imagine the future as they like. Thus, prefactual thinking is ripe for motivated reasoning. We suggest that when people want to imagine that a falsehood might become true—because this prediction fits with their pre-existing motivations and beliefs—then prefactual thinking will have a greater effect on reducing their condemnation of the falsehood now. For example, when Trump made false claims about the wide availability of COVID-19 tests, the potential for these tests to become widely available in the future fit with his supporters’ pre-existing belief that he was successfully managing the pandemic, rather than his opponents’ pre-existing belief that he was not. Thus, Trump’s supporters would more easily imagine a prefactual world in which Trump successfully makes COVID-19 tests available to all than would his opponents. In that case, imagining the future would reduce how much Trump’s supporters condemned his falsehood, but have a smaller effect on condemnation from opponents.

In summary, we propose that imagining how a falsehood might become true in the future makes the broader meaning the falsehood conveys seem more truthful, and therefore the falsehood seem less unethical to tell. Moreover, we propose that this effect will be more

pronounced when the prefactual fits, rather than conflicts, with individuals' motivations and beliefs because individuals more easily imagine how the falsehood might become true.

We aim to make three main theoretical contributions. First, whereas prior theorizing emphasizes that prefactual thinking is adaptive because it helps us plan for the future (Bagozzi et al., 2004; Epstude et al., 2016; Hammell & Chan, 2016), we reveal that prefactual thinking has a dark side in that it encourages us to excuse dishonesty in the present. Second, our findings demonstrate that people judge the morality of falsehoods based not only on the facts they know (Levine & Schweitzer, 2014), but also on the scenarios they can imagine. In so doing, our research offers a new perspective on how moral judgment depends on *mental simulation* – “the act of imagination and the generation of alternative realities” (Markman, Klein, & Suhr, 2012, vii). Third, we contribute to research on *moral flexibility* – a tendency to apply moral standards inconsistently to reach desired conclusions (Bartels, 2008; Bartels et al., 2015; Gino, 2016; Gino & Ariely, 2012; Uhlmann et al., 2009). We argue that people's freedom to imagine the future as they want helps them to excuse lies that fit with their partisan beliefs. In this way, mental simulation facilitates moral flexibility.

The next sections develop our hypotheses by drawing on theorizing about moral judgments, prefactual thinking, and the distinction between recognizing a statement is verbatim false and perceiving its gist is true. We also link our theorizing to prior work on how counterfactual thinking about the past affects moral judgments. We then test our predictions across six experiments in the context of falsehoods about products, professional skills, and politics.

People Judge a Falsehood's Morality Based on the Truthfulness of its Gist

Scholars have long argued that it is unethical to tell falsehoods (Aquinas, 1273; Harris, 2013; Kant, 1797; St Augustine, 420). From as young as four years old, we recognize and condemn lies (Wimmer et al., 1984). Moreover, individuals and groups will incur significant costs to punish dishonesty (Boles et al., 2000; Brandts & Charney, 2003; Keck, 2014; Ohtsubo et al., 2010).

Yet people do not judge all falsehoods as equally wrong (Effron & Raj, 2020; Levine & Schweitzer, 2014; Rogers et al., 2017; Schweitzer & Hsee, 2002). For example, people are more willing to excuse lies that are told with benevolent intentions; people sometimes even judge benevolent lies as more ethical than truths (Levine & Schweitzer, 2014). Moreover, people feel less compunction about lying by omission rather than commission (Levine et al., 2018). Thus, just because people recognize a falsehood as such does not mean they will harshly condemn it.

We propose that when deciding how harshly to condemn a falsehood, people also consider the truthfulness of its *gist*—the broader meaning that the statement communicates (Abadie et al., 2013, 2017, 2021; Brainerd & Reyna, 1990; Fukukura et al., 2013; Reyna & Brainerd, 1991; Stahl & Klauer, 2008). We adopt the term *gist* from cognitive psychology, where it is used to describe the “essence of information” or “fundamental meaning” in a statement (Reyna, 2020). A statement’s *gist* is often contrasted with its verbatim details—the exact words or numbers the statement communicates (Abadie et al., 2017; Abadie & Camos, 2018; Reyna & Brainerd, 1995; Reyna, 2020). For example, in the claim, “Eating yogurt reduces the risk of catching the common cold by 30%,” the verbatim details include the words “common cold” and the figure “30%,” while the *gist* is the broad idea that “Yogurt is healthy.” Past research suggests that even when people know the verbatim details of a statement, they rely on *gist*

representations—often more than verbatim details—when making decisions (Reyna, 2004, 2012).

Whereas prior research in cognitive psychology has examined how *memory* for gist explains patterns of memory retention, false memory, and complex decision making (Abadie et al., 2013, 2017; Adolphs et al., 2005; Bookbinder & Brainerd, 2017; Brainerd & Gordon, 1994; Brainerd & Reyna, 2002; Friedman, 1979; Fukukura et al., 2013; Reyna, 2012; Stahl & Klauer, 2008), we introduce the gist construct to the moral judgment literature to examine how *perceptions of the gist's truthfulness* explain why people think a falsehood is less unethical when they imagine how the falsehood might become true. Whereas perceptions of verbatim truthfulness reflect people's views of the precise and literal meaning of a statement, perceptions of gist truthfulness reflect people's views about the broader meaning that a statement conveys. Most research on moral judgments of falsehoods has focused on statements that are both verbatim false *and* convey a clearly false gist (e.g., Levine & Schweitzer, 2014; Schweitzer et al., 2006; Tyler et al., 2006). Yet, perceptions of verbatim truthfulness and gist truthfulness may diverge. People may perceive that a statement that is literally true conveys a false broader meaning (Rogers et al., 2017). Consider Bill Clinton's claim that "there is not a sexual relationship" between himself and Monica Lewinsky (Lehrer, 1998). Although his statement was literally true—he and Lewinsky were not *currently* having a sexual relationship at the time Clinton spoke—many people nonetheless perceived that it conveyed a false message, namely that there *had never been* such a relationship. People are quick to condemn such "palters," although these statements are verbatim truthful, because they communicate a false gist (Rogers et al., 2017).

In the present research, we propose the converse—that people are more willing to condone a *verbatim false* statement when they believe the statement conveys a *truthful gist*. Consider Donald Trump’s claim that “the U.S. trade deficit with China is currently \$500 billion” (Trump, 2015). Although his statement was literally false—the trade deficit with China was estimated to be only \$336 billion at the time Trump spoke (Morrison, 2018)—people may have perceived that it conveyed a broader meaning that is true, namely that the U.S. trade deficit with China was large. More generally, people may perceive that a verbatim false statement communicates a truthful gist. We argue that the more truthful a statement’s gist seems, the less unethical people will think the statement is—even if they know the statement is literally false.

Our proposition that people are more willing to condone literally false statements when they perceive the statement’s gist as truthful fits with theorizing on conversational norms. Conversational norms encourage listeners to infer the gist that the speaker intends to communicate—not just the verbatim details of the words they use (Grice, 1975). For example, listeners are rarely bothered by idiomatic or sarcastic statements that are literally false (e.g., “I have been working 24 hours a day”; “What beautiful weather we are having with this rain”), because they understand the speaker’s broader meaning (Clark & Lucy, 1975; Gibbs, 1980, 1987). We propose that people judge the ethicality of a statement not only based on the verbatim truth of the words, but also based on how truthful they perceive the gist.

In the next section, we argue that imagining how a literally false claim might become true in the future makes its gist seem more truthful.

Prefactuals Make a Falsehood’s Gist Seem Truer

Prefactuals are conditional propositions about what might occur in the future: *If X occurs, then Y will occur* (Epstude et al., 2016). Research in cognitive, social, developmental, and

applied psychology has shown that the human ability to imagine different plausible futures – i.e., *prefactual thinking* – plays important roles in goal pursuit, motivation, and affect regulation (Bacon et al., 2020; Bagozzi et al., 2004; Epstude et al., 2016; Goerke et al., 2004; Hammell & Chan, 2016; Scholl & Sassenberg, 2015). We propose, however, that prefactual thinking can have a darker consequence: encouraging people to excuse dishonesty. The reason, we argue, is that imagining how a falsehood might become true in the future makes the gist of the falsehood seem truer in the present. And as noted, a falsehood will seem less unethical to tell if people perceive its gist has truth to it, even if its verbatim details are false.

Why would prefactual thinking about a falsehood make its gist seem truer? We propose that confirmation bias plays a key role. When people consider a hypothesis, they spontaneously seek confirmatory evidence (Klayman & Ha, 1987; Snyder & Swann, 1978, 1978; Trope & Liberman, 1996). Considering an event in the past leads people to think about reasons why it “had to occur,” which creates hindsight bias (Carli, 1999; Hawkins & Hastie, 1990; Hoffrage et al., 2000; Louie, 2005; Roese & Vohs, 2012); exposure to a numerical anchor in the present makes anchor-consistent information more cognitively accessible (Chapman & Johnson, 1999; Pohl et al., 2003; Strack & Mussweiler, 1997); and – most relevant to our research – imagining a possible future event leads people to search their knowledge for evidence that the event will actually occur (Carroll, 1978; Gregory et al., 1982; Koehler, 1991; Sherman et al., 1985). Thus, we propose that when people imagine that a falsehood might become true in the future, they recruit evidence suggesting that it will actually become true.

This evidence, we argue, not only makes the falsehood seem likely to become true in the future, but also makes the gist of the falsehood seem truer now. To illustrate this process, consider again the falsehood, “The U.S. trade deficit with China is currently \$500 billion,” the

gist of which is simply that this trade deficit is large. Imagining the prefactual, “If Congress blocks the imposition of tariffs on Chinese goods, then the U.S. trade deficit with China will grow to \$500 billion next year” may bring to mind evidence consistent with that prefactual, such as, “I heard that the deficit has recently increased,” “I see lots of ‘Made in China’ labels on goods sold in America,” or “the U.S. and China have been in a trade war.” This confirmatory evidence tends to not only support the prefactual, but also support the broader meaning communicated by the falsehood. For example, the idea that “the deficit has recently increased” is consistent not only with the deficit growing further in the future (the prefactual), but also with the broader message that the deficit is currently large (the falsehood’s gist). In this way, thinking prefactually about how the falsehood *might become true* in the future will make its gist seem truer in the present.

Finally, if the falsehood’s gist seems truer, the falsehood should feel less unethical to tell – even if people still recognize its verbatim details as false. For example, a person might think, “I know the deficit is not actually \$500 billion, but the deficit is probably still pretty large – so claiming that it’s \$500 billion isn’t so dishonest.” A literally false statement will seem less unethical if its gist seems true.

To summarize, we argue that imagining a prefactual in which a falsehood might become true in the future brings to mind evidence consistent with the prefactual. This evidence makes the falsehood’s gist seem truer – and as a result, the falsehood seems less unethical to tell, even though people may still recognize it as literally false. Thus, we formulated the following hypotheses:

Hypothesis 1 (H1). Considering how a falsehood might become true in the future will lead people to judge the falsehood as less unethical to tell in the present.

Hypothesis 2 (H2). Considering how a falsehood might become true will lead people to judge the falsehood as less unethical to tell in the present by making the gist of the falsehood seem truer.

The Role of Pre-Existing Motivations and Beliefs

Given that people have flexibility to imagine different futures, we expect that motivated reasoning will amplify the effect of prefactual thinking on moral judgments of falsehoods. Specifically, prefactual thinking should reduce condemnation of falsehoods to a greater extent among people who are already inclined to accept that the falsehood might become true. For example, Donald Trump's false claim that COVID-19 tests were available to all Americans in March, 2020 would not seem so unethical to a person who imagines how such testing might become available in the future – particularly if that person were a Trump supporter. Compared to Trump opponents, Trump supporters should be more inclined to believe that Trump can successfully roll out universal testing, making tests available to all. In this way, people's pre-existing motivations and beliefs may amplify the effect of prefactual thinking on moral judgments.

There are two reasons why such motivations and beliefs might have this amplifying effect. First, events are easier to mentally simulate if they fit with what we already want and believe (Tetlock, 1998; Tetlock & Henik, 2007). In past research, experts of world politics found it easier to imagine counterfactuals that fit with their pre-existing beliefs about history (Tetlock, 1998; Tetlock & Lebow, 2001) and religious fundamentalists found it easier to imagine counterfactuals that did not conflict with their religious beliefs (Tetlock et al., 2000). In contrast, when people are asked to imagine events that conflict with their pre-existing beliefs, they often generate counterarguments about why the imagined event could not have occurred (Tetlock & Visser, 2000). Most relevant to the present research, American partisans found it easier to

imagine counterfactuals about how a falsehood *could have been true* when that falsehood aligned with their politics (Effron, 2018). Given the similarities between mental simulation about the past (e.g., counterfactual thinking) and the future (e.g., prefactual thinking; Kappes & Morewedge, 2016), we similarly expect that people will find it easier to imagine prefactuals about how a falsehood *might become true* if that possibility fits with their pre-existing motivations and beliefs.

Second, the easier a prefactual is to imagine, the larger an effect it has on judgments and decisions (Petrocelli et al., 2012; see also Petrocelli et al., 2011). People's judgments are disproportionately affected by events that easily come to mind, regardless of whether these events are real or imagined, because easily imagined events increase the cognitive accessibility of supporting evidence (Kappes & Morewedge, 2016; Tversky & Kahneman, 1973). Thus, just as easily recalled past events have disproportionate effects on people's judgments, so too will easily imagined future events (MacInnis & Price, 1987).

In one classic example, instructing people to imagine contracting a disease increased how likely they thought they were to actually contract it—but only if the disease had symptoms that were easy to imagine (Sherman et al., 1985). More recent research has examined ease of imagination through the perceived *vividness* or *plausibility* of imagined events, with analogous results. For example, the more vividly people imagined taking an action, the greater their willingness to take that action in the future (Gaesser et al., 2018, 2019). Similarly, mentally simulating an event has a larger effect on judgments and decisions when people think the event is plausible (Petrocelli et al., 2011, 2012). In the context of falsehoods, the more plausible people thought it was that a falsehood could have been true, the less unethical they thought the falsehood was to tell (Effron, 2018). Thus, we expected that the easier a prefactual is to imagine,

the larger an effect it should have on moral judgments of falsehoods. Consistent with past theorizing, we measured ease of imagination as participants' ratings of the plausibility or vividness of the prefactual.

To summarize, prefactuals are easier to imagine when they fit with a person's motivations and beliefs, and easily imagined prefactuals should have a larger effect on moral judgments. Therefore, we hypothesized:

Hypothesis 3 (H3). Considering a prefactual about how a falsehood might become true will have a greater effect on reducing condemnation of the falsehood when that prefactual fits with their pre-existing motivations and beliefs.

Advancing Theory on Mental Simulation and Moral Judgment

By investigating how prefactual thinking affects moral judgments of falsehoods, we build on research on counterfactual thinking – thoughts of “what might have been” (e.g., Roese, 1997). Like prefactuals, counterfactuals are a form of mental simulation in that they involve imagining alternatives to reality. But whereas prefactuals are simulations of how an event might occur in the future, counterfactuals are simulations of how it could have occurred in the past.

Several studies speak to the important role counterfactual thinking plays in moral judgments (see Byrne, 2017). People let themselves off the hook when they imagine bad deeds they *could have done* but chose not to do (Effron et al., 2012, 2013). They judge another person's character not only based on the lies he told, but also based on the lies they imagine he *would have* told if given the chance (Miller et al., 2005). And they are less inclined to punish someone for wrongdoing if they imagine the wrongdoing “could have been worse” (Markman et al., 2008).

Most relevant to the present research, prior work shows that lies seem more morally permissible when it is easy to imagine that they *could have been* true if past circumstances had

been different. For example, Donald Trump's supporters thought it was less unethical to falsely claim that his 2016 inauguration crowd was larger than Barack Obama's in 2008 when they considered the counterfactual *Donald Trump's inauguration would have been bigger if the weather had been nicer* (Effron, 2018). Individuals who exhibit a greater tendency to spontaneously generate counterfactual thoughts also exhibit a greater propensity to tell falsehoods (Briazu et al., 2017). Moreover, people are more likely to lie about the roll of a die if they have previously observed – or merely imagined – the winning roll (Lelieveld et al., 2016; Shalvi et al., 2011; Shalvi & Leiser, 2013). Apparently, imagining a counterfactual about how a falsehood could have been true makes it seem more justified (Shalvi et al., 2011).

We aim to advance this prior work in two key ways. First, we argue that it is not counterfactual thinking per se, but mental simulation more broadly, that increases people's inclination to excuse falsehoods. Regardless of whether the mental simulation involves undoing the past or predicting the future, the mere act of imagining a falsehood being true should bring to mind information consistent with the gist of the falsehood. Thus, our research on prefactual thinking and morality expands the scope of prior theorizing about counterfactual thinking and morality.

Second, we theorize and test why mentally simulating a falsehood increases people's inclination to excuse it. As noted, prior research shows that counterfactual thinking makes falsehoods seem more justifiable or less unethical (e.g., Effron, 2018; Shalvi et al., 2011). However, previous research has not examined *why* counterfactual thinking has this effect. We suggest that mentally simulating a falsehood brings to mind evidence that makes the gist—or broader meaning communicated by the falsehood—seem truer. And, even if a statement is

literally false, it will not seem so unethical to tell if its gist has some truth. Our studies test whether prefactuals increase beliefs in a statement's gist, which in turn predict moral judgments.

The Present Research

Six studies (four pre-registered) tested our hypotheses. We showed participants a variety of statements, clearly identified as false, and manipulated whether participants reflected on prefactuals about how the falsehoods might become true in the future. In some studies, participants considered pre-written prefactuals; in other studies, they wrote their own. Our central prediction was that participants would rate the falsehoods as less unethical to tell if they had reflected prefactually on how the falsehoods might become true in the future (H1).

Study 1 sought to establish this basic effect with moral judgments of false advertisements about consumer products. Study 2 aimed to replicate the effect with moral judgments of professionals' lies on resumes, and to test whether perceptions of the falsehood's gist truthfulness mediated this effect (H2).

To test our prediction that the effect would be stronger when the prefactuals fit with participants' pre-existing motivations and beliefs (H3), Studies 3–6 asked political partisans to judge falsehoods that either Democrats or Republicans would be more inclined to accept might become true. We expected that imagining how these falsehoods might become true would have a stronger effect on judgments of falsehoods when the prefactuals fit with participants' politics, and a weaker effect when the prefactuals conflicted with participants' politics. Studies 4 and 5 also tested a potential downstream consequence of judging falsehoods as less unethical – increased willingness to promote the falsehoods oneself on social media. Finally, Studies 5 and 6 provided further evidence for the hypothesized role of gist judgments (H2).

Open Practices

We pre-registered Studies 1, 4, 5, and 6. Pre-registration documents, stimulus materials, data, and code are posted on the Open Science Framework (OSF; https://osf.io/p3y8r/?view_only=a520d5798855469081c4480072ee1f8a).¹

Study 1

Study 1 tested whether prefactual thinking could make falsehoods seem less unethical (H1). Participants judged a series of falsehoods, adapted from actual false advertisements about household products. By random assignment, half the participants considered a prefactual about how each falsehood might become true in the future. Then all participants indicated how unethical they found each falsehood.

Method

Study 1 had a two-cell, between-participants design. We pre-registered the study at <https://aspredicted.org/blind.php?x=7ap62y>

Participants. Participants were 276 residents of a large U.K. city who were enrolled in a behavioral lab's subject pool, and who signed up to complete an hour-long series of studies from multiple researchers, in which Study 1 was embedded. Based on the size of the participant pool and budgetary considerations, the lab-standard practice for these sessions is to aim for 250 participants, posting more timeslots to guard against no-shows. We received responses from 276 people (168 women, 104 men, 1 non-binary, 3 did not report; $M_{age} = 32$, $SD = 13$), and did not exclude any observations in our analyses. Of the 276 participants, 23 identified as Black, 29 as East Asian, 5 as Middle Eastern, 67 as South Asian, 120 as White, 29 as "Other," and 3 chose not to report their ethnicity.

¹ The Research Ethics Committee overseeing this work withheld permission to post Study 2's data because we did not explicitly obtain participants' consent to post it. However, Study 2's data are available from the corresponding author upon request.

Statistical power. A sensitivity analysis showed that the smallest effect size this sample size could detect with $> 80\%$ power, two-tailed $\alpha = .05$, and the study's 8 repeated measures was $d = .22$, according to the PANGAEA web app (see <https://jakewestfall.shinyapps.io/pangea/> and Westfall, 2016). This effect size approximates the average effect of counterfactual thinking on moral judgments across several studies with similar paradigms (i.e., $d = .18$; Effron, 2018). We conclude that Study 1 was adequately powered to detect an effect size we could reasonably expect in this paradigm. (Study 1's actual statistical power was somewhat higher than 80% because PANGAEA did not allow us to specify that we had pre-registered a one-tailed test of our directional hypothesis).

Materials. The stimuli were eight falsehoods about household products (e.g., "Gerber's good start formula prevents allergies in children"). We created the falsehoods based on false claims actually used in advertising these products (see Appendix A). For each falsehood, we wrote a fact that contradicted the falsehood (e.g., "Gerber's good start formula does NOT prevent allergies in children") and a prefactual statement about how the falsehood might become true in the future (e.g., "If Gerber develops its good start formula, then it will prevent allergies in children;" see Appendix A for complete materials).

Procedure and measures. Participants sat at private computer cubicles in a large lab room. The studies in the session were presented in random orders. First, all participants read a fact about one of the household products ("It has been shown that..."). Then, participants randomly assigned to the *prefactual condition* were shown a prefactual about how a falsehood contradicting that fact might become true in the future. We described the prefactual as a "prediction." To encourage engagement with the manipulation and to communicate that whether the falsehood would or would not become true was uncertain, participants rated the likelihood of

that the prediction (1 = *Not at all likely* to 11 = *Extremely likely*). Participants in the control condition neither saw nor judged a prefactual.

Next, for the dependent measure, participants responded to a three-item measure of their perceptions that the falsehood was unethical. Participants reported how dishonest, unethical, and acceptable (reverse-coded) it would be to make the false claim by moving a slider initially appearing at the scale midpoint to select a response from 0 (*Not at all*) to 100 (*Extremely*; \checkmark s > .71 for each falsehood). We predicted lower ratings on this measure in the prefactual condition than in the control condition.

Participants repeated this procedure for all eight falsehoods, presented in randomized orders. Then they responded to a fact-check measure to test that they acknowledged that the falsehoods were indeed false. Specifically, for each product, we showed participants either the fact or the falsehood they had seen earlier, and participants had to categorize it as true or false. (Three statements were facts; five were falsehoods). The purpose of this measure was to ensure that participants believed the facts we presented and to rule out the possibility that prefactuals could lead people to forget that the falsehoods were indeed false (cf. Gerlach et al., 2014; Petrocelli & Crysel, 2009). Thus, we expected that our main results would be robust after excluding responses to items on which participants did not correctly identify fact from falsehood.

At the end of the study session, participants provided basic demographics.

Results

We had strong directional predictions so we pre-registered one-tailed significance tests to increase statistical power (Cho & Abe, 2013).

Unethicality of telling the falsehoods. On average, across conditions, participants thought it was moderately unethical to tell the falsehoods ($M = 72.54$ on a 100-point scale, $SD =$

19.92). More importantly, we observed a condition difference that supported H1. Participants who considered prefactuals about how the falsehoods might become true in the future thought they were less unethical to tell ($M = 70.45$, $SD = 19.00$) than participants in the control condition ($M = 74.66$, $SD = 20.66$). This mean difference was significant in a mixed-effect regression model, with condition as a fixed effect (1 = prefactual condition, 0 = control condition), random effects for participants, item fixed effects, and a pre-registered one-tailed test of our directional prediction, $d = -0.21$, $z = -1.78$, $p = .038$. This mixed-effect model accounts for the fact that each participant judged eight falsehoods.

Fact-check. We predicted that the prefactual manipulation would lead people to rate falsehoods as less unethical despite knowing that they were false. Our fact-check measure showed that people correctly differentiated fact from falsehood 88% of the time, and this percentage was not statistically distinguishable between prefactual (88.86%) and control (87.32%) conditions, $b = 0.22$, $z = 0.67$, $p = .502$, in a mixed logistic regression analysis with random intercepts for participants (this test was not pre-registered, so the p -value is two-tailed). Moreover, as a pre-registered robustness check we repeated our prior analyses retaining only responses to the dependent measure that corresponded correctly to our fact-checks. Indicating our previous results were robust, the effect of prefactual thinking remained significant and in the same direction, $b = -3.97$, $z = -1.69$, $p = .045$. In addition, the prefactual effect remained significant when we retained all data but statistically controlled for the fact-check measure, $b = -4.33$, $z = -1.88$, $p = .030$. Thus, we found no evidence that prefactual thinking reduced the moral condemnation of falsehoods simply by convincing people the falsehoods were true.

Discussion

Participants in Study 1 thought it was less unethical to tell falsehoods about consumer products when they imagined how these falsehoods might become true (H1). Our fact-check measure suggested that such prefactual thinking did not lead people to confuse fact and fiction; instead, it made them express milder condemnation of falsehoods they knew were factually inaccurate. Although Study 1 did not assess why this effect occurred, we posit that thinking about how the falsehood might become true reduces the condemnation it receives because it leads people to perceive the falsehood's gist as truer, even though its factual claims are literally false (H2). Study 2 tested this mechanism, and assessed generalizability to a new context (lying on a resume) and a new population (international MBA students).

Study 2

Study 2 recruited a sample of MBA students, who were current or soon-to-be job seekers, to examine whether prefactual thinking could make it seem less unethical to make false claims on one's resume (H1). Moreover, we examined whether this effect was mediated by participants' perceptions that the gist of the falsehood was true (H2).

Method

Study 2 had a 2-cell, between-participants design.

Participants. We invited all the MBA students in a required organizational behavior course at a UK business school to complete Study 2 online as part of a class survey. We could not analyze the data from 16 participants because they did not consent, and we dropped 14 responses because they were from participants who had previously submitted a response. The resulting sample was 447 students from 59 different countries (271 men, 163 women, 1 non-binary, 15 unknown gender).

Statistical power. A sensitivity analysis with the PANGEA app showed that, with 447 participants, two repeated measures, and two-tailed $\alpha = .05$, Study 2 offered approximately 80% power to detect an effect size of $d = .2$.

Procedure. Participants read two fictional scenarios about a friend who made a false claim on his resume (e.g., “Imagine you see that a friend in your stream lists financial modelling as a skill on his resume despite the fact that he has no experience with financial modelling.”). Participants randomly assigned to the *prefactual condition* then considered a prefactual statement about how the false claim might become true in the future (e.g., “Now, consider that *if* the same friend enrolls in a financial modelling course that the school offers in the summer, *then he could* develop experience with financial modelling.”). We include emphasis on the “if” and “could” portions of the prediction to highlight that it was not certain that the prefactual would indeed occur in the future. Participants randomly assigned to the *control condition* were not shown a prefactual. Then, participants completed the measures described below and repeated the procedure for a second scenario about a different friend who made a different false claim on his resume (i.e., “lists ‘data visualization’ as a skill on his resume despite the fact that he has no experience with data visualization.”). Participants were in the same condition for both scenarios, and we randomized the order of the scenarios.

Measures.

Dependent measure: unethicity of telling the falsehood. Participants indicated how unethical the relevant falsehood was, using the same 3-item scale from Study 1 ($\alpha > .72$ for each vignette).

Mechanism: gist truthfulness. Participants used four items to rate how true they perceived the gist of the falsehood: “Regardless of whether your friend’s claim that he is skilled

in financial modelling [data visualization] is literally true, how much do you agree or disagree that...” (a) “The larger point is correct.” (b) “The gist of it is true.” (c) “It is true in spirit.” and (d) “The general idea is accurate.” (The bracketed text varied by vignette). Participants indicated their agreement with each statement from -3 (*Strongly disagree*) to $+3$ (*Strongly agree*; $\alpha > .96$ in each vignette).

Results

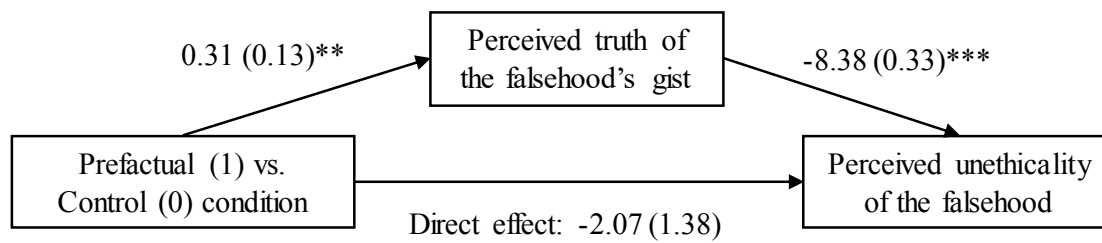
Unethicality of telling the falsehood. Across conditions, participants believed it was moderately unethical to tell the falsehoods ($M = 64.54$ on a 100-point scale, $SD = 19.08$). Replicating the results from Study 1 and supporting H1, participants thought the falsehoods were less unethical when they imagined how the falsehood might become true in the future (prefactual condition: $M = 62.40$, $SD = 19.75$) than when they did not (control condition: $M = 66.99$, $SD = 18.03$). This difference was significant in a mixed-effect regression model, with item and condition as fixed effects (1 = prefactual condition, 0 = control condition), and random intercepts for participants, $d = -0.24$, $z = -2.52$, $p = .012$. Although we had a directional prediction, we report a two-tailed significance test because Study 2 was not pre-registered.

Mediation through gist truthfulness. We predicted that considering a prefactual statement about how the falsehood might become true in the future would lead people to judge it as less unethical to tell by making the falsehood’s broader meaning—its gist—seem truer (H2). Consistent with this hypothesis, our prefactual manipulation had a significant negative indirect effect on unethicality judgments through gist ratings, $b = -2.58$, 95% CI $[-4.72, -0.43]$ (see Figure 1). When we included gist truthfulness in the model predicting unethicality judgments, the direct effect of the prefactual condition was not significant, $b = -2.07$, 95% CI $[-4.76, 0.63]$. We conducted this analysis as a generalized structural equation model with prefactual condition

as the independent variable (1 = prefactual, 0 = control), gist as the mediator, and perceptions of unethicity as the dependent variable, plus random effects for participants and fixed effects for item to account for the data's multilevel structure. We computed the indirect effect by multiplying the *a*- and *b*-paths together (i.e., using the *gsem* and *nlcom* functions in Stata).

Figure 1

Indirect Effect of Prefactual Condition on Unethicality Judgments through Gist in Study 2



Note. Indirect effect: $b = -2.58$, 95% CI $[-4.72, -0.43]$. Unstandardized coefficients shown.

Model includes item fixed effects and participant random effects.

** $p < .01$, *** $p < .001$.

Discussion

Conceptually replicating Study 1's results, Study 2 found that lying on a resume seemed less unethical to MBA students when they imagined how the lie might become true in the future (H1). Furthermore, consistent with our theorizing about mechanism (H2), such imagination made the lies' gist seem truer, which in turn predicted more lenient moral judgments.

Study 3

Studies 1 and 2 provide evidence that considering a prefactual about how a falsehood might become true leads people to judge the falsehood as less unethical to tell in the context of false advertisements and claims about professional skills. In Study 3, we examine whether prefactual thinking reduces condemnation of falsehoods in the context of American politics. One possibility is that partisan effects on moral judgments are so large in this context that prefactual thinking will have little effect (Anduiza et al., 2013; Carlson, 2015; Mueller & Skitka, 2018). In contrast, we propose that political partisanship will amplify the effect of prefactual thinking on moral judgments. Specifically, we propose that the effect of prefactual thinking on condemnation of falsehoods will be larger when the possibility that the falsehood might become true is consistent with what people want to believe (H3). Moreover, we test our theorizing that this moderation effect occurs because prefactuals that are consistent with what people want to believe are easier to imagine. In Study 3, we operationalized ease of imagination as judgments of the prefactual's plausibility; people who find a prefactual easy to imagine should rate it as more plausible (Effron, 2018; Petrocelli et al., 2012).

Method

The study's design was a mixed 2 (condition: prefactual versus control; between-subjects) x 2 (partisan fit of "it might become true": fits versus conflicts with participants' political beliefs; within-subjects) factorial design with six repeated measures.

Participants. We posted spots for 800 American participants on Amazon Mechanical Turk (MTurk), and 889 participants began the study. Participants could not begin the study if they failed a simple reading-comprehension question. We dropped data submitted from duplicate MTurk worker IDs, IP addresses, or geolocations (evidence of fraudulent data; Dennis et al., 2019; Kennedy et al., 2020). We determined these exclusion criteria before testing any of our hypotheses. The resulting sample was 735 participants (322 men, 375 women, 38 missing gender data; $M_{age} = 41$, $SD_{age} = 13$). Of these participants, 706 reported that they considered themselves or leaned Democrat or Republican so were included in our analyses of political fit: 409 individuals reported that they considered themselves or leaned Democrat, 297 individuals reported that they considered themselves or leaned Republican. Another 29 individuals reported they did not lean towards either party.

Statistical power. A sensitivity analysis with the PANGAEA app showed that, with 735 participants, six repeated measures, and two tailed $\alpha = .05$, Study 3 provided $> 99\%$ power to detect an effect size of $d = .2$ of the prefactual manipulation.

Materials. The stimuli were six falsehoods that focused on controversial issues in contemporary American politics (e.g., gun control; immigration; inequality). The falsehoods were based on actual false claims made by politicians and the media (e.g., "The average top CEO currently makes 500 times more money than the average American worker."). For each falsehood, we selected a verified fact that clearly identified the falsehood as untrue (e.g., "It's a proven fact that in 2017, the average top CEO made 265 times more money than the average

American worker”) and we wrote a prefactual statement in which the falsehood might become true in the future (e.g., “If the Trump administration keeps making pro-corporate decisions, then the average top CEO will soon make 500 times more than the average worker;” see Appendix B for all stimuli). Of the six falsehood-prefactual pairs, three fit with political stances associated with Republicans, such that Republicans would be more inclined to accept that the falsehood might become true, and three with political stances associated with Democrats, such that Democrats would be more inclined to accept that the falsehood might become true.

Procedure. Before beginning the study, participants responded to a comprehension check question to promote data quality. Participants saw a list of words (car, clock, dog, rock, stove, guitar, shoe, and painting) and selected the word describing something used to cook. We prevented participants from beginning the study if they failed to select the word “stove.”

At the beginning of the study, participants reported their political affiliation: *Democrat*, *Republican*, *Independent*, or *None of the above*. Participants who responded *Independent* or *None of the above* next reported whether they leaned Democrat, Republican, or towards neither party. We categorized participants as Democrats or Republicans if they self-identified as – or leaned towards – one of these parties.²

Next, participants viewed the fact about one of the political issues (e.g., “It’s a proven fact that in 2017, the average top CEO made 265 times more money than the average American worker.”). Participants randomly assigned to the *prefactual condition* then considered a prefactual statement about how a falsehood that contradicted the fact might become true in the future (e.g. “If the Trump administration keeps making pro-corporate decisions, then the average top CEO will soon make 500 times more than the average worker”) and rated how much they

² We also included a measure of participants’ recent voting behavior for exploratory purposes (see Online Supplement).

accepted this prefactual (see below). Participants in the *control condition* did not see a prefactual. Next, for the dependent measure, all participants rated how unethical it would be to tell a falsehood that contradicted the fact (e.g. “The average top CEO currently makes 500 times more money than the average American worker.”).

Participants repeated this procedure for all six falsehoods, presented in randomized orders. At the end of the study, participants responded to a fact-check measure to see if they recalled that the falsehoods were indeed false (see below). Then participants provided demographic information.

Measures.

Dependent measure: Unethicality of telling the falsehood. Participants rated the unethicality of telling each falsehood on a six-item scale (Effron, 2018). Specifically, they moved a slider to indicate how dishonest, justified (reverse-coded), unethical, acceptable (reverse-coded), and problematic it would be to make the false statement, and how much of a lie they considered the statement to be ($\alpha > .93$; 0 = *Not at all*, 100 = *Extremely*). The slider initially appeared at the scale midpoint.

Ease of imagining the prefactual: Prefactual potency. Recall that the prefactuals were all conditional statements about how *if* an event occurs *then* the falsehood will become true. For each prefactual, participants in the prefactual condition made two plausibility judgments: the likelihood that (a) the *if* and (b) the *then* components of would come true (from 1 = *Not at all likely* to 11 = *Extremely likely*). For example, for the prefactual, “If the Trump administration keeps making pro-corporate decisions, then the average top CEO will soon make 500 times more than the average worker,” participants were asked, “What do you perceive is the likelihood that the Trump administration will keep making pro-corporate decisions?” (*if-likelihood*) and

“Suppose that the Trump administration keeps making pro-corporate decisions. Given that, what do you perceive is the likelihood that the average top CEO will soon make 500 times more than the average American worker?” (*then-likelihood*).

Following prior work on prefactual thinking, we combined judgments of the *if* and *then* likelihoods by multiplying them together into a measure called *prefactual potency* (Petrocelli et al., 2012; see also Petrocelli et al., 2011). More-potent prefactuals tend to have a larger effect on judgments (Petrocelli et al., 2012). Although our *a priori* predictions were about potency (see Effron, 2018), we also explored judgments of the *if* and *then* likelihoods individually.

Fact-check. To ensure that participants correctly distinguished fact from fiction, we asked participants to indicate “true” or “false” to each of six statements – one about each of the political issues. Three statements were the facts and three were the falsehoods that participants had seen earlier. As in Study 1, we expected that prefactual thinking would not affect this fact-check measure, and that prefactual thinking would influence participants’ ratings of falsehoods even when limiting analysis to only correctly identified falsehoods.³

Results

Unethicity of telling the falsehood. Across conditions, participants believed it would be moderately unethical to tell the falsehoods ($M = 71.27$ on a 100-point scale, $SD = 20.51$). In line with H1, participants who thought prefactually about how the falsehoods might become true in the future thought the falsehoods were less unethical to tell ($M = 67.04$, $SD = 19.74$) than participants in the control condition ($M = 75.40$, $SD = 20.44$). This difference was significant in a mixed-effect regression model, with condition as a fixed effect (1 = prefactual condition, 0 =

³ At the end of the survey, we included additional survey questions on participants’ political knowledge for exploratory purposes (see Online Supplement).

control condition), participant random effects, and item fixed effects, $d = -0.42$, $z = -5.74$, $p < .001$ (see Table 1, step 1).

Table 1*Regression Analyses Predicting Unethicality Ratings in Study 3*

Predictor	<i>b</i>	<i>SE(b)</i>	<i>z</i>	<i>p</i>	95% CI of <i>b</i>
Step 1					
Condition	-8.48	1.48	-5.74	< .001	[-11.38, -5.59]
Constant	69.42	1.24	56.13	< .001	[67.00, 71.85]
Step 2					
Condition	-6.22	1.62	-3.84	< .001	[-9.40, -3.04]
Political fit	-3.71	0.86	-4.33	< .001	[-5.39, -2.03]
Condition x political fit	-5.21	1.21	-4.32	< .001	[-7.57, -2.84]
(Constant)	72.00	1.34	53.77	< .001	[69.38, 74.63]

Note. Condition is coded 1 = prefactual, 0 = control. Political fit with “it might become true”

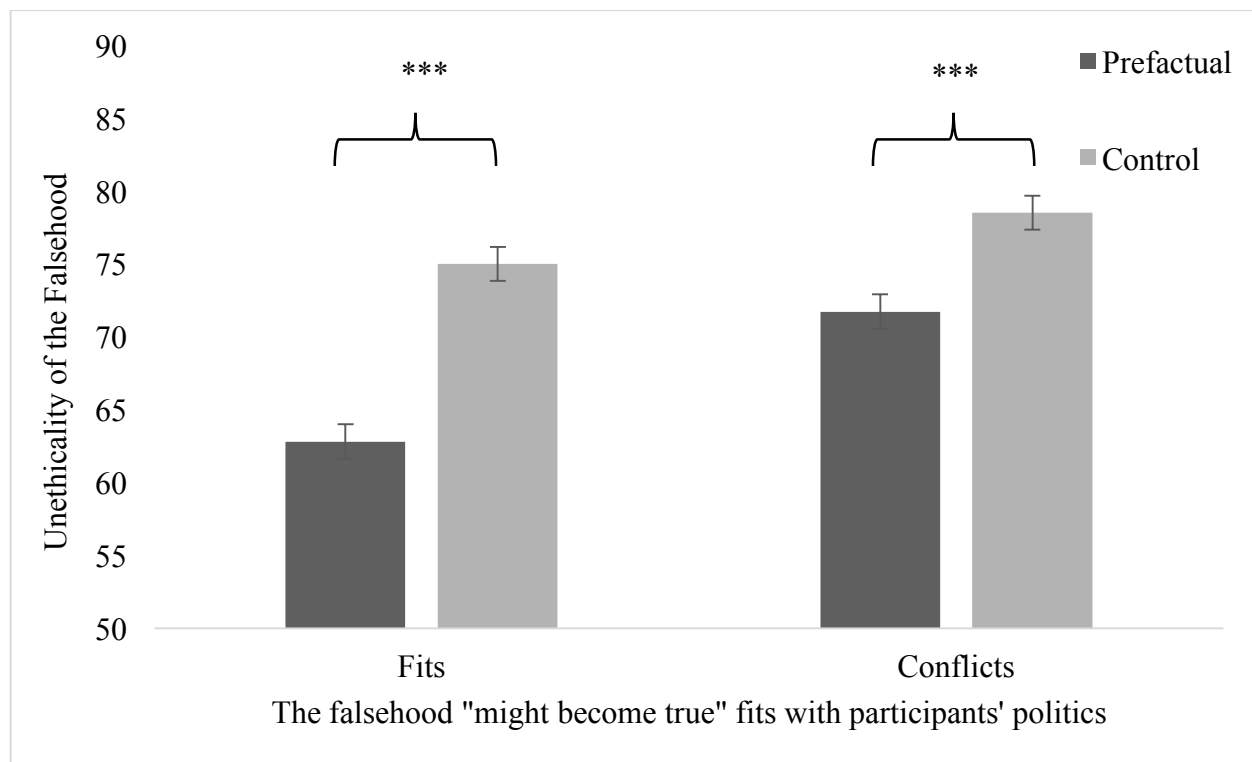
coded 1 = fits, 0 = conflicts, with a participant’s political beliefs. The mixed regression model also included participant random effects and item fixed effects. $N = 735$ in Step 1, and $N = 706$ in Step 2 because this analysis only included participants who considered themselves or leaned Republican or Democrat.

In addition, prefactual thinking reduced condemnation of falsehoods significantly more when the falsehood-prefactual pair fit, rather than conflicted, with participants' pre-existing political beliefs. Specifically, when we added a dummy code for political fit (1 = fits, 0 = conflicts) plus its interaction with condition to the regression model described above, the interaction term was significant and negative $b = -5.21$, $z = -4.32$, $p < .001$ (see Table 1, step 2; Figure 2). (This analysis included $N = 706$ participants; it omitted those who did not identify with or lean towards either political party because we could not code whether the falsehood-prefactual pairs fit with their politics). This finding supports H3 – that prefactuals reduce condemnation of falsehoods more effectively when people are already inclined to accept that the falsehood might become true in the future.

Decomposing this interaction using simple-slopes analyses, we found that the prefactual manipulation significantly reduced moral condemnation of falsehood when the falsehood-prefactual pair fit with participants' political views ($M_{\text{prefactual}} = 61.99$, $SD = 22.95$; $M_{\text{control}} = 73.16$, $SD = 22.18$), $d = -0.49$, $z = -7.05$, $p < .001$, as well as when it conflicted with their views ($M_{\text{prefactual}} = 72.09$, $SD = 21.96$; $M_{\text{control}} = 78.56$, $SD = 21.11$), $d = -0.30$, $z = -3.84$, $p < .001$. However, the size of the prefactual effect was about 60% larger when the falsehood-prefactual pair fit with their views.

Figure 2

Mean Unethicality Ratings by Condition and Political Fit in Study 3



Note. Full scale of unethicality ratings is 0-100. Plotted values are the estimated marginal means and their standard errors from the mixed regression model described in the main text.

*** $p < .001$.

Ease of imagining the prefactual: Prefactual potency. Finally, the results were consistent with our prediction that it is easier to imagine a falsehood might become true when that possibility fits with one's beliefs and motivations – and that the easier it is to imagine becoming true, the less unethical it seems to tell. First, participants thought prefactuals about falsehoods becoming true were more plausible (i.e., “potent”) when the falsehood-prefactual pair fit ($M = 62.36$, $SD = 32.57$) rather than conflicted ($M = 37.60$, $SD = 26.10$) with their politics, $d_z = 0.55$, $z = 16.25$, $p < .001$. Second, the more plausible (potent) that participants found the falsehood-prefactual pair, the less unethical they found the corresponding falsehood, $b = -0.29$, $z = -17.85$, $p < .001$. Finally, there was a significant indirect effect from fit with participants' politics, to the prefactuals' potency, to moral judgments, $b = -5.10$, 95% CI [-6.00, -4.20]. (Exploratory analyses suggested that this mediating effect of prefactual plausibility was driven by participants' judgements of the *then*-likelihood; see Online Supplement). These results are from mixed models with participants modelled as random effects and item as fixed effects; the indirect effect was computed by multiplying the a and b paths together in a multilevel mediation model computed with Stata's *gsem* command. The analyses are limited to participants in the prefactual condition because they were the only ones who saw and rated prefactuals about how the falsehood might become true.

Fact-check. We predicted that the prefactual manipulation would lead people to judge the falsehood as less unethical to tell despite knowing that it was false. The results were consistent with this prediction. Our fact-check measure showed that people correctly differentiated fact from falsehood 85% of the time. Some evidence suggests that the prefactual manipulation affected participants' beliefs about the falsehoods. In a mixed logistic regression analysis with random intercepts for participants, participants in the control condition correctly

differentiated fact from falsehood a greater proportion of the time (87.17%) than participants in the prefactual condition (81.98%), $b = -0.65$, $z = -3.59$, $p < .001$. The difference in response to the fact-check between prefactual and control condition did not depend on the political fit of the falsehood-prefactual pair, $b = -0.04$, $z = -0.19$, $p = .849$.

To examine whether these differences in acknowledging the falsehood was false could have accounted for the effect of prefactual thinking on participants' perceptions that the falsehood was unethical to tell, we repeated our prior analyses retaining only responses to the dependent measure that corresponded to correct fact-checks. Indicating our previous results were robust, the main effect of prefactual thinking, as well as its interaction with political fit, both remained significant and in the same direction, $b = -7.60$, $z = -5.13$, $p < .001$, and $b = -5.15$, $z = -4.05$, $p < .001$, respectively. Providing further evidence of robustness, these two effects also remained significant when we retained all data but statistically controlled for the fact-check measure, $b = -8.13$, $z = -5.76$, $p < .001$, and $b = -5.34$, $z = -4.42$, $p < .001$, respectively. Together, these results suggest that participants who reflected on how a falsehood might become true in the future thought the falsehood was less unethical to tell even when they knew the falsehoods were indeed false.

Discussion

In Study 3, participants who thought about how a political falsehood might become true in the future judged the falsehood as less unethical to tell – even though they knew it was false (H1). This finding replicated in a supplemental study, attesting to its robustness (see Online Supplement, Study S1). Moreover, this effect occurred even when the possibility that the falsehood might become true did *not* fit with participants' politics, but as predicted, was significantly larger when it did fit with their politics (H3). The reason, we have argued, is that it

is easier to imagine a falsehood becoming true if that possibility fits with one's pre-existing beliefs and motivations. Consistent with our theorizing, prefactuals about how a falsehood might become true seemed more plausible to participants when they fit with their politics – and the more participants thought it was plausible that the falsehood might become true, the less unethical they found the falsehood.

Study 4

The results of Studies 1–3 are consistent with our theorizing that prefactual thinking about how a falsehood might become true reduces how unethical that falsehood seems. However, the designs of Studies 1–3 cannot rule out the possibility that *any* act of prefactual thinking – regardless of whether it relates to the falsehood becoming true – is sufficient to produce this effect. To address this limitation, Study 4 used a more-stringent control condition. As in Study 3, participants rated the ethicality of telling various political falsehoods (e.g., “Millions of people voted illegally in the last presidential election”). Before doing so, all participants read if-then statements that invited prefactual thinking. The *if* part of the prefactual statement was identical in each of our two conditions (e.g., “If the United States does not tighten its border security...”). However, we manipulated the *then* part of the statement so that participants in the experimental group considered how the falsehood might become true in the future (*relevant-prefactual* condition; e.g., “... then millions of people will vote illegally in the upcoming presidential election”), whereas participants in the control group did not (*irrelevant-prefactual condition*; e.g., “... then millions of Americans will be out of a job before the next presidential election”). Our theorizing predicts that people will rate the falsehoods as less unethical to tell in the relevant-prefactual condition than in the irrelevant-prefactual condition (H1). That is, we predict the effects are specific to prefactuals relevant to how the falsehood might become true (see

Smallman, 2013). Moreover, we predict that this effect will be larger when people are inclined to accept that the falsehood might become true because this possibility fits, rather than conflicts, with their politics (H3).

Study 4's control condition also addresses an alternative explanation. It is possible that the *if* part of the prefactuals in Studies 1–3 highlight factual information about the present that influenced responses to the dependent measure. For example, the prefactual, “If the Trump administration keeps making pro-corporate decisions, then the average top CEO will soon make 500 times more than the average worker” highlights that the Trump administration was currently making pro-corporate decisions. This alternative explanation is not viable in Study 4 because participants in both conditions considered prefactuals that convey the same factual information because they have the same *if*-part.

Finally, Study 4 examined a potential downstream consequence of judging falsehoods as less unethical to tell: being more inclined to spread the falsehoods oneself. Past research suggests that imagination can—through influencing people's judgments—affect their intentions for future behavior (Anderson, 1983; Gregory et al., 1982; Sherman et al., 1981).

Method

Study 4 had a mixed 2 (condition: *relevant prefactual* vs. *irrelevant prefactual*; between participants) x 2 (partisan fit of “it might become true”: fits versus conflicts with participants' political beliefs; within participants) factorial design with 6 repeated measures. We pre-registered the study at <https://aspredicted.org/blind.php?x=6jx7rr>

Participants. We aimed to recruit 800 American MTurk participants, and 840 participants began the study. To promote data quality, we only allowed participants to begin the study if they passed a simple reading-comprehension question, and if they had not participated in

Study 3. After applying our pre-registered exclusion criteria (non-US or duplicate IP addresses; duplicate geocodes; people who failed to answer all measures for at least one falsehood), 803 people remained (479 women, 296 men, 28 did not report gender; $M_{\text{age}} = 37$, $SD_{\text{age}} = 12$). Of these participants, 723 reported that they considered themselves or leaned Democrat or Republican, so were included in our analyses of political fit: 493 identified as or leaned Democrat, 230 identified as or leaned Republican. Another 80 participants did not identify with or lean towards either party.

Statistical power. A sensitivity analysis with the PANGAEA app showed that, with 803 participants, six repeated measures, and two tailed $\alpha = .05$, Study 4 provided $> 99\%$ power to detect an effect size of $d = .2$ of the prefactual manipulation.

Materials. The stimuli were six political falsehoods, and for each falsehood, a fact that contradicted it, and a prefactual about how it might become true. Additionally, for each falsehood, we wrote a prefactual that did *not* involve imagining how the falsehood might become true. Four of the stimuli were the same as Study 3, and two were new (see Appendix B). Half of the falsehood-prefactual pairs fit with political stances associated with Republicans, such that Republicans would be more inclined to accept that the falsehood might become true, and half fit with political stances associated with Democrats, such that Democrats would be more inclined to accept that the falsehood might become true.

Procedure. Before beginning the study, participants responded to the same comprehension check question from Study 3. We prevented participants from beginning the study if they failed the comprehension check.

The procedure was identical to Study 3, except it used a different control condition. After reporting their political affiliation⁴, participants read a factual statement about a political issue (e.g., “It’s a proven fact that there have been just four documented cases of people voting illegally in the 2016 presidential election”). Then, participants randomly assigned to the *relevant-prefactual condition* read and rated the plausibility of an if-then statement about how a falsehood that contradicted the fact might become true (e.g., “If the United States does not tighten its border security, then millions of people will vote illegally in the upcoming presidential election”). Participants in the new *irrelevant-prefactual condition* read and rated the plausibility of an if-then statement that did not involve that falsehood becoming true (e.g., “If the United States does not tighten its border security, then millions of Americans will be out of a job before the next presidential election”). The *if* part of the prefactual statement was identical in both conditions. In both conditions, we measured how easy participants found it to imagine the prefactual. Finally, participants indicated their moral judgments of the falsehood and their intentions to promote it on social media, repeated this procedure for the remaining five falsehoods in randomized orders, responded to a fact-check measure, and provided demographics.

Measures.

Ease of imagining the prefactual: Prefactual potency. Using Study 3’s measure, participants separately indicated how likely they found the *if* component and *then* component of each prefactual they saw. As in Study 3, we multiplied judgments of the *if* and *then* likelihood together into a single measure of plausibility called prefactual potency (Petrocelli et al., 2011, 2012).

⁴ We also included a measure of participants’ recent voting behavior for exploratory purposes (see Online Supplement).

Dependent measure: Unethicality of telling the falsehood. Participants responded to the same three-item measure of the unethicality of telling the falsehood from Study 1 ($\alpha s > .76$ for each falsehood).

Intentions to promote the falsehood. We used four items to examine people's behavioral intentions to promote the falsehood (from Effron & Raj, 2020): "if one of your acquaintances shared an article with this headline on social media... how likely would you be to ..." (a) "like" it, (b) share it, (c) respond by posting a negative comment [reversed], or (d) block or unfollow this person [reversed]; 1 = *Not at all*, 7 = *To a great extent*). The items did not form a reliable composite ($\alpha s < .32$ across falsehoods), so we followed our pre-registered plan and analyzed each item separately.

Fact-check. Participants responded to the same true/false fact-check measure from Study 3 to test whether they remembered that the falsehoods were indeed false. We predicted that our manipulation would not affect the fact-check measure and that our results would remain reliable when excluding participants who failed the fact-check.⁵

Results

We had strong directional predictions so we pre-registered one-tailed tests for all confirmatory analyses (Cho & Abe, 2013).

Unethicality of telling the falsehood. Study 4 replicated Study 3's results with a more stringent control condition. As predicted (H1), participants judged falsehoods as significantly less unethical when they reflected on how the falsehoods might become true in the future ($M = 65.57$, $SD = 21.23$) compared to when they reflected on how a similar statement might become

⁵ At the end of the study, we included additional survey questions on participants' political knowledge for exploratory purposes (see Online Supplement).

true in the future ($M = 70.22$, $SD = 20.26$), $d = -0.22$, $z = -3.20$, $p = .001$ (see Table 2; step 1).

This difference between conditions was significant in a mixed-effect regression model, with condition as a fixed effect (1 = relevant-prefactual, 0 = irrelevant-prefactual), participant random effects, and fixed effects for the six items.

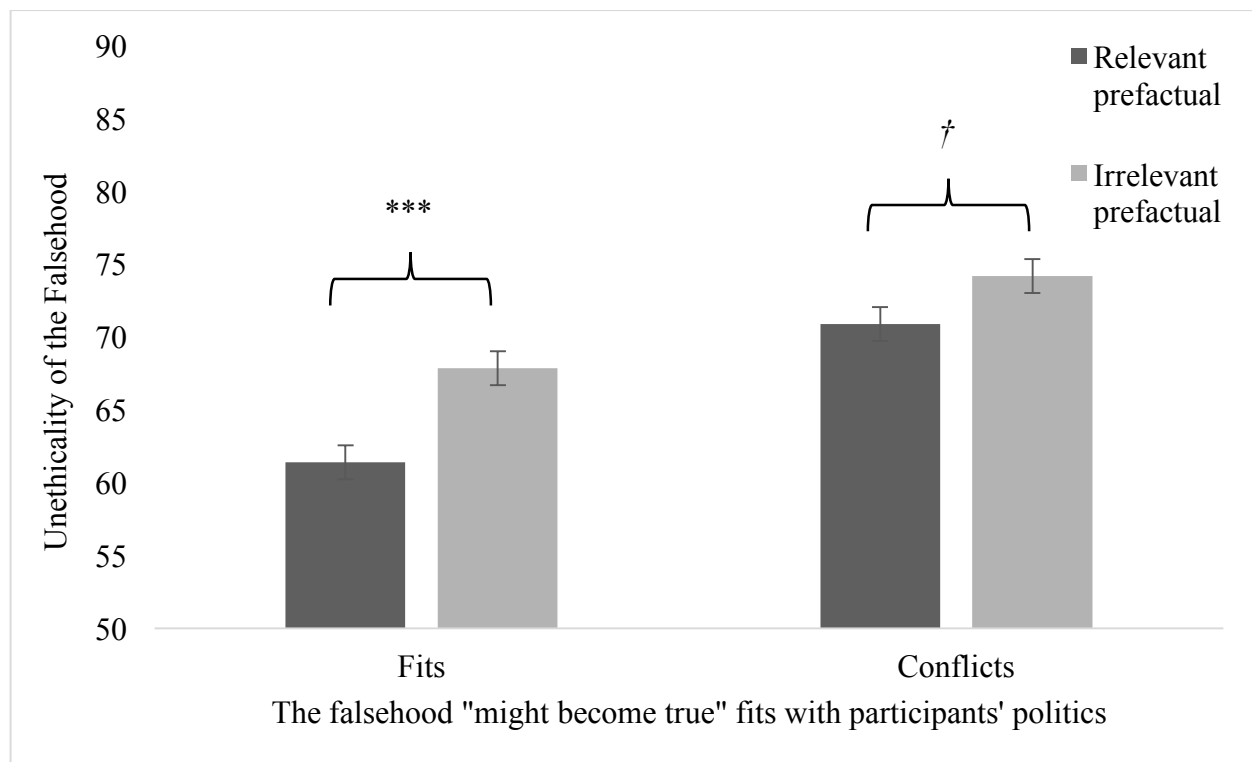
Table 2*Regression Analyses Predicting Unethicality Ratings in Study 4*

Predictor	<i>b</i>	<i>SE(b)</i>	<i>z</i>	<i>p</i>	95% CI
Step 1					
Condition	-4.67	1.46	-3.20	.001	$[-\infty, -2.27]$
Constant	63.11	1.21	52.13	< .001	[60.74, 65.48]
Step 2					
Condition	-3.13	1.66	-1.89	.059	[-6.37, 0.12]
Political fit	-4.90	0.87	-5.64	< .001	[-6.60, -3.20]
Condition x political fit	-3.47	1.18	-2.93	.002	$[-\infty, -1.52]$
(Constant)	67.87	1.38	49.18	< .001	[65.17, 70.58]

Note. Condition is coded 1 = relevant-prefactual, 0 = irrelevant-prefactual. Political fit with “it might become true” coded 1 = fits, 0 = conflicts, with a participant’s political beliefs. The mixed regression model also included participant random effects and item fixed effects. $N = 803$ in Step 1. $N = 723$ for Step 2 because this analysis only included participants who considered themselves or leaned Republican or Democrat. We report one-tailed 95% confidence intervals for pre-registered one-tailed tests. Values of ∞ reflect that one-tailed tests, by definition, cannot detect an effect in the opposite direction as predicted.

Supporting H3, the effect of the prefactual manipulation was significantly larger when the falsehood-prefactual pair fit, rather than conflicted, with participants' politics, $b = -3.47$, $z = -2.93$, $p = .002$ (see Table 2, step 2; Figure 3). For this analysis, we added a dummy code for political fit and its interaction with condition to the mixed model. (This analysis included $N = 723$ participants; it omitted those who did not identify with or lean towards either political party because we could not code whether the prefactual fit with their politics). This finding is consistent with our hypothesis that prefactual thinking would reduce participants' condemnation of falsehoods more when they were inclined to accept that the falsehood might become true because it fit with their politics.

Simple slopes analysis showed that when the falsehood-prefactual pair fit with participants' political views, considering how the falsehood might become true led participants to rate the falsehood as less unethical to tell ($M = 61.04$, $SD = 24.69$) than considering an irrelevant prefactual ($M = 67.61$, $SD = 22.92$), $d = -0.28$, $z = -3.99$, $p < .001$. When the falsehood-prefactual pair conflicted with participants' political views, considering a relevant prefactual also led participants to rate the falsehood as less unethical to tell ($M = 70.64$, $SD = 22.50$) than considering an irrelevant prefactual ($M = 73.91$, $SD = 21.15$) – but this effect was only marginally significant, $z = -1.89$, $p = .059$, and about half as large, $d = -0.15$. We report two-tailed significance tests of the simple slopes because we did not pre-register these analyses.

Figure 3*Mean Unethicality Ratings by Condition and Political Fit in Study 4*

Note. Full scale of unethicality ratings is 0-100. Plotted values are the estimated marginal means and their standard errors from the mixed regression model described in the main text.

† $p < .10$; *** $p < .001$.

Ease of imagining the prefactual: Prefactual potency. The results were also consistent with our theorizing that it is easier to imagine a falsehood might become true when that possibility fits with one's motivations and beliefs, and that the easier it is to imagine become true, the less unethical the falsehood will seem. First, as in Study 3, scores on the prefactual potency measure showed that participants rated the relevant prefactuals as more plausible when the falsehood-prefactual pair fit ($M = 64.80$, $SD = 28.24$) rather than conflicted ($M = 35.15$, $SD = 20.52$) with their political beliefs, $d_z = 0.90$, $z = 19.25$, $p < .001$. Second, the more plausible participants rated the relevant prefactual on the potency measure, the less unethical they found the falsehood, $b = -0.26$, $z = -17.26$, $p < .001$. Finally, we observed a significant indirect effect from fit with participants' politics, to prefactual potency, to moral judgments in the relevant-prefactual condition, $b = -5.52$, 95% CI $[-\infty, -4.73]$ ⁶, computed using the *gsem* and *nlcom* commands in Stata. (Exploratory analyses showed that – as in Study 3 – the mediating effect of prefactual potency was driven by judgments of the *then*-likelihood; see Online Supplement). These results are from mixed models with participants modelled as random effects and item as a fixed effect; the indirect effect was computed by multiplying the *a* and *b* paths together in a multilevel mediation model. The analyses are limited to participants in the relevant prefactual condition because they were the only ones who saw and rated prefactuals relevant to the falsehood.

Intentions to promote the falsehood. We next examined a potential downstream consequence of more-lenient moral judgments: stronger intentions to promote the falsehoods on social media. As predicted, we observed significant indirect effects of the prefactual manipulation on intentions to promote the falsehood. Specifically, imagining how a falsehood

⁶ We report one-tailed 95% confidence intervals for pre-registered one-tailed tests. Values of ∞ reflect that one-tailed tests, by definition, cannot detect an effect in the opposite direction as predicted.

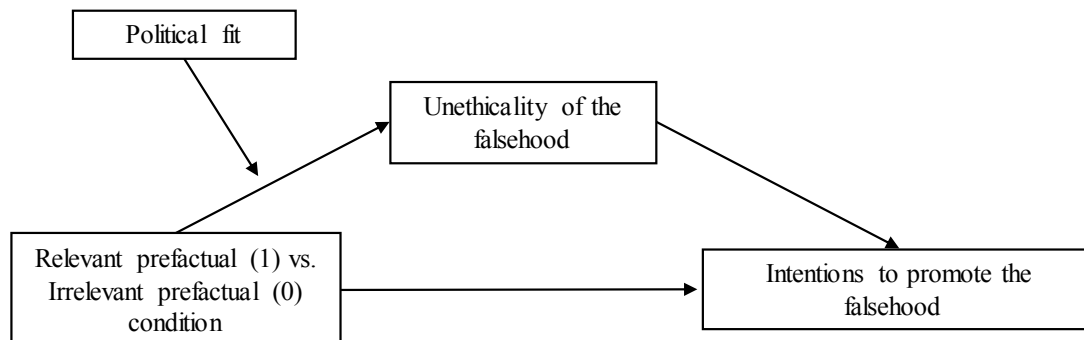
might become true made it seem less unethical to tell, which in turn predicted stronger intentions to “like” and share the falsehood, as well as weaker intentions to respond to someone who posted the falsehoods with a negative comment or by blocking/unfollowing this person; respectively, $b = 0.10 [0.05, \infty]$, $b = 0.09 [0.04, \infty]$, $b = -0.04 [-\infty, -0.02]$, and $b = -0.06 [-\infty, -0.02]$ for the indirect effects and their 95% CIs. We conducted these analyses as generalized structural equation models with prefactual condition as the independent variable (1 = relevant-prefactual, 0 = irrelevant-prefactual), perceptions of unethicality as the mediator, intentions to promote the falsehood as the dependent variable, and random effects for participants and fixed effects for item to account for the data’s multilevel structure. We computed the indirect effect by multiplying the a - and b -paths together using the *gsem* and *nlcom* functions in Stata.

Moreover, as predicted, these indirect effects were significantly greater when the falsehood-prefactual pair fit (vs. conflicted) with participants’ politics (Figure 4, Table 3). We added moderation by political fit on the a -path of the mediation models with prefactual condition as the independent variable, perceptions of unethicality as the mediator, and intentions to promote the falsehood as the dependent variable.

The total effect of the prefactual condition was significant for reducing participants’ intentions to block/unfollow the person who posted the falsehood, $b = -0.17$, $z = -1.80$, $p = .036$, but was not statistically significant for any of the other promoting intentions (see Online Supplement).

Figure 4

Conditional Indirect Effect Model of Prefactual Manipulation on Intentions to Promote the Falsehood on Social Media in Study 4



Note. We computed separate conditional indirect effects models for the four promoting behaviors: “like” the falsehood, share the falsehood, respond by posting a negative comment [reversed], or block or unfollow this person. We report the conditional indirect effects and indices of moderated mediation for each promoting behavior in Table 3.

Table 3*Indirect Effects of Prefactual Manipulation on Promoting the Falsehood through Unethicality in Study 4*

Effects	“Like”		Share		Negative comment		Block/unfollow	
	<i>b</i>	95% CI	<i>b</i>	95% CI	<i>b</i>	95% CI	<i>b</i>	95% CI
Conditional indirect effects								
“It might become true” fits with politics	0.15	[0.08, 0.23]	0.13	[0.06, 0.19]	-0.05	[-0.08, -0.02]	-0.06	[-0.09, -0.03]
“It might become true” conflicts with politics	0.07	[-0.003, 0.15]	0.06	[-0.002, 0.12]	-0.02	[-0.05, 0.002]	-0.03	[-0.06, 0.002]
Index of moderated mediation	0.08	[0.04, ∞]	0.07	[0.03, ∞]	-0.03	[-∞, -0.01]	-0.03	[-∞, -0.01]

Note. Indirect effects and index of moderated mediation specifying Prefactual condition as the independent variable (1 = relevant-prefactual, 0 = irrelevant-prefactual), perceived judgments of unethicality as the mediator, moderation by political fit on the a-path, and intentions to promote the falsehood (“like”, share, negative comment, block/unfollow) as the dependent variables. The model includes item fixed effects and participant random effects to account for the multi-level nature of the data. Coefficients that are statistically significant at one-tailed $p < .05$ are in bold. We report one-tailed 95% confidence intervals for pre-registered one-tailed tests. Values of ∞ reflect that one-tailed tests, by definition, cannot detect an effect in the opposite direction as predicted.

Fact-check. We hypothesized that relevant prefactuals would lead people to judge the falsehoods as less unethical to tell despite knowing that they were false. The fact-check measure showed that people correctly identified fact from falsehood 86% of the time – a proportion that did not differ significantly between the relevant-prefactual (84.83%) and irrelevant-prefactual (86.77%) conditions, $b = -0.21$, $z = -1.50$, $p = 0.133$ (two-tailed test). This analysis was conducted in a mixed logistic regression analysis with random intercepts for participants and fixed effects for falsehood. Furthermore, there was not a significant interaction between prefactual condition and political fit on recognition of the falsehood, $b = -0.07$, $z = -0.33$, $p = .743$.

As a robustness check, we repeated our prior analyses with only responses to the dependent measures that corresponded to correct fact-checks. Indicating our previous results were robust, the main effect of the prefactual manipulation, as well as its interaction with political fit, both remained significant and in the same direction, $b = -4.95$, $z = -3.28$, $p = .001$, and $b = -4.35$, $z = -3.06$, $p = .002$, respectively. As a second robustness check, we repeated our prior analyses controlling for the fact-check measure. The main effect of prefactuals was significant when controlling for the fact-check, $b = -4.39$, $z = -2.98$, $p = .003$, as was the interaction between the prefactual manipulation and political fit, $b = -4.56$, $z = -3.43$, $p = .001$, providing further evidence of robustness. Thus, as in Study 3, we found no evidence that prefactual thinking reduced the moral condemnation of falsehoods simply by convincing people the falsehoods were true.

Discussion

Study 4 replicated Study 3's results with a more stringent control condition. In so doing, it ruled out two alternative explanations for our prior results. First, the possibility that *any* act of

prefactual thinking reduces moral condemnation cannot explain Study 4's results, because both conditions prompted participants to think prefactually. Instead, the results were consistent with our claim that thinking prefactually *about how a falsehood might become true* is what made the falsehoods seem less unethical (H1). Second, the results cannot be explained by the possibility that the specific prefactuals in the study highlighted factual information about the present, because participants in both conditions were shown prefactuals that conveyed the same factual information. Study 4 also found evidence for potential downstream consequences of moral judgments. Imagining how a falsehood could become true made it seem less unethical to tell, which in turn predicted a greater inclination to promote the falsehood on social media.

Study 5

Study 5 tested whether our previous results would replicate if, instead of instructing participants to consider prefactuals that we provided, we prompted participants to generate their own prefactuals. A successful replication would advance our research in two ways. First, it would suggest that our previous results generalize beyond the specific prefactuals we used in Studies 1-4. Second, along with Study 4's results, it would cast further doubt on the possibility that these specific prefactuals highlighted factual information that influenced participants' judgments. By manipulating whether participants themselves generate ways in which the falsehoods could become true, we could ensure that all participants were exposed to the same factual information.

In addition, Study 5 sought additional support for our proposed mechanism (H2). As in Study 2, participants reported their perceptions that the gist of the falsehood was true. We expected that participants who imagined prefactuals about how the falsehoods might become true

would perceive the gist of the falsehood as truer, which in turn would predict them judging the falsehood as less unethical to tell.

Participants in Studies 3 and 4 expressed less condemnation of a falsehood when they imagined how it might become true – especially when that prefactual fit with their political views. Consistent with our explanation for these results, Studies 3 and 4 suggest that when people are inclined to accept that a falsehood might become true they find falsehoods about how the falsehood might become true easier to imagine – and that falsehoods that were easier to imagine seemed less unethical. To operationalize the ease of imagination, those studies asked participants to rate the likelihood of the relevant prefactuals occurring (see Effron, 2018; Petrocelli et al., 2011, 2012). Study 5 sought convergent evidence with a different operationalization of ease of imagination: Participants rated how vividly they had imagined the prefactuals (see Gaesser et al., 2018, 2019). The easier it is to imagine a prefactual, the more vividly people should say they imagined it.

Finally, in Study 5 we again measured participants' intentions to promote the falsehood online. We predicted that when imagining how a falsehood might become true led participants to judge the falsehood as less unethical, participants would report greater intentions to share the falsehood on social media.

Method

Study 5 had a 2 (condition: prefactual vs irrelevant prediction; between–subjects) x 2 (partisan fit of “it might become true”: fits versus conflicts with participants' political beliefs; within-subjects) factorial design with six repeated measures. We pre-registered the study at <https://aspredicted.org/blind.php?x=sf98wy>.

Participants. We aimed to recruit 600 American political partisans on Prolific Academic (Peer et al., 2017), and 646 participants began the study. To be eligible to complete the study, participants had to be listed as either a Democrat or a Republican in Prolific Academic's database, be located within the US, correctly answer a reading comprehension question, pass a CAPTCHA check to verify they are human, and complete the study on a computer rather than a mobile device. After applying our pre-registered exclusion criteria (responses from duplicate Prolific participant IDs and people who failed to complete all the dependent measures for at least one falsehood),⁷ our final sample size was 599 participants (271 men, 298 women; $M_{age} = 33$, $SD = 12$; 441 identified as or leaned Democrat and 158 identified as or leaned Republican).

Statistical power. A sensitivity analysis with the PANGEA app showed that, with 599 participants, six repeated measures, and two tailed $\alpha = .05$, Study 5 provided $> 99\%$ power to detect an effect size of $d = .2$ of the prefactual manipulation.

Procedure. The stimuli were the six political falsehoods and associated facts from Study 4. The procedure was identical to Study 4. Participants first answered a comprehension check question, reported their political affiliation, and read that they would be considering “a number of facts that have been verified by reputable, non-partisan fact-checking websites”. They then read a “proven fact,” completed the prefactual manipulation, responded to the dependent measures about a falsehood that contradicted the fact, and repeated this procedure for the remaining five stimuli. Half of the falsehood-prefactual pairs fit with political stances associated with Republicans and half fit with political stances associated with Democrats. Finally, participants answered fact-check questions.

⁷ We also pre-registered plans to exclude participants with duplicate or non-US IP addresses or duplicate geolocations. However, we inadvertently neglected to collect IP addresses and geolocations due to an error in Qualtrics, so we could not execute our pre-registered intention to exclude this data.

This time, however, the manipulation did not provide participants with specific prefactuals. Instead, participants generated their own predictions either about how the falsehoods might become true (*prefactual condition*) or about how unrelated events might unfold in the future (*irrelevant-prediction condition*). Appendix B shows the prefactuals and irrelevant predictions.

For example, for one of the six stimuli, participants in both conditions read, “It’s a proven fact that the average top CEO currently makes 265 times more money than the average American worker.” In the *prefactual condition*, participants then typed a response to the following open-ended prompt:

However, it is possible to imagine that the average top CEO will soon make 500 times more money than the average American worker. Please complete the prediction below:

The average top CEO will soon make 500 times more money than the average American worker if... _____.

Participants in the *irrelevant-prediction* condition instead responded to the following prompt:

Now consider current U.S. economic sanctions on North Korea. It is possible to imagine different consequences of this. Please complete the prediction below:

Current U.S. economic sanctions on North Korea will lead to... _____.

Participants then responded to measures of prefactual vividness, unethicity of the falsehood, gist truthfulness, and intentions to promote the falsehood. Unlike Study 2, Study 5 administered the unethicity dependent measure prior to ratings of the gist truthfulness to test robustness of this mediation pathway.

Measures.

Dependent measure: Unethicity of telling the falsehood. We used the same three-item measure of the unethicity of telling the falsehood from Study 1 ($\alpha > .77$).

Gist truthfulness. Participants reported their perceptions that the gist of the falsehood was true with the same four-item measure from Study 2 ($\alpha > .96$).

Ease of imagining the prefactual: Prefactual vividness. After participants wrote their prefactual or irrelevant predictions, they responded to three items assessing how vivid, intense, and detailed their prediction felt from 1 = *Not at all* to 5 = *Extremely* ($\alpha > .91$), adapted Johnson, Foley, Suengas, and Raye (1988).

Intentions to promote the falsehood. To reduce study length, we administered just one of Study 4's items assessing intentions to promote the falsehood: "If one of your acquaintances posted an article with this headline on social media, how likely would you be to share it?" (1 = *Not at all likely* to 7 = *Extremely likely*).

Fact-check. As in Study 3 and 4, participants responded to six true/false questions assessing their understanding that the political falsehoods were false.

Results

We had strong directional predictions so we pre-registered and report one-tailed significant tests (Cho & Abe, 2013). We used two-tailed tests for any exploratory or non-pre-registered analyses, as described below.

Main effect of prefactual thinking. Conceptually replicating Studies 1-4 and supporting H1, participants rated falsehoods as less unethical to tell when they wrote about how they might become true (*prefactual condition*: $M = 66.63$, $SD = 22.26$) versus how irrelevant claims might become true (*irrelevant prediction condition*: $M = 73.84$, $SD = 19.28$). This difference was significant in a mixed regression model with condition (1 = prefactual, 0 = irrelevant prediction) as a fixed effect, item as a fixed effect, and participant as a random effect, $d = -0.35$, $z = -4.23$, $p < .001$ (see Table 4, step 1).

Table 4*Regression Analyses Predicting Unethicality Ratings in Study 5*

Predictor	<i>b</i>	<i>SE(b)</i>	<i>z</i>	<i>p</i>	95% CI of <i>b</i>
Step 1					
Condition	-7.14	1.69	-4.23	<.001	$[-\infty, -4.36]$
Constant	84.08	1.38	61.04	<.001	[81.38, 86.78]
Step 2					
Condition	-6.20	1.80	-3.44	.001	[-9.74, -2.67]
Political fit	-3.65	0.93	-3.91	<.001	[-5.47, -1.82]
Condition x political fit	-1.84	1.25	-1.47	.071	$[-\infty, 0.21]$
(Constant)	84.81	1.42	59.69	<.001	[82.02, 87.59]

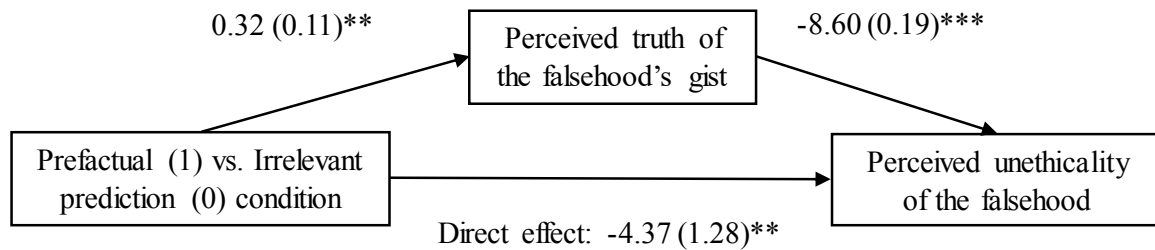
Note. Condition is coded 1 = prefactual, 0 = irrelevant prediction. Political fit with “it might become true” coded 1 = fits, 0 = conflicts, with a participant’s political beliefs. The mixed regression model also included participant random effects and item fixed effects. We report one-tailed 95% confidence intervals for pre-registered one-tailed tests. Values of ∞ reflect that one-tailed tests, by definition, cannot detect an effect in the opposite direction as predicted.

Mediation by gist truthfulness. We predicted that considering a prefactual statement would reduce participants' condemnation of falsehoods by increasing their perceptions that the gist of the falsehood is true. The results were in line with this prediction (H2), conceptually replicating Study 2's results. Specifically, when participants imagined how the falsehood might become true, they thought the falsehood's gist was truer, and the truer they found the gist, the less unethical they thought it was to tell the falsehood (see Figure 5) – a significant indirect effect, $b = -2.77$, 95% CI $[-4.70, -0.83]$. When we included gist truthfulness in the model predicting unethicality judgments, the direct effect of the prefactual condition was significant, $b = -4.37$, 95% CI $[-6.88, -1.87]$. We conducted this analysis as a generalized structural equation model with prefactual condition as the independent variable (1 = prefactual, 0 = irrelevant prediction), perceived truth of the falsehood's gist as the mediator, and perceptions of unethicality as the dependent variable, with fixed effects for item and random intercepts for participants to account for the multilevel structure of the data. We computed the indirect effect by multiplying the a - and b -paths together using the *gsem* and *nlcom* functions in Stata⁸.

⁸ We report two-tailed tests for this analysis because our pre-registration document included an error in explaining how we would analyze mediation by gist. We mistakenly pre-registered testing whether gist mediated the effect of political fit rather than the effect of prefactual thinking.

Figure 5

Indirect Effect of Prefactual Manipulation on Unethicality Judgments through Gist in Study 5



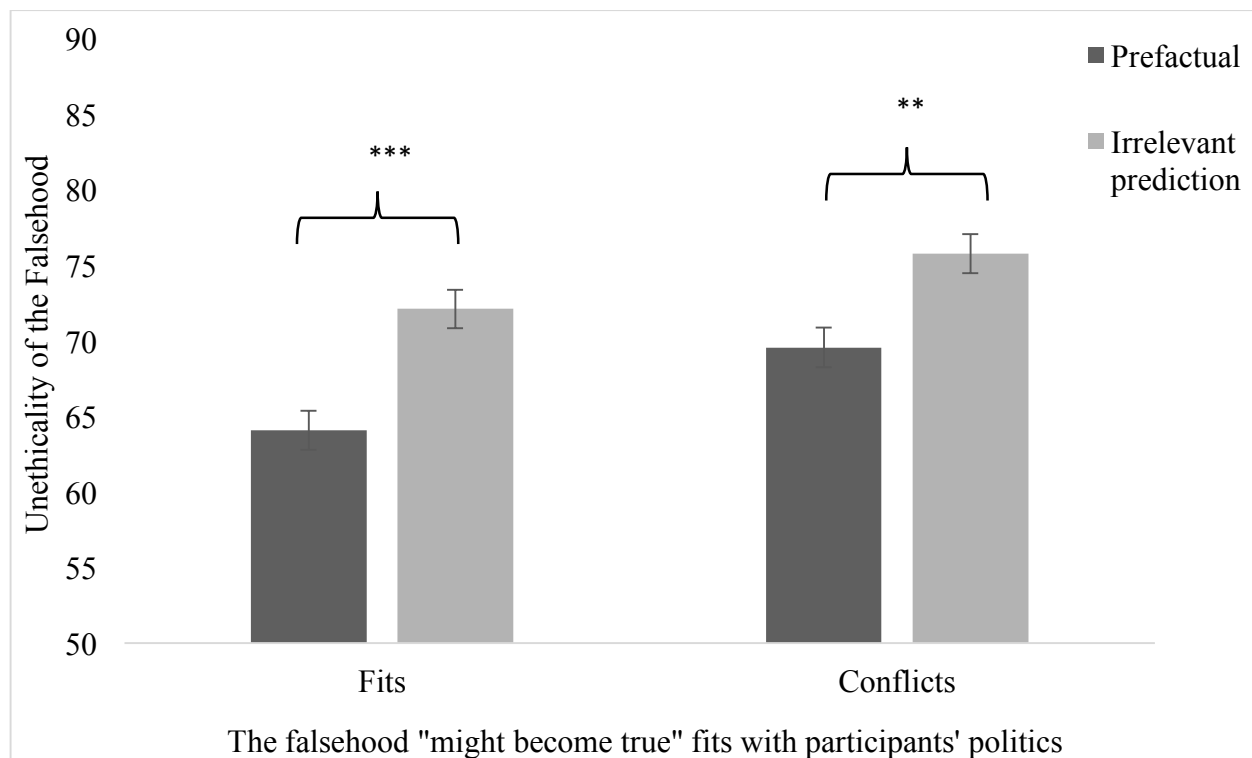
Note. Indirect effect: $b = -2.77$, 95% CI $[-4.70, -0.83]$ **. Unstandardized coefficients shown.

Model includes item fixed effects and participant random effects.

* $p < .05$; ** $p < .01$, *** $p < .001$

Moderation by political fit. To test our prediction that prefactual thinking would have a greater effect on reducing condemnation of falsehoods when participants were inclined to accept that the falsehood might become true based on their political views, we next added a dummy code for political fit (1 = fit, 0 = conflicts, with participants' political beliefs), and its interaction with the prefactual manipulation, to the regression model (Table 4, step 2). Consistent with Studies 3 and 4 and our directional prediction (H3), the coefficient on the interaction term was negative; however, it was not statistically significant, $b = -1.84$, $z = -1.47$, $p = .071$ (see Figure 6).

Nonetheless, because it was marginally significant, we decomposed the interaction with simple slopes to examine the overall pattern of results. When a falsehood-prefactual pair fit with participants' political views, they rated it as less unethical to tell if they wrote about how it might become true in the future ($M = 62.73$, $SD = 24.95$) than if they wrote about an irrelevant prediction ($M = 71.19$, $SD = 21.09$), $d = -0.37$, $z = -4.47$, $p < .001$. The same effect was also significant but directionally smaller when the falsehood-prefactual pair conflicted with participants' political views (*prefactual condition*: $M = 70.94$, $SD = 22.80$; *irrelevant prediction condition*: $M = 76.92$, $SD = 19.78$), $d = -0.28$, $z = -3.44$, $p = .001$. These simple slope analyses are two-tailed, because we did not pre-register these analyses.

Figure 6*Mean Unethicality Ratings by Condition and Political Fit in Study 5*

Note. Full scale of unethicality ratings is 0-100. Plotted values are the estimated marginal means and their standard errors from the mixed regression model described in the main text.

** $p < .01$; *** $p < .001$.

Ease of imagining the prefactual: Prefactual vividness. The results were also consistent with our theorizing that it would be easier to imagine a falsehood becoming true if that possibility fits with one's pre-existing beliefs and motivations. Participants in the prefactual condition reported that they had more vividly imagined the falsehood becoming true when the falsehood-prefactual pair fit ($M = 2.90$, $SD = 0.99$) rather than conflicted ($M = 2.62$, $SD = 0.99$) with their politics, $d_z = 0.39$, $z = 4.54$, $p < .001$. We examined this difference in a mixed-effect regression model, with political fit as a fixed effect (1 = fits, 0 = conflicts), participant random effects, and item fixed effects. This analysis was limited to participants in the prefactual condition because they were the only ones who saw and rated prefactuals about how the falsehood might become true.

As an exploratory analysis, we examined whether prefactual vividness and gist mediated the effect of political fit in the prefactual condition, such that participants found it easier to imagine that a falsehood might become true when that possibility fit with their politics, and the easier the prefactual was to imagine, the more the gist of the falsehood seemed true, and the less unethical they judged the falsehood. There was a significant indirect effect from political fit to vividness to gist to unethicity in a serial mediation model, $b = -0.54$, 95% CI $[-0.83, -0.26]$, computed with item fixed effects and participant random effects, and an independent covariance structure (Stata's default). We did not pre-register this analysis so we consider these results preliminary. This analysis was again limited to participants in the relevant-prefactual condition because they were the only ones who saw and rated prefactuals about how the falsehood might become true.

Intentions to promote the falsehood. As in Study 4, imagining how a falsehood might become true made it seem less unethical to tell, which in turn predicted stronger intentions to

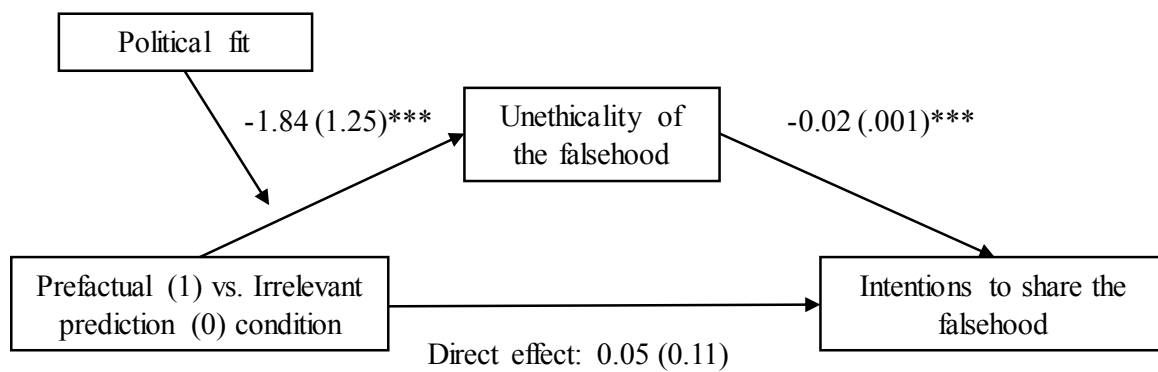
share the falsehood online – a significant indirect effect, $b = 0.16$, 95% CI $[0.09, \infty]$. We conducted this analysis as a generalized structural equation model with prefactual manipulation as the independent variable (1 = prefactual, 0 = irrelevant prediction), perceptions of unethicity as the mediator, intentions to promote the falsehood as the dependent variable, plus random effects for participants and fixed effects for item to account for the data's multilevel structure. We computed the indirect effect by multiplying the a - and b -paths together.

This indirect effect was significant both when the falsehood-prefactual pair *fit* with participants' politics, $b = 0.17$, 95% CI $[0.09, 0.25]$, and when it *conflicted* with participants' politics, $b = 0.13$ 95% CI $[0.06, 0.21]$. But, as predicted, it was larger when the falsehood-prefactual pair fit with participants' politics, $b = 0.04$, 95% CI $[-0.005, \infty]$ for the index of moderated mediation (Figure 7). For this analysis, we added political fit and its product with the prefactual manipulation on the a -path of the mediation model described above.

Analyses of the manipulation's total effect on sharing intentions revealed that imagining how the falsehood might become true increased participants' intentions to share the falsehood on social media, but only if the possibility that the falsehood might become true fit with their politics (see Online Supplement).

Figure 7

Conditional Indirect Effect of Prefactual Manipulation on Intentions to Share the Falsehood on Social Media in Study 5



Note. Unstandardized coefficients shown. Model includes item fixed effects and participant random effects.

*** $p < .001$.

Fact-check. We predicted that the prefactual manipulation would lead people to judge the falsehood as less unethical to tell despite them knowing that the statement was false. The fact-check measure showed that people correctly differentiated fact from falsehood 74% of the time – a proportion that did not differ between prefactual (74.26%), and irrelevant prediction (74.19%) conditions, $b = 0.00$, $z = 0.00$, $p > .99$, nor did it depend on the interaction between prefactual condition and political fit, $b = -0.30$, $z = -1.21$, $p = .227$ (these tests were not pre-registered, so the p -values are for two-tailed tests). We ran these analyses as a mixed logistic regression with random intercepts for participants and fixed effects for items. Thus, as expected, people recognized the falsehoods as factually incorrect even though imagining how they might become true in the future made the falsehoods' gist seem truer.

As a robustness check, we also repeated our main analyses excluding responses to the dependent measures that corresponded to incorrect fact-checks. After these exclusions, the main effect of the prefactual manipulation on judgments of how unethical the falsehood is to tell remained significant in the predicted direction, $b = -8.42$, $z = -4.95$, $p < .001$. Further, the main effect of the prefactual manipulation on judgments of how unethical the falsehood is to tell remained significant in the predicted direction when we retained all data but statistically controlled for the fact-check measure, $b = -7.37$, $z = -4.25$, $p < .001$. However, the interaction between prefactual condition and political fit was not significant when excluding responses that failed the fact-check, $b = 1.44$, $z = 0.85$, $p = .198$, or when controlling for the fact-check, $b = -0.98$, $z = -0.73$, $p = .233$.

Together, results from the fact-checks support the hypothesis that prefactual thinking reduces the moral condemnation of falsehoods despite knowing their falsity. However, these

results provide equivocal support for the prediction that the prefactual effect would depend on whether participants were inclined to accept that the falsehood might become true.

Discussion

Study 5 – in which participants generated their own prefactuals – replicated several key results. Considering how a falsehood might become true in the future decreased participants' condemnation of the falsehood (H1), and this effect was significantly mediated by beliefs about the falsehoods' gist (H2). Additionally, as in Study 4, the more that the manipulation reduced moral condemnation, the more inclined participants were to share the falsehoods on social media. These results suggest that our previous results generalize beyond the specific prefactuals we used in Studies 1-4, and cast further doubt on the possibility that those specific prefactuals influenced judgments by highlighting factual information. Moreover, the irrelevant-prediction control condition ruled out the possibility that our effects are the result of imagining the future in general, rather than specifically imagining how the falsehood might become true.

One limitation of this study is that the prefactual and control condition differed in their form as well as their content. Specifically, participants in the prefactual condition generated a prefactual ("*If X, then Y*"), whereas participants in the control condition generated a prediction ("*X will lead to Y*"). Therefore, the differences we observed in moral condemnation between conditions could in part be the result of differences in the form of statements each group generated. We address this limitation in Studies 4 and 6 by including a control condition in which participants consider prefactuals with identical form, but that do not involve imagining how the falsehood might become true.

The results were directionally consistent with our hypothesis that imagining how the falsehood might become true reduces condemnation of falsehoods more strongly when the

prefactual fits with people's pre-existing beliefs (H3). Consistent with our argument that ease of imagination explains this pattern, participants said they imagined prefactuals more vividly when the falsehoods fit with their politics, and the more vividly they imagined the prefactuals, the more they thought the gist of the falsehoods were true, and the less unethical they found the falsehoods. Moreover, Study 5 established convergent validity for the role of ease of imagination in the moderating effect of political fit with a different operationalization—participants' ratings of the vividness of their imagination. An important caveat, however, is that the key statistical test of H3 was not significant in this study ($p = .071$ for the main analysis). It is possible that random error variance explains why the evidence for H3 was stronger in Studies 3 and 4. Another possibility is that when asked to generate prefactuals themselves, participants imagined ways the falsehood might become true that most easily came to mind for them. As a result, even though Study 5's participants may have found prefactuals easiest to imagine when they were inclined to accept that the falsehood might become true, they may have still found their prefactuals at least moderately easy to imagine when they were not inclined to accept this possibility.

Study 6

Study 6 improved our paradigm in two ways to provide a more-conservative test of our theorized mechanism – that imagining how a falsehood might become true makes the falsehood seem less unethical by making its gist seem truer (H2).

First, the gist truthfulness measure we used in Studies 2 and 5 did not specify what the gist communicated by the falsehood was, leaving some ambiguity about what participants meant when they endorsed these items. To address this limitation, participants in Study 6 rated the truthfulness of specific statements that captured the gist of each falsehood. Second, Studies 2 and 5 did not give participants the opportunity to rate on a continuous scale how truthful they

perceived the verbatim details of the falsehood, which prevented our analyses from clearly distinguishing between perceptions of gist and verbatim truthfulness. Study 6 addresses both of these limitations with new measures. Our theorizing predicts that the prefactual manipulation will have a significant indirect effect on moral judgments through perceptions of *gist truthfulness* but not through perceptions of *verbatim truthfulness*.

In addition, in Study 6 we test the moderating role of participants' pre-existing belief with a new set of partisan falsehoods.

Method

Study 6 had a mixed 2 (condition: *relevant prefactual* vs. *irrelevant prefactual*; between participants) x 2 (partisan fit of "it might become true": fits versus conflicts with participants' political beliefs; within participants) factorial design with 8 repeated measures. We pre-registered the study at https://aspredicted.org/FQX_HTX

Participants. Using Prolific Academic's pre-screen data, we aimed to recruit 800 regular users of Twitter, Facebook, Instagram, or Reddit (defined as those who use the sites at least once a month). We targeted these individuals because they are particularly likely to encounter political falsehoods through their use of these platforms. We separately recruited 400 Democrats and 400 Republicans to obtain an equal number of participants who supported each political party, and 800 participants began the study. To promote data quality, we only allowed participants to begin the study if they were located within the US, correctly answered a simple reading-comprehension question, passed a CAPTCHA check to verify they are human, completed the study on a computer rather than a mobile device, and if they had not participated in one of our previous studies. After applying our pre-registered exclusion criteria (duplicate IP addresses, duplicate Prolific participant IDs, or duplicate geocodes; people who failed to answer all measures for at

least one falsehood), 747 people remained (438 women, 297 men, 12 did not report gender; $M_{\text{age}} = 36$, $SD_{\text{age}} = 14$). Of these participants, 396 reported that they considered themselves or leaned Democrat and considered themselves or leaned 351 Republican. Participants who neither considered themselves nor leaned towards either party were prevented from starting the study.

Statistical power. A sensitivity analysis with the PANGEA app showed that, with 747 participants, eight repeated measures, and two tailed $\alpha = .05$, Study 6 provided >99% power to detect an effect size of $d = .2$ of the prefactual manipulation.

Materials. The stimuli were eight political falsehoods, and for each falsehood, a fact that contradicted it, a prefactual about how it might become true, a prefactual that was *unrelated* to how it might become true, and a description of the falsehood's gist (see Appendix C for all stimuli). These descriptions were written by the first author and a research assistant following the definition of gist as "the broader meaning that the statement communicates." For example, the gist description of "Every day 500 people die from gun violence in the United States" was "Many lives are lost to gun violence in the United States." Half of the falsehood-prefactual pairs fit with political beliefs associated with Republicans and half fit with political beliefs associated with Democrats.

Procedure. The procedure was identical to Study 4, except that we used different stimuli and included additional measures. Before beginning the study, participants responded to the comprehension check and reported their political affiliation on the same question from Studies 3-5. We prevented participants from beginning the study if they failed the comprehension check or if they did not consider themselves or lean Democrat or Republican.

Next, participants read that they would be considering "a number of facts that have been verified by reputable, non-partisan fact-checking websites" and then seeing and rating some

statements; they also read that at the end of the study we would ask true/false questions to test their understanding of these facts. Participants read a factual statement about a political issue (e.g., “It’s a proven fact that the number of people who die from gun violence each day in the United States is LESS than 500”). Each fact was accompanied by a statement indicating “this fact has been verified by reputable, non-partisan fact-checkers”. Then, participants randomly assigned to the *relevant-prefactual condition* read an if-then statement about how a falsehood that contradicted the fact might become true (e.g., “If Republicans loosen gun protection laws, then every day 500 people might die from gun violence”). Participants in the *irrelevant-prefactual condition* read an if-then statement that did not involve the falsehood becoming true (e.g., “If Republicans loosen gun protection laws, then the NRA might throw a celebration at its annual convention”). The *if* part of the prefactual statement was identical in both conditions, and in both conditions participants rated how easy it was to imagine the prefactual.

After considering the relevant or irrelevant prefactual, participants read the falsehood (e.g., “Every day 500 people die from gun violence in the United States.”) and rated its verbatim truthfulness, gist truthfulness, unethicity, and responded to the ‘liking’ measure (see below). Participants repeated this procedure for the remaining seven falsehoods in randomized orders, responded to a fact-check measure (described below), and provided demographics.

Measures.

Ease of imagining the prefactual: Prefactual potency. Participants separately indicated how likely they found the *if* component and *then* component of each prefactual they saw on the same measures as Studies 3 and 4. As in Studies 3 and 4, we multiplied judgments of the *if* and *then* likelihood together into a single measure of plausibility called prefactual potency (Petrocelli et al., 2011, 2012).

Gist truthfulness. To measure the falsehood's gist truthfulness, participants rated the following: "Consider the statement's broader message: [gist statement]. How much do you agree or disagree that this broader message is true?" (from -3 = *Strongly disagree* to 3 = *Strongly agree*).

Verbatim truthfulness. Participants rated the truthfulness of the falsehood's verbatim details: "How much do you agree or disagree that this statement is literally and precisely true?" (from -3 = *Strongly disagree* to 3 = *Strongly agree*)

Dependent measure: Unethicality of telling the falsehood. Participants responded to a single-item measure of the unethicality of telling the falsehood: How unethical would it be to make that statement? (from 0 = *Not at all unethical* to 100 = *Extremely unethical*).

"Liking". Study 6 included a new behavioral measure of participants' inclination to promote the falsehood – a lab analogue to "liking" content on social media. After reading each falsehood, participants had an opportunity to click a button to "like" it. They were told that the content that receives the most likes would be shared with participants in our next study, and thus that their "likes will determine which statements future participants read."

Fact-check. Participants responded to eight true/false questions to indicate what they thought of the statements presented earlier in the study. We instructed participants to "Please answer the questions below to test your memory and judgment of the facts presented earlier in the study." We predicted that our results would remain reliable when excluding participants who failed the fact-check measure, indicating that they either did not remember or did not believe that the falsehoods presented earlier in the study were false.

Results

We had strong directional hypotheses so we pre-registered one-tailed tests for all confirmatory analyses (Cho & Abe, 2013), and two-tailed tests for non-preregistered analyses.

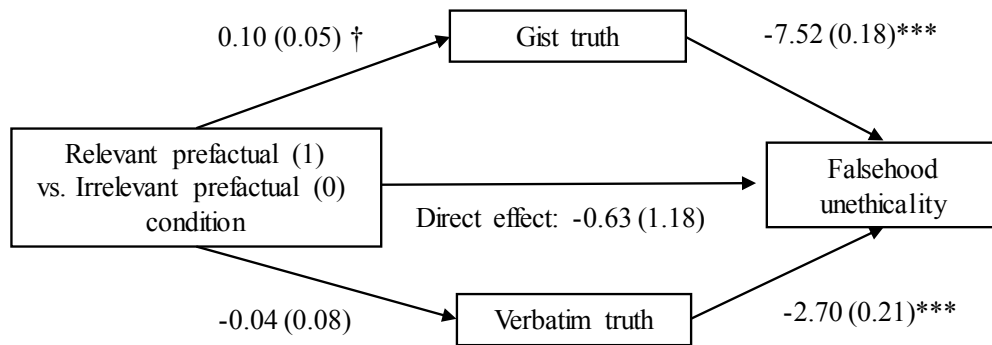
Gist truthfulness mechanism. The main aim of Study 6 was to test the theorized mechanism of gist truthfulness and distinguish this mechanism from perceptions of verbatim truthfulness. Results supported the unique mediating role of perceived gist truthfulness. Imagining how a falsehood might become true made the gist of the falsehood seem truer (relevant-prefactual: $M = -0.01$, $SD = 0.67$; irrelevant-prefactual: $M = -0.11$, $SD = 0.73$), $d = 0.14$, $z = 1.94$, $p = .027$, but did not make the falsehood seem verbatim truer (relevant-prefactual: $M = -1.28$, $SD = 1.05$; irrelevant-prefactual: $M = -1.24$, $SD = 1.05$), $d = -0.04$, $z = -0.57$, $p = .567$. These results are from mixed-effect regression models with condition as a fixed effect (1 = relevant prefactual, 0 = irrelevant prefactual), fixed effects for the eight items, and participant random effects.

As predicted, there was a significant indirect effect of the prefactual manipulation on unethicity through perceptions of gist truthfulness, but not verbatim truthfulness. When participants imagined how the falsehood might become true, they thought the gist of the falsehood was truer, and the truer they found the gist, the less unethical they thought it was to tell the falsehood – a significant indirect effect, $b = -0.74$, one-tailed 95% CI $[-\infty, -0.11]$ (see Figure 8). By contrast, there was not a significant indirect effect through perceptions of verbatim truthfulness, $b = 0.12$, 95% CI $[-0.29, 0.53]$ (this indirect effect was not significant by one- or two-tailed tests). When we included gist truthfulness and verbatim truthfulness in the model predicting unethicity judgments, the direct effect of the prefactual condition was not significant, $b = -0.63$, 95% CI $[-2.93, 1.68]$.⁹ We conducted this analysis as a generalized

⁹ Similarly, this direct effect of the prefactual condition was also not significant when we included gist truthfulness, but not verbatim truthfulness, in the model predicting unethicity judgments, $b = -0.41$, 95% CI $[-2.73, 1.92]$.

structural equation model with prefactual condition as the independent variable (1 = relevant prefactual, 0 = irrelevant prefactual), gist truthfulness and verbatim truthfulness as parallel mediators, and perceptions of unethicity as the dependent variable, with random effects for participants and fixed effects for item to account for the data's multilevel structure. We computed the indirect effects by multiplying the *a*- and *b*-paths together using the *gsem* and *nlcom* functions in Stata. Together, these results supported our theorizing that imagining how a falsehood might become true makes it seem less unethical to tell not by making the falsehood's verbatim details seem truer, but by making the broader message it conveys—the gist—seem truer.

In contrast to previous studies, the total effect of the prefactual manipulation was not significant in this study (relevant-prefactual condition: $M = 50.36$, $SD = 18.13$; irrelevant-prefactual condition: $M = 51.63$, $SD = 17.25$), $d = -0.07$, $z = -0.98$, $p = .163$. This result is from a mixed-effect regression model, with condition as a fixed effect (1 = relevant prefactual, 0 = irrelevant prefactual), participant random effects, and fixed effects for the eight items (see Table 5; step 1). As discussed below, this result could be due to random error variance across studies, or a result of prompting participants to focus on the falsehood's verbatim truth prior to judging its unethicity in this study.

Figure 8*Indirect Effect of Prefactual Manipulation on Unethicality Judgments through Gist in Study 6*

Note. Unstandardized coefficients shown. Indirect effect through perceptions of gist truthfulness: $b = -0.74$, one-tailed 95% CI $[-\infty, -0.11]$. Indirect effect through perceptions of verbatim truthfulness: $b = 0.12$, 95% CI $[-0.29, 0.53]$. Model includes item fixed effects and participant random effects.

* $p < .05$; ** $p < .01$, *** $p < .001$ by two-tailed tests

Moderation by political fit. Supporting H3, the effect of the prefactual manipulation was significantly larger when the falsehood-prefactual pair *fit*, rather than *conflicted*, with participants' politics, $b = -2.61$, $z = -1.87$, $p = .031$ (see Figure 9). For this analysis, we added a dummy code for political fit and its interaction with condition to the mixed model predicting unethicity judgments (see Table 5; step 2). These results are consistent with the hypothesis that prefactual thinking would have a greater effect on reducing people's condemnation of falsehoods when that possibility that the falsehood might become true fits with an individual's politics.

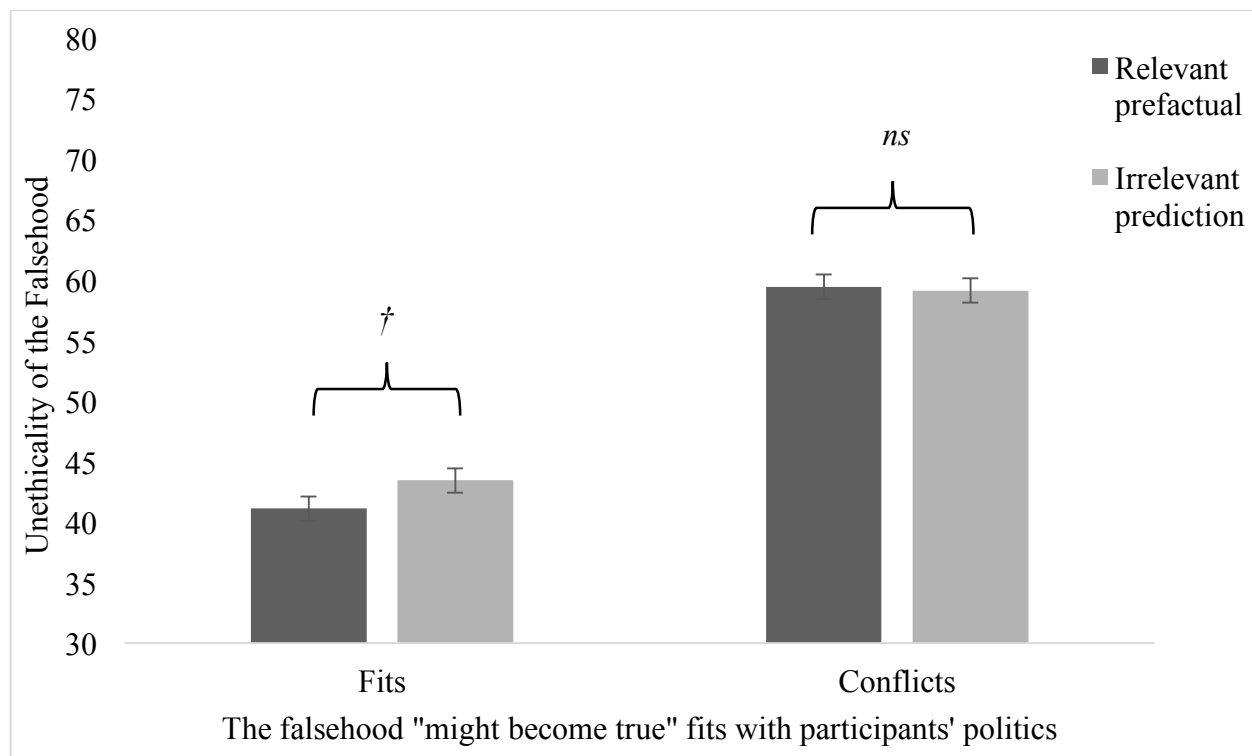
Decomposing this interaction with simple slopes analyses, when the falsehood-prefactual pair *fit* with participants' politics, considering a relevant prefactual led participants to rate the falsehood as less unethical to tell ($M = 40.86$, $SD = 22.83$) than considering an irrelevant prefactual ($M = 42.57$, $SD = 20.27$), $d = -0.08$, $z = -1.76$. This effect was only marginally significant, two-tailed $p = .078$, but – following pre-registered robustness checks – was statistically significant when including only responses that corresponded to correct fact-checks, $b = -3.71$, $z = -2.36$, $p = .018$, or when retaining all data but statistically controlled for the fact-check measure, $b = -2.80$, $z = -1.97$, $p = .049$. Thus, we can be most confident that prefactual thinking reduces condemnation of falsehoods that fit with one's politics when people acknowledge that the falsehood is false.

When the falsehood-prefactual pair *conflicted* with participants' political views, considering a relevant prefactual did not lead participants to rate the falsehood as less unethical to tell ($M = 59.86$, $SD = 22.49$) than considering an irrelevant prefactual ($M = 60.69$, $SD = 22.52$), $d = -0.04$, $z = 0.04$, $p = .968$. These tests are two-tailed, because we did not pre-register hypotheses for these simple slopes.

Table 5*Regression Analyses Predicting Unethicality Ratings in Study 6*

Predictor	<i>b</i>	<i>SE(b)</i>	<i>z</i>	<i>p</i>	95% CI of <i>b</i>
Step 1					
Condition	-1.27	1.29	-0.98	.163	$[-\infty, 0.86]$
Constant	31.73	1.32	24.13	<.001	[29.15, 34.31]
Step 2					
Condition	0.06	1.44	0.04	0.968	[-2.77, 2.89]
Political fit	-15.69	0.98	-15.96	<.001	[-17.62, 13.76]
Condition x political fit	-2.61	1.40	-1.87	.031	$[-\infty, -0.31]$
(Constant)	40.07	1.37	29.34	<.001	[37.40, 42.75]

Note. Condition is coded 1 = relevant prefactual, 0 = irrelevant prefactual. Political fit with “it might become true” coded 1 = fits, 0 = conflicts, with a participant’s political beliefs. The mixed regression model also included participant random effects and item fixed effects. We report one-tailed 95% confidence intervals for pre-registered one-tailed tests. Values of ∞ reflect that one-tailed tests, by definition, cannot detect an effect in the opposite direction as predicted.

Figure 9*Mean Unethicality Ratings by Condition and Political Fit in Study 6*

Note. Full scale of unethicality ratings is 0-100. Plotted values are the estimated marginal means and their standard errors from the mixed regression model described in the main text.

† $p < .10$.

Ease of imagining the prefactual: Prefactual potency. Results supported our theorizing that it would be easier to imagine a falsehood becoming true if that possibility fits with one's pre-existing motivations and beliefs. Participants in the relevant-prefactual condition rated the prefactuals as more plausible when the falsehood-prefactual pair fit ($M = 44.09$, $SD = 20.04$) rather than conflicted ($M = 23.59$, $SD = 14.57$) with their politics, $d_z = 0.88$, $z = 21.21$, $p < .001$. We examined this difference in a mixed-effect regression model, with political fit as a fixed effect (1 = fits, 0 = conflicts), participant random effects, and item fixed effects. This analysis was limited to participants in the relevant-prefactual condition because they were the only ones who saw and rated prefactuals about how the falsehood might become true

Moreover, in an exploratory analysis, we examined whether prefactual vividness and gist mediated the effect of political fit on judgments of unethicity in the relevant-prefactual condition. Our theorizing predicts that participants would find it easier to imagine that a falsehood might become true when that possibility fit with their politics, and the easier the prefactual was to imagine, the more truthful the gist of the falsehood would seem, and the less unethical they would judge the falsehood. In line with our theorizing and results of Study 5, there was a significant indirect effect from political fit to potency, to gist, to ratings of unethicity in a serial mediation model, $b = -3.74$, 95% CI $[-4.31, -3.17]$, computed with item fixed effects and participant random effects, an independent covariance structure (Stata's default), and two-tailed test because the analysis was not pre-registered. This analysis was again limited to participants in the relevant-prefactual condition.

“Liking.” We did not find evidence that the prefactual manipulation increased participants' likelihood of promoting the falsehood. Participants in the Prefactual condition were not significantly more likely to “like” the falsehood (20.63%) than participants in the Control

condition (24.21%), $b = -0.31$, $z = -1.97$, $p = .976$, in a mixed-effects logistic regression model, with condition as a fixed effect (1 = prefactual condition, 0 = control condition), participant random effects, and fixed effects for the eight items. Similarly, there was not a significant indirect effect of the prefactual manipulation on liking through unethicity ratings (see Online Supplement).

Fact-check. We hypothesized that relevant prefactuals would lead people to judge the falsehoods as less unethical to tell by making the gist of the falsehood seem truer, despite them acknowledging that the falsehoods were false. The fact-check measure showed that people correctly differentiated fact from falsehood 80% of the time – a proportion that did not differ significantly between the prefactual (80.79%) and control (78.80%) conditions, $b = 0.16$, $z = 1.44$, $p = 0.151$ (we did not pre-register this analysis, so this is a two-tailed test). This analysis was conducted in a mixed logistic regression analysis with random intercepts for participants. Furthermore, there was not a significant interaction between prefactual condition and political fit on participants' responses to the fact-check measure, $b = -0.04$, $z = -0.30$, $p = .768$, when we add a dummy-code for political fit and its interaction with the prefactual manipulation to the mixed logistic regression model.

As a robustness check, we repeated our prior analyses (a) with only responses to the dependent measures that corresponded to correct fact-checks, and (b) when controlling for the fact-check measure. All results remained in the same direction and same statistical significance except where noted in the main text above (see Online Supplement).

Discussion

Results from Study 6 address its two main aims. First, our findings conceptually replicate the mediation findings from Study 2 and Study 5 and support our theorizing about the

mechanism of the prefactual effect (H2). Considering a prefactual about how a falsehood might become true in the future made the gist of the falsehood seem truer, which predicted rating the falsehood as less unethical to tell. This result emerged even when accounting for judgments of verbatim truthfulness. Moreover, we did not find evidence that prefactual thinking made participants perceive greater verbatim truthfulness of the falsehood. These findings suggest that it is specifically by making the gist of the falsehood seem truer that relevant prefactuals make falsehoods seem less unethical to tell.

Second, results provide additional support for the moderating role of whether the falsehood becoming true in the future fit with people's pre-existing motivations and beliefs (H3). Results were consistent with our theorizing that this moderation occurred because people's pre-existing motivations and beliefs affected their ease of imagining the prefactual. Participants rated prefactuals as more plausible when they fit with their politics, and the more plausible they found the prefactuals, the more they thought the gist of the falsehoods were true, and the less unethical they found the falsehoods.

Interestingly, although we found an indirect effect from prefactual condition to perceptions of gist truth to perceptions of unethicality, we did not find a main effect of prefactual condition on perceptions of unethicality. However, we did find evidence that prefactuals reduce condemnation of falsehoods amongst falsehoods that fit with participants' political beliefs, consistent with past studies. One difference between this study and our previous studies is that in Study 6, participants rated the degree to which the falsehood was verbatim—literally and precisely—true before rating the unethicality of the falsehood. Having participants rate the verbatim truth of the falsehood prior to judging its unethicality may have discouraged participants from condoning statements that were verbatim false. Future research might examine

whether prompting people to reflect on the verbatim details of a falsehood reduces the extent to which imagining how the falsehood might become true makes the falsehood seem less unethical to tell.

One limitation of this study is that we found no evidence that prefactuals increased participants' promoting of the falsehood to other research participants, perhaps because our behavioral measure was not particularly sensitive. Future research might examine how imagining how a falsehood might become true affects people's willingness to promote falsehoods in the real-world context of social media, and especially amongst partisan ingroup members.

Meta-Analysis

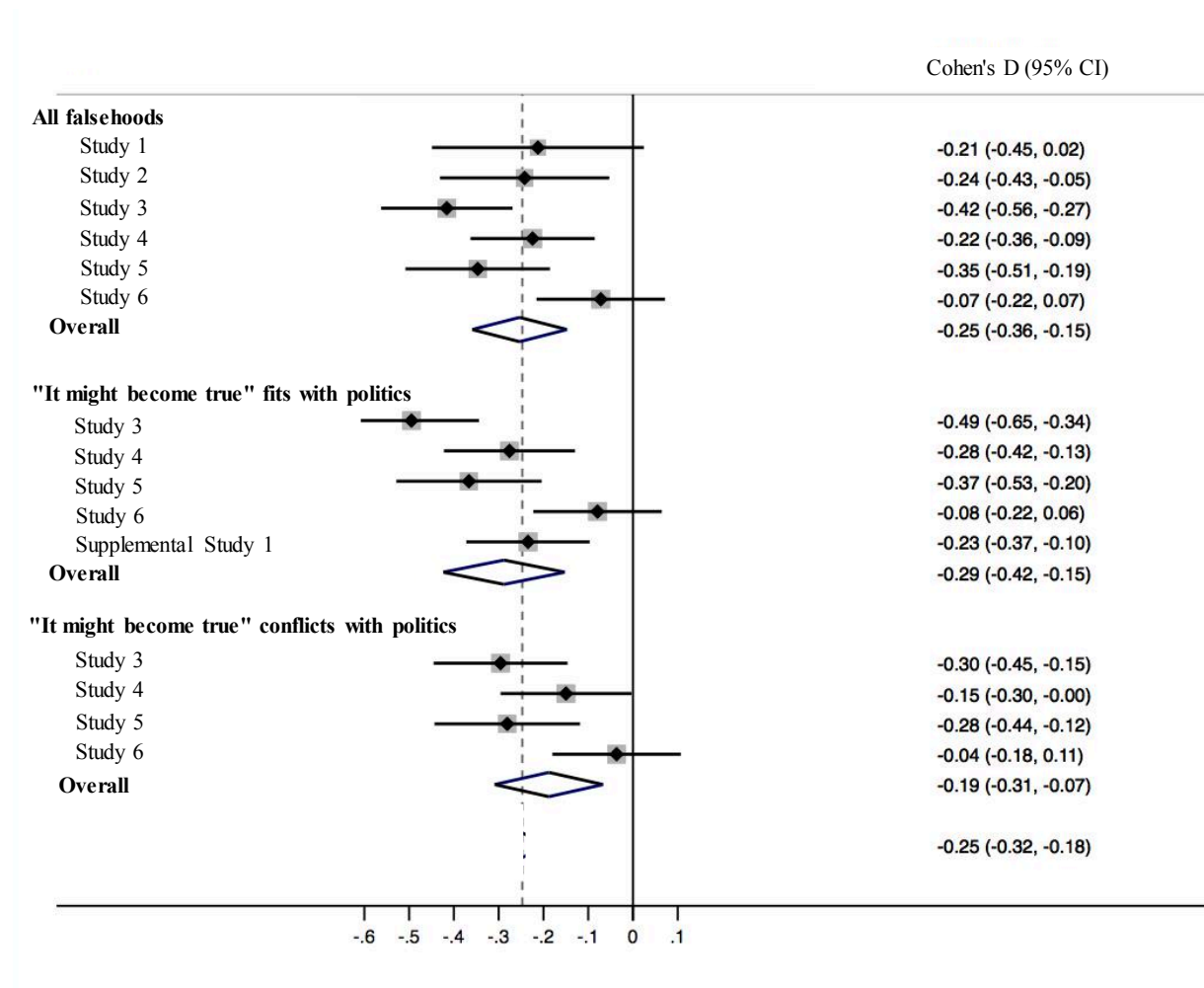
To inform our discussion of our studies, we next report a meta-analysis of their results. Our goal is to estimate the size of the prefactual effect overall, as well as when the falsehood-prefactual pair was aligned versus misaligned with participants' political views. In addition to the six studies reported in the main text, the analysis included a Supplemental Study, discussed in the General Discussion and reported in full in the Online Supplement, in which all the falsehood-prefactual pairs were aligned with participants' politics. We decided *a priori* to treat study as a random effect in the meta-analysis, because the methods and stimuli differed across studies.

Results from this meta-analysis show that imagining how a falsehood might become true reduced condemnation of the falsehood, with an effect that was significant and modest in size, $d = -0.25$, 95% CI $[-0.36, -0.15]$ (See Figure 10), putting it between the 35th and 40th percentile for social-psychological effects that have been meta-analyzed. As a benchmark, the median effect size in meta-analyzed research on interpersonal relationships is $d = .28$ (Lovakov & Agadullina, 2021).

The effect size was descriptively larger when the falsehood-prefactual pairs fit with participants' political views, $d = -0.29$, 95% CI $[-0.42, -0.15]$, and smaller when the falsehood-prefactual pairs did not fit with their views, $d = -0.19$, 95% CI $[-0.31, -0.07]$. Recall that across Studies 3-6, the evidence was mixed about whether the prefactual effect would emerge at all when the prefactuals and falsehoods did not fit with participants' politics. The meta-analysis suggests that the prefactual effect is indeed reliable even in that case.

Figure 10

Meta-Analysis: Imagining How a Falsehood Might Become True Reduces the Condemnation It Receives



Note. The dashed line illustrates the estimate of the overall effect size. Lines show 95% CIs for the individual studies and diamonds show 95% CIs for meta-analytic effects. The 95% CIs underestimate statistical significance for the studies in which we pre-registered one-tailed tests (Studies 1, 4, 5, and 6). The size of the grey squares reflects the sample size of the study.

General Discussion

Overall, the evidence from our six studies (four pre-registered) suggests that when people consider how a falsehood might become true in the future, they think the falsehood is not so bad to tell in the present (H1). This effect emerged with participants from 59 countries judging false claims about consumer products, job seekers' work experience, and controversial political issues. This effect also emerged regardless of whether participants were provided with specific prefactuals (Studies 1–4 and 6) or generated the prefactuals themselves (Study 5), and it may have important downstream consequences: The less unethical participants found the falsehood, the less inclined they were to censure someone who shared the falsehood on social media, and the more inclined they were to share the falsehood themselves (Studies 4 and 5). Together, these results suggest that imagining prefactuals can license dishonesty.

The results also shed light on why. Imagining how a falsehood might become true led people to view the falsehood's gist as truer—and the truer they found the gist, the less unethical they judged the falsehood (H2; Studies 2, 5, and 6). This mediation effect emerged with two different measures of how true people perceived the falsehood's gist, and occurred independently of how true people found the falsehood's verbatim details. Moreover, this mediation occurred even among the large majority of participants who correctly acknowledged that the falsehood's literal claims were incorrect. Thus, prefactual thinking can make a falsehood's broader message seem truer, even when one knows the falsehood's specific details are false, which in turn may mitigate moral condemnation of the falsehood. Consider a person who imagines how the falsehood “the average CEO makes 500 times more money than the average American worker” might become true in the future. Our results suggest that this person would not come to believe that the pay disparity between CEOs and workers is *actually* this large, but would become more

convinced of the general idea that “CEOs make much more money than workers” – and thus judge the falsehood as less unethical. Although our mediation analyses cannot specify the causal order of the gist mediator and judgments of unethicality (Bullock et al., 2010; Fiedler et al., 2011), our results are nonetheless consistent with this interpretation.

The studies also suggest that prefactuals help excuse falsehoods especially—but not exclusively—when people are inclined to believe that the falsehood might become true because it fits with their pre-existing motivations and beliefs (H3). Specifically, prefactuals more effectively reduced people’s moral condemnation of falsehoods when the possibility the falsehood might become true fit, rather than conflicted, with their politics. This moderation effect was statistically significant in the three studies in which participants considered pre-written prefactuals (Studies 3, 4, and 6), and directionally consistent but not significant in the study in which participants generated their own prefactuals (Study 5, $p = .071$). Thus, we can be most confident in the robustness of this moderation effect when people consider prefactuals that others have provided.

We hypothesized that this moderating role of political fit occurs because prefactuals are easier to imagine when they fit with one’s motivations and beliefs (Effron, 2018; Tetlock, 1998; Tetlock & Henik, 2007), and because the easier a prefactual is to imagine, the larger its effect on judgment (Gaesser et al., 2018, 2019; Petrocelli et al., 2011, 2012; Sherman et al., 1985). Our results were consistent with this theorizing regardless of whether we operationalized ease of imagination as judgments of plausibility or vividness of mental simulation. Specifically, when the possibility that the falsehood might become true fit with participants’ politics, they judged a prefactual in which it became true in the future as more plausible (Studies 3, 4, and 6), and they said they had imagined this prefactual more vividly (Study 5); moreover, the more plausible and

vivid they found the prefactual, the more they perceived the gist of the falsehood as true, and the less unethical they found the falsehood. Together, our findings suggest that prefactuals may be particularly effective in helping people excuse falsehoods that are consistent with what they want to believe. As a result, prefactuals amplified partisan disagreement in moral judgments of politically-charged falsehoods.

Together, our results suggest that prefactuals offer people a degree of freedom when judging morality. It may be relatively hard to convince yourself that a falsehood is actually true, even if you are motivated to do so. It may be comparatively easy to convince yourself that the falsehood *might become* true. Thus, people's ability to imagine the future in a way that supports their present beliefs may help them to excuse falsehoods that fit with their politics.

Alternative Explanations

We rule out four alternative explanations for our findings. First, we find no evidence that our effects occurred because people believed that the falsehoods we presented were factual rather than fictional. Most participants accurately identified the falsehoods as false when asked to do so at the end of each study, and the results remained robust when we dropped the small number of cases in which participants believed a falsehood was true. Moreover, we find no evidence that prefactual thinking affected people's beliefs that the falsehood was literally true when we include a continuous measure of verbatim truthfulness (Study 6). Together, these findings suggest that our results did not occur because participants rejected the facts we presented them nor because prefactual thinking led people to forget that the falsehood were false (Murphy et al., 2019).

Second, the results cannot be explained by the possibility that the prefactuals communicated *factual* information that made the falsehood seem more justified. Participants who considered how the falsehood might become true judged the falsehood as less unethical to tell

than participants in the control condition even when we included any factual information from the prefactual condition in the control condition (Study 4 and 6) and when participants in the prefactual condition generated their own prefactuals rather than reading a given prefactual (Study 5).

Third, our findings cannot be explained by participants interpreting the speakers' claims as commitments for the future, rather than a false claim about the present. Across studies, we did not tell participants that the individual who told the falsehood claimed that it might become true in the future; we simply offered the prefactual as a possible "prediction" that was not attributed to any particular person, and about which participants could indicate agreement or disagreement. Moreover, in Studies 3-6, participants judged falsehoods about events beyond anyone's personal control (e.g., the number of illegal voters or average CEO compensation), including falsehoods about undesirable events. Thus, it is unlikely that participants interpreted the speaker's claims as promises for the future.

Fourth, we cannot account for the results by positing that *any* sort of prefactual thinking affects moral judgments by putting people in a *mental simulation mindset* (Galinsky & Moskowitz, 2000; Hirt et al., 2004; Kray et al., 2006; Wong et al., 2008). We find that a *specific* type of prefactual thought reduces moral condemnation of a falsehood: imagining how that falsehood might become true. People expressed less condemnation of falsehoods when they imagined futures in which the falsehoods became true as opposed to futures that were unrelated to the falsehoods' truth (Studies 4-6).

Theoretical Contributions

Our findings make several contributions to the literatures on mental simulation and moral psychology. Whereas past research on prefactual thinking emphasizes its functional outcomes,

such as helping people to plan for the future (e.g., Epstude et al., 2016), we reveal a dysfunctional outcome – that prefactuals can encourage people to relax their moral standards about lying. A just and well-functioning society arguably depends on holding people accountable for spreading false claims in the context of advertisements, work, and politics. Yet our studies found that prefactual thinking can not only reduce how much people condemn falsehoods in these contexts, but also increase people’s inclination to spread such falsehoods themselves.

We also contribute to work on moral flexibility, which demonstrates that people apply their moral standards inconsistently so that they can let themselves and the people they like off the hook for bad behavior (see Bartels et al., 2015; Effron, 2016; Gino, 2016; Uhlmann et al., 2009). Past research has shown that mental simulation facilitates moral flexibility in the self (Briazu et al., 2017; Shalvi et al., 2011). Yet, motivation has been absent from most previous research on other-oriented mental simulation (e.g., Alicke et al., 2008; MacRae, 1992; Miller et al., 2005). The present research demonstrates how mental simulation facilitates moral flexibility in judging others (see also Effron, 2018).

Moreover, whereas past work demonstrates how moral flexibility can result from selective memory for the past (Kouchaki & Gino, 2016; Shu et al., 2011; Shu & Gino, 2012), our results suggest that it also results from imagination of the future. Prefactual thinking may offer a particularly appealing strategy for motivated moral reasoning because, unlike memories for the past, predictions about the future cannot be fact-checked in the present. Vague predictions may even be impossible to falsify. Thus, a person whose politics lead them to believe that *a lie will become true eventually* may be difficult to convince otherwise.

Our findings also advance understanding of how mental simulation affects moral judgments. Prior work in this area focused primarily on counterfactual thinking (Alicke et al.,

2008; Branscombe et al., 1996; Byrne, 2017; MacRae, 1992; Mandel & Dhami, 2005; Parkinson & Byrne, 2017), neglecting prefactual thinking. For example, prior work finds that people judge falsehoods more leniently when they imagine that they *would have been* true if circumstances had been different (Briazu et al., 2017; Shalvi et al., 2011) – particularly if they like the falsehoods (Effron, 2018). Our results suggest, however, that to justify giving a moral pass to a falsehood, people need not mentally undo its falsity in the past; they need only imagine it might become true in future. In this way, mental simulation in general—and not just counterfactual thinking in particular—can facilitate moral flexibility.

Whereas previous research on counterfactual thinking demonstrates that imagining how a falsehood could have been true makes people judge the falsehood as less unethical to tell (Effron, 2018) and be more likely to tell them (Shalvi et al., 2011), our studies are the first to our knowledge to examine *how* imagination affects moral judgments of falsehoods. Our findings suggest that mentally simulating a falsehood makes people perceive the gist of the falsehood as truer, which was associated with reduced perceptions of the falsehood’s unethicality. More broadly, our findings suggest that people judge falsehoods not only on their literal content, but also on what they view to be the broader meaning communicated by the claim.

Our findings also have implications for understanding why misinformation spreads. Current research on this question assumes that people spread misinformation because they believe it, or fail to think carefully about its accuracy (e.g., Bago et al., 2020; Lazer et al., 2018; Newman et al., 2020; Scheufele & Krause, 2019). By contrast, we argue that people sometimes spread misinformation because they find it morally permissible. Indeed, Studies 4 and 5 found that when considering how a falsehood might become true made a falsehood seem less unethical to tell, people expressed stronger intentions to spread the falsehood on social media, particularly

when the possibility that the falsehood might become true fit with their political beliefs. In this way, our research contributes to a growing body of work examining when and why individuals excuse falsehoods despite knowing they are false (e.g., Effron, 2018; Effron & Raj, 2020; Levine & Schweitzer, 2014).

Finally, our work speaks to the growing political divide in the United States. Some pundits suggest that American partisans suffer from a “reality gap” (Neimand & Griffin, 2017). Indeed, research reveals that Americans on opposite ends of the political spectrum disagree about basic facts (Kteily et al., 2016). Our research highlights a different source of disagreement. Even when political partisans in our studies agreed that a statement was not factual, they disagreed about whether it was prefactual. That is, considering how a falsehood might become true in the future provided a more effective justification for excusing it when participants’ politics led them to easily imagine that the falsehood might become true. As a result, relevant prefactuals amplified partisan differences in condemning falsehoods (see Figures 2, 3, 6, and 9).

Limitations and Future Directions

Our research has several limitations that provide opportunities for future research. First, future research is needed to explore the generalizability of our findings to a broader range of samples and falsehoods. Our results generalize to multinational MBA students from 59 different countries, residents of a major UK city, and Americans from both sides of the political spectrum on two online sampling serves (MTurk and Prolific Academic); however, future research should examine these effects in broader participant samples. Moreover, our findings generalize to falsehoods in the contexts of consumer products, work experiences, and political issues; however, future research should examine whether our observed effects are limited to falsehoods with particular characteristics. For example, prefactual thinking may not reduce condemnation of

falsehoods that have immediate and severe negative consequences or falsehoods that people fail to believe might become true. If people find it inconceivable that a falsehood might become true, then trying and failing to imagine it may make the falsehood seem more, rather than less, unethical (for reversal effects in imagination, see Sherman et al., 1985).

Second, future research should examine how our findings extend beyond moral judgments and behavioral intentions in experimental settings to real behaviors of promoting falsehoods online. We find evidence in two studies that imagining how a falsehood might become true increases people's intentions to share the falsehoods online (Study 4 and 5), and recent research has shown that these measures of behavioral intentions to share content online are positively associated with actual sharing of content on Twitter (Mosleh et al., 2020). Moreover, in a supplemental study, we find that imagining how a falsehood might become true also leads to greater intentions to promote the falsehood on the popular discussion website, Reddit (see Study S1 in Online Supplement). However, we did not find evidence that prefactual thinking led participants to promote the falsehood to other research participants by 'liking' it (Study 6). One possibility is that our lab-based measure was not sensitive enough to capture 'liking' behavior. Another possibility is that lab-based behavioral measures of promoting content to other participants differ significantly from real-world liking behavior in which people promote content to a curated audience of likeminded individuals (Bakshy et al., 2015). Future research might examine whether prefactual thinking increases sharing of falsehoods on social media sites such as Twitter and Facebook.

Third, future research should investigate additional processes by which prefactual thinking can make falsehoods seem less unethical. Our results were consistent with the proposed gist truthfulness mechanism, but do not preclude the possibility that other, unmeasured mediators

may operate as well (Rucker et al., 2011; Zhao et al., 2010).¹⁰ One possibility is that relevant prefactuals can make a falsehood seem justified by increasing people's confidence that it will indeed become true in the future, even without making its gist seem truer in the present.

Fourth, we studied the causal effects of prefactual thinking by randomly assigning some participants to consider how a falsehood might become true. Future research should examine when and how people spontaneously generate these prefactual thoughts. Our research was inspired by real cases in which leaders encouraged others to imagine how their falsehoods might become true in the future. For example, when Donald Trump was criticized for falsely claiming that COVID-19 cases were on a downward trend in March, 2020, Trump encouraged supporters to imagine how his falsehood might become true in the future—"I'll be right eventually. [COVID-19 is] going to disappear." (Fox News, 2020)—and his press secretary did the same—"No one is lying to the American people. One day, COVID will go away." (McEnany, 2020). Similarly, entrepreneurs in Silicon Valley use lofty dreams of what a product *might do in the future* to justify deceiving investors about the truth of the product's current capabilities (Carreyrou, 2018). These examples highlight another important avenue for future research. Whereas we examined how prefactual thinking affected participants' judgments of the unethicity of others' falsehoods, future research might examine how prefactuals affect people's willingness to tell falsehoods themselves. For example, entrepreneurs may be more willing to lie about their product's capabilities and job applicants may be more willing to lie about their professional skills when they can imagine that those lies might become true in the future.

¹⁰ We find mixed evidence for a direct effect of the prefactual condition on judgments of unethicity, when controlling for perceptions of gist truthfulness: $b = -2.07$, 95% CI $[-4.76, 0.63]$ (Study 2), $b = -4.37$, 95% CI $[-6.88, -1.87]$ (Study 5), and $b = -0.63$, 95% CI $[-2.93, 1.68]$ (Study 6).

Fifth, our results reveal that the more truthful people found a falsehood's gist, the less unethical they thought it was to tell, despite recognizing the falsehood as literally untrue. Future research should examine the causal relationship between perceptions of gist truthfulness and moral judgments of falsehoods, and explore what other factors besides prefactual thinking can make a statement that is literally false seem true in gist.

Finally, future research should examine how to mitigate the effect of prefactuals on excusing falsehoods. Excusing and propagating falsehoods has dangerous consequences. In the case of Theranos, employees excusing Holmes' lies could have threatened the lives of patients: "people would have died from missed diagnoses or wrong medical treatments" (Carreyrou, 2018, p. 708). Similarly, Trump's false claims about COVID-19 spread misinformation that may have hindered public responses to reduce the spread of the virus. In a supplementary study (see Study S1 in Online Supplement), we tested whether encouraging people to think carefully and deliberately could attenuate the effect of prefactual thinking on excusing falsehoods, but found that this effect persisted irrespective of whether participants thought intuitively or deliberately. Alternatively, prompting people to focus on the literal and precise truth of a statement may reduce the effect of prefactual thinking. For instance, in Study 6, participants rated the verbatim truth of the falsehood prior to rating its unethicality, and the estimate of the effect size of prefactual thinking on moral condemnation of falsehoods was smaller in this study than in each of our previous studies. Thus, future research should examine whether prompting people to focus on the literal and precise truth of statements reduces the effects of prefactual thinking.

Conclusion

In business and in politics, people seem to frequently get away with telling falsehoods, even when their lies are discovered. Pundits blame this apparent trend on society's loosening

grip on reality (e.g., *The Economist*, 2016). Yet when people let others off the hook for dishonesty, the reason may not only be that our society is post-truth; it may also be that people's focus is prefactual.

Acknowledgments

We thank Laura Giurge for helpful feedback on an earlier draft, and Justine Murray and Stephanie Rodriguez for helping compile and create the falsehoods used as stimuli in our studies.

References

- Abadie, M., & Camos, V. (2018). False memory at short and long term. *Journal of Experimental Psychology: General*, 148(8), 1312–1334. <https://doi.org/10.1037/xge0000526>
- Abadie, M., Gavard, E., & Guillaume, F. (2021). Verbatim and gist memory in aging. *Psychology and Aging*. <https://doi.org/10.1037/pag0000635>
- Abadie, M., Waroquier, L., & Terrier, P. (2013). Gist memory in the unconscious-thought effect. *Psychological Science*, 24(7), 1253–1259. <https://doi.org/10.1177/0956797612470958>
- Abadie, M., Waroquier, L., & Terrier, P. (2017). The role of gist and verbatim memory in complex decision making: Explaining the unconscious-thought effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(5), 694–705. <https://doi.org/10.1037/xlm0000336>
- Adolphs, R., Tranel, D., & Buchanan, T. W. (2005). Amygdala damage impairs emotional memory for gist but not details of complex stimuli. *Nature Neuroscience*, 8(4), 512–518. <https://doi.org/10.1038/nn1413>
- Alicke, M. D., Buckingham, J., Zell, E., & Davis, T. (2008). Culpable Control and Counterfactual Reasoning in the Psychology of Blame. *Personality and Social Psychology Bulletin*, 34(10), 1371–1381. <https://doi.org/10.1177/0146167208321594>
- Anderson, C. A. (1983). Imagination and expectation: The effect of imagining behavioral scripts on personal influences. *Journal of Personality and Social Psychology*, 45(2), 293–305. <https://doi.org/10.1037/0022-3514.45.2.293>
- Anduiza, E., Gallego, A., & Muñoz, J. (2013). Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption. *Comparative Political Studies*, 46(12), 1664–1692. <https://doi.org/10.1177/0010414013489081>

- Aquinas, T. (1273). *Summa theologiae*. Benzinger Brothers.
- Bacon, A. M., Walsh, C. R., & Briazu, R. A. (2020). Looking behind and looking ahead: Personality differences in counterfactual and prefactual thinking. *Imagination, Cognition and Personality, 40*(2), 111–125. <https://doi.org/10.1177/0276236620905810>
- Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General, 149*(8), 1608–1613. <https://doi.org/10.1037/xge0000729>
- Bagozzi, R. P., Moore, D. J., & Leone, L. (2004). Self-control and the self-regulation of dieting decisions: The role of prefactual attitudes, subjective norms, and resistance to temptation. *Basic and Applied Social Psychology, 26*(2), 199–213. https://doi.org/doi:10.1207/s15324834basp2602&3_7
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science, 348*(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
- Bartels, D. M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition, 108*(2), 381–417. <https://doi.org/10.1016/j.cognition.2008.03.001>
- Bartels, D. M., Bauman, C. W., Cushman, F. A., Pizarro, D. A., & McGraw, A. P. (2015). Moral judgment and decision making. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 478–515). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118468333.ch17>

- Boles, T. L., Croson, R. T. A., & 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>
- Bookbinder, S. H., & Brainerd, C. J. (2017). Emotionally negative pictures enhance gist memory. *Emotion*, 17(1), 102–119. <https://doi.org/10.1037/emo0000171>
- Brainerd, C. J., & Gordon, L. L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology*, 30(2), 163–177. <https://doi.org/10.1037/0012-1649.30.2.163>
- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, 10(1), 3–47. [https://doi.org/10.1016/0273-2297\(90\)90003-M](https://doi.org/10.1016/0273-2297(90)90003-M)
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11(5), 164–169. <https://doi.org/10.1111/1467-8721.00192>
- Brandts, J., & Charness, G. (2003). Truth or consequences: An experiment. *Management Science*, 49(1), 116–130. <https://doi.org/10.1287/mnsc.49.1.116.12755>
- Branscombe, N. R., Owen, S., Garstka, T. A., & Coleman, J. (1996). Rape and accident counterfactuals: Who might have done otherwise and would ot have changed the outcome? *Journal of Applied Social Psychology*, 26(12), 1042–1067. <https://doi.org/10.1111/j.1559-1816.1996.tb01124.x>
- Briazu, R. A., Walsh, C. R., Deeprose, C., & Ganis, G. (2017). Undoing the past in order to lie in the present: Counterfactual thinking and deceptive communication. *Cognition*, 161, 66–73. <https://doi.org/10.1016/j.cognition.2017.01.003>

- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, 98(4), 550–558.
<https://doi.org/10.1037/a0018933>
- Byrne, R. M. J. (2017). Counterfactual Thinking: From Logic to Morality. *Current Directions in Psychological Science*, 26(4), 314–322. <https://doi.org/10.1177/0963721417695617>
- Carli, L. L. (1999). Cognitive reconstruction, hindsight, and reactions to victims and perpetrators. *Personality and Social Psychology Bulletin*, 25(8), 966–979.
<https://doi.org/10.1177/01461672992511005>
- Carlson, E. (2015). Ethnic voting and accountability in Africa: A choice experiment in Uganda. *World Politics*, 67(2), 353–385.
- Carreyrou, J. (2018). *Bad Blood: Secrets and Lies in a Silicon Valley Startup*. Pan Macmillan.
- Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14(1), 88–96. [https://doi.org/10.1016/0022-1031\(78\)90062-8](https://doi.org/10.1016/0022-1031(78)90062-8)
- Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes*, 79(2), 115–153.
<https://doi.org/10.1006/obhd.1999.2841>
- Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research*, 66(9), 1261–1266.
<https://doi.org/10.1016/j.jbusres.2012.02.023>
- Clark, H. H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 14(1), 56–72. [https://doi.org/10.1016/S0022-5371\(75\)80006-5](https://doi.org/10.1016/S0022-5371(75)80006-5)

- Dennis, S. A., Goodson, B. M., & Pearson, C. A. (2019). Online worker fraud and evolving threats to the integrity of MTurk data: A discussion of virtual private servers and the limitations of IP-based screening procedures. *Behavioral Research in Accounting*, 32(1), 119–134. <https://doi.org/10.2308/bria-18-044>
- Effron, D. A. (2016). Beyond “being good frees us to be bad”: Moral self-licensing and the fabrication of moral credentials. In Prooijen, Jan-Willem van & P. A. M. van Lange (Eds.), *Cheating, corruption, and concealment: Roots of unethical behavior*. Cambridge, UK: Cambridge University 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.
- Effron, D. A., Miller, D. T., & Monin, B. (2012). Inventing racist roads not taken: The licensing effect of immoral counterfactual behaviors. *Journal of Personality and Social Psychology*, 103(6), 916–932. <https://doi.org/10.1037/a0030008>
- Effron, D. A., Monin, B., & Miller, D. T. (2013). The unhealthy road not taken: Licensing indulgence by exaggerating counterfactual sins. *Journal of Experimental Social Psychology*, 49(3), 573–578. <https://doi.org/10.1016/j.jesp.2012.08.012>
- Effron, D. A., & Raj, M. (2020). Misinformation and morality: Encountering fake-news headlines makes them seem less unethical to publish and share. *Psychological Science*, 31(1), 75–87. <https://doi.org/10.1177/0956797619887896>
- Epstude, K., Scholl, A., & Roese, N. J. (2016). Prefactual thoughts: Mental simulations about what might happen. *Review of General Psychology*, 20(1), 48–56. <https://doi.org/10.1037/gpr0000064>

- Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231–1236.
<https://doi.org/10.1016/j.jesp.2011.05.007>
- Fox News. (2020, July 19). *President Trump goes one-on-one with Chris Wallace | Full Interview*. <https://www.youtube.com/watch?v=W6XdpDOH1JA>
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108(3), 316–355.
<https://doi.org/10.1037/0096-3445.108.3.316>
- Fukukura, J., Ferguson, M. J., & Fujita, K. (2013). Psychological distance can improve decision making under information overload via gist memory. *Journal of Experimental Psychology: General*, 142(3), 658–665. <https://doi.org/10.1037/a0030730>
- Gaesser, B., Keeler, K., & Young, L. (2018). Moral imagination: Facilitating prosocial decision-making through scene imagery and theory of mind. *Cognition*, 171, 180–193.
<https://doi.org/10.1016/j.cognition.2017.11.004>
- Gaesser, B., Shimura, Y., & Cikara, M. (2019). Episodic simulation reduces intergroup bias in prosocial intentions and behavior. *Journal of Personality and Social Psychology*, 118(4), 683. <https://doi.org/10.1037/pspi0000194>
- Galinsky, A. D., & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology*, 36(4), 384–409. <https://doi.org/10.1006/jesp.1999.1409>
- Gerlach, K. D., Dornblaser, D. W., & Schacter, D. L. (2014). Adaptive constructive processes and memory accuracy: Consequences of counterfactual simulations in young and older adults. *Memory*, 22(1), 145–162. <https://doi.org/10.1080/09658211.2013.779381>

- Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition*, 8(2), 149–156. <https://doi.org/10.3758/BF03213418>
- Gibbs, R. W. (1987). Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(3), 524–533. <https://doi.org/10.1037/0278-7393.9.3.524>
- Gino, F. (2016). How moral flexibility constrains our moral compass. In J.-W. van Prooijen & P. A. M. van Lange (Eds.), *Cheating, Corruption, and Concealment*. Cambridge University Press.
- Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, 102(3), 445–459. <https://doi.org/10.1037/a0026406>
- Goerke, M., Möller, J., Schulz-Hardt, S., Napiersky, U., & Frey, D. (2004). “It’s not my fault-but only I can change it”: Counterfactual and prefactual thoughts of managers. *Journal of Applied Psychology*, 89(2), 279–292. <https://doi.org/10.1037/0021-9010.89.2.279>
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology*, 43(1), 89–99. <https://doi.org/10.1037/0022-3514.43.1.89>
- Grice, H. P. (1975). Logic and Conversation. *Speech Acts*, 41–58. https://doi.org/10.1163/9789004368811_003
- Hammell, C., & Chan, A. Y. C. (2016). Improving physical task performance with counterfactual and prefactual thinking. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0168181>

- Harris, S. (2013). *Lying*. Four Elephants Press.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107(3), 311–327.
<https://doi.org/10.1037/0033-2909.107.3.311>
- Hirt, E. R., Kardes, F. R., & Markman, K. D. (2004). Activating a mental simulation mind-set through generation of alternatives: Implications for debiasing in related and unrelated domains. *Journal of Experimental Social Psychology*, 40(3), 374–383.
<https://doi.org/10.1016/j.jesp.2003.07.009>
- Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 566–681.
- Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, 117(4), 371–376. <https://doi.org/10.1037/0096-3445.117.4.371>
- Kant, I. (1787). On a supposed right to lie from benevolent motives. In L. W. Beck (Ed.), *The critique of practical reason and other writings in moral philosophy* (pp. 346–350). University of Chicago Press.
- Kappes, H. B., & Morewedge, C. K. (2016). Mental simulation as substitute for experience. *Social and Personality Psychology Compass*, 10(7), 405–420.
<https://doi.org/10.1111/spc3.12257>
- Keck, S. (2014). Group reactions to dishonesty. *Organizational Behavior and Human Decision Processes*, 124(1), 1–10. <https://doi.org/10.1016/j.obhdp.2013.12.006>

- Kennedy, R., Clifford, S., Burleigh, T., Waggoner, P. D., Jewell, R., & Winter, N. J. G. (2020). The shape of and solutions to the MTurk quality crisis. *Political Science Research and Methods*, 8(4), 614–629. <https://doi.org/10.1017/psrm.2020.6>
- Keyes, R. (2004). *The Post-Truth Era: Dishonesty and Deception in Contemporary Life*. St. Martin's Press.
- Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94(2), 211–228. <https://doi.org/10.1037/0033-295X.94.2.211>
- Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, 110(3), 499–519. <https://doi.org/10.1037//0033-2909.110.3.499>
- Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences*, 113(22), 6166–6171. <https://doi.org/10.1073/pnas.1523586113>
- Kray, L. J., Galinsky, A. D., & Wong, E. M. (2006). Thinking within the box: The relational processing style elicited by counterfactual mind-sets. *Journal of Personality and Social Psychology*, 91(1), 33–48. <https://doi.org/10.1037/0022-3514.91.1.33>
- Kteily, N. S., Sheehy-Skeffington, J., & Ho, A. K. (2016). Hierarchy in the eye of the beholder: (Anti-)egalitarianism shapes perceived levels of social inequality. *Journal of Personality and Social Psychology*, 112(1), 136–159. <https://doi.org/10.1037/pspp0000097>
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>

- Lehrer, J. (1998, January 21). *Newshour with Jim Lehrer*. PBS.
- Lelieveld, G.-J., Shalvi, S., & Crone, E. A. (2016). Lies that feel honest: Dissociating between incentive and deviance processing when evaluating dishonesty. *Biological Psychology*, *117*, 100–107. <https://doi.org/10.1016/j.biopsycho.2016.03.009>
- Levine, 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, & 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>
- Louie, T. A. (2005). Hindsight bias and outcome-consistent thoughts when observing and making service provider decisions. *Organizational Behavior and Human Decision Processes*, *98*(1), 88–95. <https://doi.org/10.1016/j.obhdp.2005.03.004>
- Lovakov, A., & Agadullina, E. R. (2021). Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*, *51*(3), 485–504. <https://doi.org/10.1002/ejsp.2752>
- MacInnis, D. J., & Price, L. L. (1987). The role of imagery in information processing: Review and extensions. *Journal of Consumer Research*, *13*(4), 473–491. <https://doi.org/10.1086/209082>
- MacRae, C. N. (1992). A tale of two curries: Counterfactual thinking and accident-related judgments. *Personality and Social Psychology Bulletin*, *18*(1), 84–87. <https://doi.org/10.1177/0146167292181012>

- Mandel, D. R., & Dhimi, M. K. (2005). “What I did” versus “what I might have done”: Effect of factual versus counterfactual thinking on blame, guilt, and shame in prisoners. *Journal of Experimental Social Psychology*, 41(6), 627–635.
<https://doi.org/10.1016/j.jesp.2004.08.009>
- Markman, K. D., Klein, W. M. P., & Suhr, J. A. (2012). *Handbook of Imagination and Mental Simulation*. Psychology Press.
- Markman, K. D., Mizoguchi, N., & McMullen, M. N. (2008). “It would have been worse under Saddam.” Implications of counterfactual thinking for beliefs regarding the ethical treatment of prisoners of war. *Journal of Experimental Social Psychology*, 44(3), 650–654. <https://doi.org/10.1016/j.jesp.2007.03.005>
- McEnany, K. (2020, September 9). *White House Briefing*. <https://www.c-span.org/video/?475583-1/white-house-denies-president-downplayed-coronavirus-threat>
- Miller, D. T., Visser, P. S., & Staub, B. D. (2005). How surveillance begets perceptions of dishonesty: The case of the counterfactual sinner. *Journal of Personality and Social Psychology*, 89(2), 117.
- Morrison, W. M. (2018). *China-U.S. Trade Issues*.
<https://ecommons.cornell.edu/handle/1813/77813>
- Mosleh, M., Pennycook, G., & Rand, D. G. (2020). Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter. *PLOS ONE*, 15(2), e0228882. <https://doi.org/10.1371/journal.pone.0228882>
- Mueller, A. B., & Skitka, L. J. (2018). Liars, damned liars, and zealots: The effect of moral mandates on transgressive advocacy acceptance. *Social Psychological and Personality Science*, 9(6), 711–718. <https://doi.org/10.1177/1948550617720272>

- Murphy, G., Loftus, E. F., Grady, R. H., Levine, L. J., & Greene, C. M. (2019). False memories for fake news during Ireland's abortion referendum. *Psychological Science*, 30(10), 1449–1459. <https://doi.org/10.1177/0956797619864887>
- Neimand, A., & Griffin, L. (2017, January 20). *Why each side of the partisan divide thinks the other is living in an alternate reality*. The Conversation. <http://theconversation.com/why-each-side-of-the-partisan-divide-thinks-the-other-is-living-in-an-alternate-reality-71458>
- Newman, Eryn. J., Jalbert, M. C., Schwarz, N., & Ly, D. P. (2020). Truthiness, the illusory truth effect, and the role of need for cognition. *Consciousness and Cognition*, 78, 102866. <https://doi.org/10.1016/j.concog.2019.102866>
- Ohtsubo, Y., Masuda, F., Watanabe, E., & Masuchi, A. (2010). Dishonesty invites costly third-party punishment. *Evolution and Human Behavior*, 31(4), 259–264. <https://doi.org/10.1016/j.evolhumbehav.2009.12.007>
- Parkinson, M., & Byrne, R. M. J. (2017). Counterfactual and semi-factual thoughts in moral judgements about failed attempts to harm. *Thinking & Reasoning*, 23(4), 409–448. <https://doi.org/10.1080/13546783.2017.1345790>
- Parloff, R. (2014, April 12). Theranos' Elizabeth Holmes: Young entrepreneurs need “a mission.” *Fortune*. <https://fortune.com/2014/12/03/theranos-elizabeth-holmes-young-fortune-mpw-next-gen/>
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>

- Petrocelli, J. V., & Crysel, L. C. (2009). Counterfactual thinking and confidence in blackjack: A test of the counterfactual inflation hypothesis. *Journal of Experimental Social Psychology, 45*(6), 1312–1315. <https://doi.org/10.1016/j.jesp.2009.08.004>
- Petrocelli, J. V., Percy, E. J., Sherman, S. J., & Tormala, Z. L. (2011). Counterfactual potency. *Journal of Personality and Social Psychology, 100*(1), 30–46.
- Petrocelli, J. V., Seta, C. E., & Seta, J. J. (2012). Prefactual potency: The perceived likelihood of alternatives to anticipated realities. *Personality and Social Psychology Bulletin, 38*(11), 1467–1479. <https://doi.org/10.1177/0146167212453747>
- Pohl, R., Eisenhauer, M., & Hardt, O. (2003). SARA: A cognitive process model to simulate the anchoring effect and hindsight bias. *Memory, 11*(4–5), 337–356. <https://doi.org/10.1080/09658210244000487>
- Reyna, & Brainerd, C. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making, 4*(4), 249–262. <https://doi.org/10.1002/bdm.3960040403>
- Reyna, & Brainerd, C. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1–75. [https://doi.org/10.1016/1041-6080\(95\)90031-4](https://doi.org/10.1016/1041-6080(95)90031-4)
- Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13*(2), 60–66. <https://doi.org/10.1111/j.0963-7214.2004.00275.x>
- Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-trace theory. *Judgment and Decision Making, 7*(3), 332–359.

- Reyna, V. F. (2020). A scientific theory of gist communication and misinformation resistance, with implications for health, education, and policy. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1912441117>
- Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, 121(1), 133–148. <https://doi.org/10.1037/0033-2909.121.1.133>
- Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, 7(5), 411–426. <https://doi.org/10.1177/1745691612454303>
- 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>
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371. <https://doi.org/10.1111/j.1751-9004.2011.00355.x>
- Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, 116(16), 7662–7669. <https://doi.org/10.1073/pnas.1805871115>
- Scholl, A., & Sassenberg, K. (2015). Better know when (not) to think twice: How social power impacts prefactual thought. *Personality and Social Psychology Bulletin*, 41(2), 159–170. <https://doi.org/10.1177/0146167214559720>
- 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>

- Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty*, 25(2), 185–201. <https://doi.org/10.1023/A:1020647814263>
- Shalvi, S., Dana, J., Handgraaf, M. J. J., & De Dreu, C. K. W. (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., & Leiser, D. (2013). Moral firmness. *Journal of Economic Behavior & Organization*, 93, 400–407. <https://doi.org/10.1016/j.jebo.2013.03.014>
- Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin*, 11(1), 118–127. <https://doi.org/10.1177/0146167285111011>
- Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology*, 17(2), 142–158. [https://doi.org/10.1016/0022-1031\(81\)90011-1](https://doi.org/10.1016/0022-1031(81)90011-1)
- Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology*, 102(6), 1164–1177. <https://doi.org/10.1037/a0028381>
- 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>

- Smallman, R. (2013). It's what's inside that counts: The role of counterfactual content in intention formation. *Journal of Experimental Social Psychology*, 49(5), 842–851. <https://doi.org/10.1016/j.jesp.2013.05.004>
- Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36(11), 1202–1212. <https://doi.org/10.1037/0022-3514.36.11.1202>
- St. Augustine. (421). *Enchiridion on faith, hope, and love*. Regnery Publishing.
- Stahl, C., & Klauer, K. C. (2008). A simplified conjoint recognition paradigm for the measurement of gist and verbatim memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(3), 570–586. <https://doi.org/10.1037/0278-7393.34.3.570>
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437–446. <https://doi.org/10.1037/0022-3514.73.3.437>
- Tetlock, P. E. (1998). Close-call counterfactuals and belief-system defenses: I was not almost wrong but I was almost right. *Journal of Personality and Social Psychology*, 75(3), 639–652. <https://doi.org/10.1037/0022-3514.75.3.639>
- Tetlock, P. E., & Henik, E. (2007). Theory-versus imagination-driven thinking about historical counterfactuals. In D. R. Mandel, D. J. Hilton, & P. Catellani (Eds.), *The psychology of counterfactual thinking* (pp. 199–216). Routledge.
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals.

Journal of Personality and Social Psychology, 78(5), 853–870.

<https://doi.org/10.1037//0022-3514.78.5.853>

Tetlock, P. E., & Lebow, R. N. (2001). Poking counterfactual holes in covering laws: Cognitive styles and historical reasoning. *American Political Science Review*, 95(4), 829–843.

<https://doi.org/10.1017/S0003055400400043>

Tetlock, P. E., & Visser, P. S. (2000). Thinking about Russia: Plausible pasts and probable futures. *British Journal of Social Psychology*, 39(2), 173–196.

<https://doi.org/10.1348/014466600164417>

The Economist. (2016, October 11). Yes, I'd lie to you. *The Economist*.

<https://www.economist.com/briefing/2016/09/10/yes-id-lie-to-you>

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In *Social psychology: Handbook of basic principles* (pp. 239–270). The Guilford Press.

Trump, D. (2015, October 27). *Road to the White House 2016: Presidential Candidate Donald Trump Campaign Rally in Sioux City, Iowa*. <https://www.c-span.org/video/?329010-1/donald-trump-campaign-rally-sioux-city-iowa>

Trump, D. (2020, July 3). *Remarks by President Trump After Tour of the Centers for Disease Control and Prevention | Atlanta, GA* [Interview]. <https://www.whitehouse.gov/briefings-statements/remarks-president-trump-tour-centers-disease-control-prevention-atlanta-ga/>

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232. [https://doi.org/10.1016/0010-0285\(73\)90033-9](https://doi.org/10.1016/0010-0285(73)90033-9)

- 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>
- Uhlmann, E. L., Pizarro, D. A., Tannenbaum, D., & Ditto, P. H. (2009). The motivated use of moral principles. *Judgment and Decision Making*, 4(6), 479–491.
- Westfall, J. (2016). *PANGEA: Power ANALysis for GEneral Anova designs*. <http://jakewestfall.org/publications/pangea.pdf>
- Wimmer, H., Gruber, S., & Perner, J. (1984). Young children's conception of lying: Lexical realism—moral subjectivism. *Journal of Experimental Child Psychology*, 37(1), 1–30. [https://doi.org/10.1016/0022-0965\(84\)90055-9](https://doi.org/10.1016/0022-0965(84)90055-9)
- Wong, E. M., Galinsky, A. D., & Kray, L. J. (2008). The Counterfactual Mind-Set: A Decade of Research. In *Handbook of Imagination and Mental Simulation*. Psychology Press.
- Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>

Appendix A

Stimuli in Study 1

Fact	Prefactual	Falsehood
Listerine antiseptic is NOT as effective as floss at reducing gingivitis.	If Listerine completes its product development plans, then Listerine antiseptic will become as effective as floss at reducing gingivitis.	Listerine antiseptic is clinically proven to be as effective as floss at reducing gingivitis.
Danactive dairy drinks do NOT reduce the risk of catching the common cold.	If Danone develops the probiotics in its dairy products, then Danactive dairy drinks will be able to reduce the risk of catching the common cold.	Danactive dairy drinks reduce the risk of catching the common cold.
Photos sent via Snapchat do NOT disappear forever.	If Snapchat develops its software, then they will ensure that photos sent via Snapchat disappear forever.	Photos sent via Snapchat disappear forever.
Gerber's good start formula does NOT prevent allergies in children.	If Gerber develops its good starter formula, then it will prevent allergies in children.	Gerber's good starter formula prevents allergies in children.
Crystal wash, an environmentally friendly laundry detergent substitute, is NOT as effective as laundry detergent for cleaning clothes.	If Crystal wash develops more environmentally-friendly bacteria killers, then it will be just as effective as laundry detergent.	Crystal wash is just as effective in cleaning clothes as laundry detergent.
Luminosity games do NOT enhance users' performance at school.	If Lumos Labs works with scientists on their product development, then Luminosity will enhance users' performance at school.	Luminosity games enhance users' performance at school.
Taking Airborne dietary supplements does NOT boost your immune system.	If Airborne dietary supplements incorporate additional vitamins into their formula, then it will be able to boost the immune system.	Taking Airborne dietary supplements boosts your immune system.
Volkswagen 3.0 liter TDI diesel cars do NOT meet emission standards.	If Volkswagen works with its engineers to create more fuel efficient engines, then its 3.0 liter TDI diesel cars will meet emission standards.	Volkswagen 3.0 liter TDI diesel cars meet all emission standards.

Appendix B

Stimuli in Studies 3, 4, and 5

Study	Fact	Relevant prefactual (Studies 3 and 4)	Irrelevant prefactual (Study 4)	Relevant prefactual (Study 5)	Irrelevant prediction (Study 5)	Falsehood	Political fit
3, 4, 5	It's a proven fact that there have been just four documented cases of people voting illegally in the 2016 presidential election.	If the United States does not tighten its border security, then millions of people will vote illegally in the upcoming presidential election.	If the United States does not tighten its border security, then millions of Americans will be out of a job before the next presidential election.	Millions of people will vote illegally in the upcoming presidential election, if...	The new United States-Mexico-Canada Agreement (USMCA), which gives the U.S. greater access to Canadian dairy and allows extra imports of Canadian cars will lead to...	Millions of people voted illegally in the last presidential election.	Rep
3, 4, 5	It's a proven fact that the United States' trade deficit with China was \$336 billion last year.	If Democrats block the imposition of tariffs on Chinese goods, then the U.S. trade deficit with China will grow to \$500 billion next year.	If Democrats block the imposition of tariffs on Chinese goods, then the U.S. will have more difficulty negotiating with China about North Korea next year.	The U.S. trade deficit with China will grow to \$500 billion next year, if...	The Jobs for Our Heroes Act, which makes it easier for veterans to apply for commercial driver's licenses will lead to...	Last year, the trade deficit with China was \$500 billion.	Rep

3	It's a proven fact that heart disease is the leading cause of death for black Americans.	If the Trump administration's proposed ban on abortion after 20 weeks of pregnancy is not passed, then abortion will become the leading cause of death for black Americans.				Abortion is the leading cause of death for black Americans.	Rep
4, 5	It's a proven fact that the company US Steel is NOT currently building any new steel mills.	If the Trump administration imposes its steel tariffs on more foreign countries, then US Steel will open six new steel mills.	If the Trump administration imposes its steel tariffs on more foreign countries, then other countries will remove the United States from the United Nations Economic and Social Council.	US Steel will build six new steel mills, if...	Current tax cuts to the top 20% of earners in the U.S. will lead to...	The company US Steel is currently building six new steel mills.	Rep
3, 4, 5	It's a proven fact that in 2017, the average top CEO made 265 times more money than the average American worker.	If the Trump administration keeps making pro-corporate decisions, then the average top CEO will soon make 500 times more than the average worker.	If the Trump administration keeps making pro-corporate decisions, then the number of small businesses in the United States will decrease.	The average top CEO will soon make 500 times more money than the average American worker, if...	Current U.S. economic sanctions on North Korea will lead to...	The average top CEO currently makes 500 times more money than the average American worker.	Dem

3, 4, 5	It's a proven fact that 13% of legal gun owners in the U.S. have purchased firearms without a background check.	If the Trump administration continues to support pro-gun policies, then over 25% of legal gun purchases in the U.S. will be made without background checks.	If the Trump administration continues to support pro-gun policies, then more Americans will begin to purchase guns.	Over 25% of legal gun owners in the U.S. will purchase firearms without a background check, if...	The new Consumer Protection Act that allows consumers to freeze and unfreeze their credit file on short notice will lead to...	Over 25% of legal gun purchases in the U.S. are currently made without background checks.	Dem
3	It's a proven fact that the number of HIV cases related to heroin use declined between 2010 and 2015.		If Donald Trump continues to cut funding from the Substance Abuse and Mental Health Services Administration, then heroin-related HIV rates will rise.			Heroin-related HIV rates rose between 2010 and 2015.	Dem
4, 5	It's a proven fact that white Americans are 10% more likely to be approved for mortgages than black or Hispanic Americans with the same qualifications.	If Republicans abandon U.S. affirmative action policies, then white Americans will be 300% more likely to be approved for mortgages than black or Hispanic applicants with the same credentials.	If Republicans abandon U.S. affirmative action policies, then the percentage of American college students who are black will fall below 10%.	White Americans will be 300% more likely to be approved for mortgages than black or Hispanic applicants with the same credentials, if...	The legalization of marijuana in California will lead to...	White Americans are 300% more likely to be approved for mortgages than black or Hispanic applicants with the same credentials.	Dem

Appendix C

Stimuli in Study 6

Fact	Relevant prefactual	Irrelevant prefactual	Falsehood	Gist	Political fit
On average, LESS than 500 people die from gun violence each day in the United States.	If Republicans loosen gun protection laws, then every day 500 people might die from gun violence.	If Republicans loosen gun protection laws, then the NRA might throw a celebration at its annual convention.	Every day 500 people die from gun violence in the United States.	Many lives are lost to gun violence in the United States.	Dem
The leading cause of death for young Black men and women in the U.S. is heart disease, NOT law enforcement violence.	If there isn't major police reform, then law enforcement violence might become the leading cause of death for young Black men and women in the US.	If there isn't major police reform, then police force recruitment might increase.	Law enforcement violence is the leading cause of death for young Black men and women in the US.	Law enforcement harms young Black people in America.	Dem
There is evidence that people of lower socio-economic class are more likely to abuse drugs than other socio-economic groups.	If poor people become properly supported by public policies, then poor people might not abuse drugs any more frequently than any other socio-economic groups.	If poor people are properly supported by public policies, then incarceration might decrease.	Poor people do not abuse drugs any more frequently than other socio-economic groups.	Poor people aren't inclined to be drug addicts.	Dem

There are large racial gaps in vaccination rates across the entire population, including those 65 and above.	If the government prioritizes vaccine equality amongst seniors, then there might be no difference between White, Black, Hispanic, and Asian American vaccine rates among Americans 65 and older.	If the government prioritizes vaccine equality amongst seniors, then many COVID-19 deaths might be prevented.	There is no difference between White, Black, Hispanic, and Asian American vaccine rates among Americans 65 and older.	The Biden Administration is achieving vaccination equality.	Dem
There has been no change in human trafficking at the Mexico Border in the time since Biden became president.	If the Biden administration defunds homeland security, then human trafficking at the Mexico border might quadruple in the time since Joe Biden became President.	If the Biden administration defunds homeland security, then extra funds might be spent on public infrastructure.	Human trafficking at the Mexico border has quadrupled in the time since Joe Biden became president.	The Biden administration enables illegal activity at the U.S.-Mexico border.	Rep
Joe Biden has NOT halted all deportations.	If liberals pressure Joe Biden, then Biden might halt all deportations.	If liberals pressure Joe Biden, then Biden might repay all students loans.	Joe Biden has halted all deportations.	Joe Biden is lenient on illegal immigrants.	Rep
The Potomac River has gotten CLEANER over the years.	If problems with immigration aren't resolved, then the Potomac River in Washington, D.C., might get dirtier from litter that is exclusively left by immigrants.	If problems with immigration aren't resolved, then we might see more fighting between politicians.	The Potomac River in Washington, D.C., has gotten dirtier from litter that is exclusively left by immigrants.	Immigrants are polluting American land.	Rep

None of the \$600 million COVID relief package given to San Francisco is being used to give alcohol and marijuana to the homeless.	If California liberals decide how the COVID economic relief package is spent, then part of it the \$600 million COVID relief package given to San Francisco will be used to give alcohol and marijuana to the homeless.	If California liberals decide how the COVID economic relief package is spent, then small business owners in California might receive a lot of support.	Part of the \$600 million COVID relief package given to San Francisco is being used to give alcohol and marijuana to the homeless.	The government is weak on drugs.	Rep
--	---	--	--	----------------------------------	-----
