

UNITED WISDOM

May We Lie to Protect Someone?

Freedom and Action

Andrey Zarinsh & Maria Evilkut

A-15-N01 · EN

@andrey.zarinsh @ilkutki

May We Lie to Protect Someone?

Epigraph

> *not every one who says a false thing lies*

Augustine, On Lying §3, H. Browne translation in the New Advent edition; a clause excerpt. The following explanation concerns someone who believes the statement true: distinguishing error from lying does not permit deliberate deception.

What this article is about

At a drawing class, participants display their work anonymously. One drawing attracts laughter. Lena knows its author: a friend attending for the first time who asked her not to identify him yet. Another participant asks whether it is his work. He is already preparing to forward a photograph with a mocking caption. Lena wants to say no. A small lie seems capable of protecting someone from undeserved humiliation, while truthfulness seems about to make her complicit in cruelty.

The hypothetical scene contains several questions. Does the questioner need the fact for a permissible decision? Can Lena decline to answer? Whom would her denial affect? How does protecting her friend differ from concealing his responsibility? We will neither treat every good purpose as permission to lie nor every disclosure as a moral duty. Deliberate deception has its own grounds, costs and limits. Philosophical and religious texts retain their disagreements; our practice remains a way to clarify an ordinary conversation between adults, rather than a procedure for danger to life or professional advice.

I. Identify what we are actually about to say

Lena might be mistaken about authorship, honestly express a supposition or knowingly provide an incorrect name. The outward sentence can be the same while the speaker's relationship to knowledge differs. If she genuinely does not know, saying so is not a trick. If she knows and says otherwise to create a false impression, the phrase cannot establish an absence of deception. Protection does not begin by renaming our action to make it easier on our conscience. First distinguish uncertainty from intentionally misleading somebody.

In On Lying §3, Augustine distinguishes a statement's falsehood from the speaker's inward position. Someone may express a mistaken judgement while believing it true; the absence of a lie does not automatically remove imprudence. That qualification matters. It helps our scene distinguish an accidental inaccuracy from a conscious denial. We neither transfer Augustine's

@andrey.zarinsh @ilkutki

entire theological account into a contemporary conversation check nor promise that a sincerely convinced person never answers for harm caused by their speech.

We must also name the intended outcome. Protecting a friend is too broad: stopping a mocking publication, preserving requested anonymity, avoiding personal conflict or preventing a conversation about a weak drawing? These purposes may coexist without being equivalent. If Lena fears being considered difficult, acknowledging that motive is useful too. It does not cancel care, but should not quietly become the only reason. Otherwise every action allowing her to leave without awkward questions or subsequent responsibility will acquire the flattering name of protecting her friend.

Finally, describe the possible harm without exaggeration. The scene involves an intention to humiliate and circulate mockery; this differs from a teacher properly asking who made the drawing in order to provide feedback. A painful comment is not always an unfair attack, and the friend's reluctance to hear criticism does not permit attributing his work to somebody else. We are considering a particular use of information. Different people may need the same information for different purposes, and those purposes matter when assessing the answer.

II. Truthfulness does not require answering every question

Inside the dilemma of lying or exposing someone, refusal often disappears. Lena can say that the author requested anonymity, that she will not help circulate mockery, and that the drawing can be discussed without a name. This does not guarantee that the questioner stops guessing. It may irritate him. But refusal preserves the distinction between a boundary on access and a false report. We need not establish that everybody will like a refusal before recognising it as an available option. Its moral significance is not measured by the speaker's convenience in the next minute.

Being able to ask in ordinary conversation does not establish a ground for demanding an answer. We are not determining legal rights or professional obligations here. We are noticing a familiar mistake: insistence alone does not make information common property. We can speak truthfully about our position without disclosing a friend's personal information. Yet a refusal should mean a boundary rather than function as a planned hint confirming the name. A convenient formulation still requires attention to the impression it intentionally creates in this conversation.

Refusal sometimes fails to prevent harm, and that cannot be concealed. The questioner may already know the author or take silence as confirmation. The choice then concerns a more difficult relationship between truthfulness and protection. First, however, check whether a boundary has genuinely been exhausted or is merely unpleasant. Lies are often chosen earlier because they seem socially smooth: nobody notices an objection or asks for an explanation.

This benefit to the speaker can coexist with care, without proving that deception was necessary.

Lena need not perform neutrality if she wants to change how the drawing is treated. She can object directly to mockery and propose discussing quality without public humiliation. Protection then addresses the harmful action rather than merely hiding a fact. Declining to provide a name and declining to circulate a photograph are different steps. The first preserves an entrusted condition; the second defines her participation. Together they offer more clarity than an invented account in which authorship supposedly disappears and nobody needs to answer for anything afterwards.

III. A good intention does not settle the question of deception

Suppose Lena nevertheless says that another person, who left the group long ago, made the drawing. She has prevented direct mockery of her friend but transferred it to someone absent. The protective purpose now operates through another person's reputational cost. Their absence makes harm less visible without eliminating it. A good motive does not explain why that person may become the material of protection. Assessment must therefore include those used by a false account, as well as the person for whose sake it was told.

Another cost affects the protected person. Her friend requested anonymity, not an invented biography or control of every conversation. After the denial, Lena might require him to sustain her account, leave the drawing unsigned and avoid the teacher. Protection becomes management. Even if the first denial seems justified, it creates no authority to arrange subsequent decisions for somebody else. Consent to help has a subject and a limit; it cannot retrospectively expand into an obligation to join whatever story the helper considers necessary to save face.

In Groundwork section I, Kant considers a lying promise made from a difficult position and distinguishes moral grounds from calculation of future advantage. This is a particular example, not a treatise on every protective lie which we claim to have read in detail. In a limited comparison, it prevents inconvenience from becoming sufficient justification: selecting a useful outcome does not automatically establish a permissible method. Nor does the example require telling every questioner every fact known. A lying promise and declining disclosure remain different actions.

Our intention needs examination after the first minute too. Lena may sincerely protect her friend while also preserving an image of herself as a good friend. If she begins concealing the lie to preserve that image, the original reason no longer explains the continuation. Ask which specific harm a new invention now prevents. If the answer is only discovery that she lied, the ground has changed. This is not a verdict on her personality, but a reason to stop extending

the story and seek correction that preserves, where possible, trust, personal boundaries and responsibility to affected people.

IV. Trust matters to an ethics of consequences too

Attention to consequences is sometimes presented as allowing any brief deception with a good outcome. Mill writes differently in Utilitarianism Chapter II: veracity supports reliance on speech, and violating it for immediate convenience can undermine a much more significant good. His argument considers consequences of a rule and a disposition, rather than only the success of one reply. We preserve that level of reasoning without assigning numerical prices to trust in each conversation. Nobody has a reliable table showing that a little lie always produces more assistance than loss.

In the same context, Mill admits possible exceptions to prevent great, undeserved evil where withholding the fact can only be achieved through denial, and calls for exceptions to remain limited by necessity. These conditions should not quietly become everyday awkwardness. Fear of criticism of a drawing and prevention of serious harm are different grounds. We do not turn the historical examples into medical or crisis recommendations. For the ordinary scene, the point is that even an approach concerned with consequences needs an explanation of an exception and its boundary, rather than a kindly label or the speaker's confident forecast.

Prediction in the scene is incomplete. A denial may stop one conversation while leaving the photograph with the questioner. Later he might learn the author's identity, add an accusation of deception and circulate the story more widely. This is possible, not guaranteed. It does not prove that truthful refusal is always safer. But deception cannot be called reliable protection without checking which route to harm it blocks. If the expected benefit rests on certainty that everybody believes us and nobody checks, we have a weak prediction, not an established advantage.

Trust also has several recipients. A questioner can use truthfulness cruelly; this does not require helping him. A friend can entrust Lena with an identity; she need not endorse all his decisions. A teacher may depend on honest authorship in a different, agreed context. Trust is not one reservoir either wholly protected or wholly destroyed. Name the relationship and its subject: who was promised what, which information is needed for which decision, and where refusal preserves honesty without serving another person's intention to humiliate.

V. Protection from humiliation differs from concealment of responsibility

Change the scene. The drawing uses somebody else's work without acknowledgement, and the teacher is investigating to avoid attributing authorship incorrectly. Now the request has a different ground. Lena can still object to public humiliation and propose a calm discussion. But denying a known fact to spare her friend every consequence protects him from responsibility for his own action. Care need not take that form. We can support a person through correction without becoming a source of false accusation, concealment or another person's loss of recognition.

The distinction is not always obvious because an error's consequences can themselves feel humiliating. A person may feel ashamed even after a careful comment. The feeling matters, without establishing that every report of a fact causes injustice. Examine grounds, form, recipient and necessary extent. If discussing the work privately is sufficient, announcing the name before everyone is excessive. If another creator has lost recognition, avoiding public mockery does not mean simply closing the question. Duties of correction and boundaries on how a person is treated must be considered together.

Closeness can make assessment one-sided. Lena more easily imagines her friend's pain than the position of someone whose authorship was taken. Reversing roles is useful without declaring everybody's feelings identical. What would she regard as protection if her work had been taken? What would she regard as unnecessary humiliation if she had made the mistake? Changing perspective provides no automatic answer, but shows whether a rule operates only for someone close. Partiality is not always wrong; it should not conceal another person's justified need for a truthful decision.

Protection from cruel treatment remains possible while acknowledging an error. Lena can decline to forward a mocking photograph, help her friend formulate a correction and support a conversation with the teacher. None of this requires denying authorship or declaring the mistake nonexistent. Such help may be harder than a short denial because it does not eliminate consequences for the friend. It preserves his participation instead. He can answer, clarify and correct what is his, while the helper does not make care a promise to cancel unwelcome facts forever or control other people's reactions.

VI. Do not conceal disagreements between sources

Augustine concludes On Lying in §42 against every lie; his grounds involve Christian truth and fidelity. In the selected context, Mill allows limited exceptions under particular conditions of harm and necessity. We do not make these positions secretly identical or claim that both permit compassionate deception. Augustine's distinction between error and lying does not cancel his conclusion. Mill's recognition of veracity does not remove his discussion of exception. These are real differences in grounds which should remain visible to the reader.

Kant's lying-promise example cannot become a shortened proof of every conclusion about conversation either. We read specific paragraphs distinguishing duty and prudence and examining the maxim's universality. We do not claim that this means we studied all Kant's writings on lying, or attribute our practice of refusal to him. The boundary of reading is part of accuracy, rather than a deficiency to hide behind a philosopher's confident name. His example raises a question about method while leaving separate investigation of more difficult cases open.

In Analects IV.22, restraint in speech is connected with concern that actions will not match words. This is a historical context of correspondence between conduct and utterance, not an independent answer concerning protective lying. Our comparison shows that a promise to protect someone needs boundaries too. Lena cannot guarantee that nobody learns the name or ever laughs. She can specify her participation: not passing on an entrusted identity in a mocking conversation and not circulating a photograph with a humiliating caption.

Our editorial position seeks truthful protection before disagreement becomes permission for an exception. Refusal, limiting disclosure, objecting to treatment and involving the protected person deserve separate examination. This does not prove that every conflict can always be resolved without deception. Nor does it direct action in a dangerous situation requiring different circumstances and competencies. We keep the difficult question open at its proper scale. In ordinary conversation, we seek actions explainable without invention and without handing all available information to someone prepared to humiliate.

VII. If the lie has already been told

Lena has already said no and now regrets it. The next question is not how to sustain the denial more convincingly, but what needs correction. If she named somebody else, she needs to stop transferring the accusation to them. If she simply denied her friend's involvement, correction need not mean immediately revealing his identity to the same mocking group. She can acknowledge that her answer was incorrect and that she will not discuss authorship in such a conversation. This is no magic formula; its adequacy depends on the misunderstanding and harm she created.

Her friend needs to know about a decision affecting him. Otherwise Lena alone determines the price of maintaining his anonymity. The conversation can be short: she answered incorrectly to stop mockery; it was her decision; she will not ask him to sustain an invention. They can then discuss whether he wants to identify himself to the teacher or group. Lena should not transfer her lie into his obligation to correct it. But she cannot treat him as uninvolved in something already affecting his participation and subsequent speech.

Acknowledgement does not require a long public confession. A large public account may spread personal information afresh and give someone attention they did not request. Correction should correspond to the misunderstanding and recipients affected. If nobody was accused and no significant decision relied on the fact, a different response may suffice than when another person's authorship was denied. We derive no universal disclosure procedure from the scene. We retain two requirements: do not preserve a useful lie for your image, and do not cause new harm merely for a feeling of complete purification.

After correction, reconsider the promise of help. An agreement might be that Lena will not disclose his name or join mockery in such conversations; when a genuine authorship question arises, she will clarify the situation and invite his own participation. This is a definite commitment instead of a boundless promise to arrange everything. He may disagree, requiring further discussion. Help remains joint action, rather than secret management of someone else's information. An honest limit can protect more than a promise deliverable only through an ever-growing story.

VIII. Keep difficulty distinct from permission

Protective lying is difficult precisely because truthfulness and care both matter. Acknowledging difficulty does not remove every boundary. We can lack an answer to an extreme case and still see that refusal is possible today and an invented name transfers harm to an uninvolved person. A decision need not await completion of moral philosophy's history. It should remain proportionate to facts actually known. If a situation exceeds ordinary conversation, a limited exercise does not thereby acquire new competence or become a guarantee of safety.

Several signs suggest protection is becoming self-justification: harm is vaguely named, alternatives are declared impossible without checking, other people's costs are omitted and the exception has no end. These signs are not a diagnostic test or proof of a bad motive. They indicate unasked questions. Lena can return to each: what am I preventing, why is a boundary insufficient, who bears the cost, and when does necessity end? Answers may leave disagreement, but make it more specific than invoking a little lie for a great good.

Moral discussion should not make the protected person a silent occasion for the helper's heroism. The friend has his own participation, including opportunities to withstand criticism, identify the work, ask for support or leave the group. Lena can help without choosing his response beforehand. If she protects him only while he confirms an image of someone fragile and grateful, care becomes a conditional role. Clarifying his wishes does not remove all her obligations, but returns the relationship to two people acting rather than one managing another's life with the best intentions.

In the original scene, Lena chooses not to name the author and not to join a mocking publication. She says this directly, without false denial. This outcome is not declared a universal victory over protective lying. It shows an available response to a particular conversation: decline a use of information while being honest about one's position and not shifting mockery to someone else. If the question changes, reassess the decision. Protection remains attention to a real person and a way of acting, rather than permission to say anything promising immediate relief.

What this means in practice

Choose an ordinary conversation where you want to say something knowingly false for someone else's sake. Do not apply the exercise to threats, violence, medical decisions or professional duties. First record what you know and what you assume. Name the specific harm you want to prevent. Distinguish unfair treatment from consequences of a real error: both can hurt, while requiring different help. If your own awkwardness is the principal motive, record it beside care without substituting it for the other person's circumstances.

Check three alternatives: declining an answer, truthfully limiting disclosure and directly objecting to harmful use of information. Do not call an option impossible merely because it is uncomfortable. If you expect refusal to reveal the fact, name the grounds and limits of that expectation. Examine the proposed lie's costs to uninvolved people, trust and the protected person's freedom. In particular, do not assign them an obligation to maintain an invention they did not choose or transfer unwanted attention to another creator.

Formulate one honest proposal defining your participation. For instance, decline to identify authors in a conversation intended to ridicule them and decline to forward the photograph with that caption. This is an example for a hypothetical scene, not a universal script. If the protected person can participate, clarify their wishes separately without promising control over everyone else. Name a limit of help you are prepared to maintain.

If an incorrect answer has already been given, identify the misunderstanding and its recipients. Correction should not enlarge the account to preserve your image or reveal unnecessary private information for relief. Afterwards check an available result: have you stopped your own

circulation and left nobody under a false accusation? This author-designed practice has no claimed research-established efficacy, determines no legal permissibility and guarantees no outcome. It helps distinguish protecting a person from conveniently justifying the selected method.

Key quotation

> *we must never at all tell a lie*

Augustine, On Lying §42, H. Browne translation in the New Advent edition; a clause from the concluding argument. This is the position of this Christian text, not a declared consensus among all sources here.

Connections between traditions

Source	Context	What it helps reveal	Difference preserved
Augustine §§3,42	Lying, inward position and Christian fidelity to truth	Error differs from deliberate deception	The conclusion is not converted into permission for a kindly lie
Kant, Groundwork I	A lying promise made from difficulty	Advantage and permissible method are different questions	The example does not establish reading of every Kantian text on lying
Mill, Utilitarianism II	Reliance on speech and the boundary of a possible exception	Immediate convenience does not equal general welfare	Possible exception remains a real disagreement with Augustine
Confucius IV.22	Actions corresponding to words	A promise of protection needs a practicable limit	The saying does not independently resolve protective lying
United Wisdom	Hypothetical drawing scene and editorial practice	Truthful refusal and costs to third parties	The editorial procedure is not attributed to primary sources

@andrey.zarinsh @ilkutki

Sources

• Augustine. On Lying, New Advent. Selected §§1–4 and §§40–43 were read, not the entire treatise. H. Browne translation, edited in the 1887 Nicene and Post-Nicene Fathers by Philip Schaff; revised for New Advent by Kevin Knight. The epigraph comes from §3 and key quotation from §42. The theological grounds and final prohibition are preserved. Worldwide rights in the edition are not declared established.

• Immanuel Kant. Fundamental Principles of the Metaphysic of Morals, Thomas Kingsmill Abbott translation, Project Gutenberg 5682. The lying-promise context in section I and paragraphs concerning duty and prudence were read. This does not replace separate investigation of Kant's other writings on lying.

• John Stuart Mill. Utilitarianism, Project Gutenberg 11224. Chapter II's selected context concerning truthfulness, trust, immediate advantage and restrictions on exception was read. Historical examples are not used as medical or crisis instructions.

• Confucius. The Analects, James Legge translation, Project Gutenberg 3330. IV.15–26 context was read; IV.22 is compared in a limited way with a practicable promise of help. No literal quotation from this source appears here.

Tags

truth, lying, deception, error, protection, trust, intention, harm, humiliation, refusal, information, anonymity, care, responsibility, authorship, consequences, exception, necessity, consent, boundaries, correction, promise, closeness, third parties, Augustine, Kant, Mill, Confucius

About the authors

Andrey Zarinsh is the author and founder of United Wisdom. He explores inner transformation, the unity of traditions, intercultural dialogue, and contemporary spiritual discipline. Instagram: @andrey.zarinsh.

Maria Evilkut is a co-author of United Wisdom. Alongside her research, she seeks to test these ideas in practice and observe how principles of unity and mutual understanding work in real life between people and cultures. Instagram: @ilkutki.