

Immoral lies and partial beliefs

Neri Marsili

*Logos – Logic, Language and Cognition Research Group,
University of Barcelona*

Abstract: In a recent article, Krauss (2017) raises some fundamental questions concerning (i) what the desiderata of a definition of lying are, and (ii) how definitions of lying can account for partial beliefs. This paper aims to provide an adequate answer to both questions. Regarding (i), it shows that there can be a tension between two desiderata for a definition of lying: ‘descriptive accuracy’ (meeting intuitions about our ordinary concept of lying), and ‘moral import’ (meeting intuitions about what is wrong with lying), vindicating the primacy of the former desideratum. Regarding (ii), it shows that Krauss’ proposed ‘worse-off requirement’ meets neither of these desiderata, whereas the ‘comparative insincerity condition’ (Marsili 2014) can meet both. The conclusion is that lies are assertions that the speaker takes to be more likely to be false than true, and their distinctive blameworthiness is a function of the extent to which they violate a sincerity norm.

Keywords: Lying, Credences, Moral Concepts, Definitions, Norms

Introduction: Insincerity and Graded Beliefs

In the philosophical literature on the definition of lying, scholars agree that the following are *necessary conditions* for lying:

- A. The speaker asserts that p
- B. The speaker believes that p is false

Recently, Marsili (2014, 2018) and Krauss (2017) have challenged this orthodoxy. They both contend that the ‘belief requirement’ (B) ¶

is off track, because it is unable to capture *graded-belief lies*¹: lies involving a graded belief, *i.e.* a proposition that the speaker neither fully believes to be true, nor fully believes to be false. For example, suppose that Kermit has a credence of 0.25 in (1), so that he believes that it is *probably false* that there are chocolate cookies in the jar:

(1) There are chocolate cookies in the jar

If Kermit tells his friend Elmo that (1), Kermit is lying. But the 'belief requirement' (B) prevents the standard definition from counting (1) as a lie, because Kermit does not fully believe (1) to be false – he merely believes (1) to be more likely to be false than true. Marsili and Krauss agree that (B) needs to be revised to capture *graded-belief lies* like (1). Marsili's (2014:162, 2018:176) proposal is to expand (B) into the 'comparative insincerity condition', so as to capture any credence perceivably lower than 0.5:

The comparative insincerity condition

(CIC) The speaker takes herself to be more confident in the falsity of p than in its truth

CIC is able to capture graded-belief lies like (1). Nonetheless, Krauss (2017) has recently alleged that CIC is mistaken, because it fails to 'account for the damage lying does'. His alternative proposal is that an assertion is a lie only if it is expected to make the addressee 'epistemically worse-off'; that is, only if it satisfies the 'worse-off requirement':

The worse-off requirement

(W-O) The expected epistemic damage to the audience, with respect to p, by the speaker's lights, conditional on the audience trusting her with respect to p, at all, is greater than 0

¹ For Marsili (2016), however, there is a further reason to reject (B), namely that it fails to capture some insincere promises: those that you believe you will likely keep, despite your intention to violate them.

In simpler words, in asserting that p you lie only if you expect your audience's credence in p to become more inaccurate than the one they previously held (if they trust your assertion that p). Applied to our example, this means that (1) is a lie whenever Kermit asserts (1) and expects Elmo to update his credence in p to a value that is more distant from 0.25 (which is Kermit's standard of accuracy) than Elmo's current credence in p . For instance, Kermit is lying if he expects Elmo to revise his credence in (1) from 0.4 to 1, because 1 is further away from 0.25 than 0.4.

Krauss prefers W-O to CIC for one crucial reason: he takes the former, but not the latter, to be able to account for the 'distinctive blameworthiness' of lying. He shows this by means of a counterexample:

Imagine a case in which the speaker is just slightly more confident that p is true than false – say she has 0.51 credence in p . And, further, imagine that the speaker knows that her audience is agnostic – that the audience has 0.5 credence in the proposition. [...] Imagine that the audience moves from 0.5 credence to 0.8 credence. If the speaker's credence is 0.51, then the speaker will expect the audience to suffer epistemic damage. If the blameworthiness for lying is grounded in expected epistemic damage, then this speaker is blameworthy in exactly the same way as liars are, even if, according to both the orthodox position and Marsili's proposal, she hasn't lied.

The effectiveness of this counterexample is conditional on two assumptions: (i) that the primary desideratum of a definition of lying is to track the distinctive blameworthiness of lying, and (ii) that this blameworthiness is best characterised in terms of expected epistemic damage, as defined by the worse-off requirement. These assumptions are not uncontroversial, as they take a firm stance on two fundamental issues in the philosophy of lying: what the primary desiderata of a definition of lying are (for assumption (i)), and what makes lying blameworthy (for assumption (ii)). If these assumptions are correct, the ones underlying the current philosophical debate on lying must be wrong, as both (i) and (ii) are overwhelmingly held to be false. In this paper, I vindicate a 'traditional' stance on these issues, and show that, at closer inspection, both assumptions are indeed wrong. Against assumption (i), I argue that capturing the

distinctive blameworthiness of lying cannot be the only desideratum of a definition of lying. Against assumption (ii), I show that the worse-off requirement (W-O) fails to track the distinctive blameworthiness of lying, which is better characterised in terms of the violation of a sincerity norm.

1. The Desiderata of a Definition of Lying

I will start by dealing with the first assumption: that the primary desideratum of a definition of lying is to track the distinctive blameworthiness of lying. To avoid misunderstandings, let me stress that this assumption should not be confused with the view that all lies are, all things considered, morally wrong. Krauss is merely assuming that a good definition should track the fact that *lies are blameworthy in the same distinctive way* (e.g. in virtue of the epistemic risk that they impose on the audience). This still allows for single acts of lying to be 'all things considered' permissible, whenever sufficient countervailing considerations apply (Margolis 1963).

The idea that a definition should track the distinctive blameworthiness of lying is not entirely new. Some authors before Krauss have assumed that a negative moral evaluation is part of the meaning of the word 'lie' (Margolis 1963, cf. also Williams 1985, 140), so that in defining lying one should also explain what makes lying blameworthy (Grotius 1625, book III.I.XI.1)². This is nonetheless a minority view, and in contemporary scholarship the prevailing assumption is that the key desideratum of a definition of lying is its 'descriptive accuracy': a good definition should match our (morally neutral) intuitions about particular cases, capturing all and only those utterances that we call lies. The task of *defining what a lie is* and *explaining why* (ceteris paribus) *lying is wrong* are on this conception two separate tasks (Kemp and Sullivan 1993, Fallis 2009, Carson 2010:13, Mahon 2015). In assuming that a definition should first of all capture the distinctive blameworthiness of lying, Krauss is

² In his *De Jure Belli ac Pacis*, Grotius (1625, III.I.XI.1) declares that his intent is not to define lying in general, but the notion of a 'Lye strictly taken, as it is naturally *unlawful*' (that is, wrong or blameworthy). On his view, this 'strict' notion of lying necessarily involves the 'Violation of a real Right' (that is, an undefeated right to be told the truth).

thus departing from the methodological conventions in the literature, setting a different priority for his proposal: to explain what makes lying *prima facie* wrong.

Let us assume that Krauss has good motivations for introducing this new desideratum of, as we may call it, 'moral import'. The crucial problem with its introduction is that the two resulting desiderata (moral import and descriptive accuracy) can in principle come apart (*i.e.* impose inconsistent constraints on the definition); when they do, a morally interesting definition will not be descriptively accurate.

The *worse-off* criterion (W-O) proposed by Krauss can be helpful to illustrate how moral import and descriptive accuracy can come apart. Suppose that we share Krauss' intuition that the distinctive moral wrong in lying is that the liar imposes an epistemic threat on the dupe, and that W-O is successful in tracking this distinctive kind of blameworthiness. This would not yet guarantee that a definition incorporating W-O is descriptively accurate. To see this, consider the following counterexample (inspired by Benton 2018).

Suppose that (i) Kermit tells Elmo that (1) is false (he tells Elmo that there are *no* chocolate cookies in the jar), (ii) Kermit is maximally certain that (1) is true (Kermit is certain that there are in fact some chocolate cookies in the jar), and (iii) Kermit is aware that Elmo is already maximally certain that (1) is true (he knows that Elmo is already convinced that what he just said is true). While Kermit is clearly lying³, he is not attempting to modify Elmo's degree of belief in (1) – he is merely providing a further (testimonial) reason in support of that belief. Since the expected epistemic damage (from the speaker's perspective, and if the hearer trusts him) is not greater than 0, W-O fails to classify this as a lie. More generally, this sort of counterexamples (cf. Benton 2018 for further ones) shows that W-O is not descriptively accurate – it is too narrow, because it fails to provide a criterion that captures *all* lies.

³ Virtually every definition on the market converges on this prediction, including those that require an intention to deceive. Note that although Elmo already believes that *p* is false, Kermit is attempting to deceive Elmo on most standard accounts of deception, as he brings about new (testimonial) evidence for that belief, contributing causally to Elmo's continuing to have that inaccurate belief (cf. Chisholm and Feehan 1977, 144, Fuller 1976; Mahon 2007, 186-7, 189-90).

This seems to be a problem for Krauss's view, for he holds true (a) and (b), which seem to be incompatible with (c), that we have just proved to be true:

- (a) Lies are blameworthy in the same way
- (b) Lies are blameworthy only if [they meet the worse-off requirement]
- (c) There are some lies that [do not meet the worse-off requirement]

Since (a), (b) and (c) cannot all be true, one of these claims needs to be abandoned. If the desideratum of moral import needs to play some role in the definition of lying, we cannot abandon (a). The conclusion is that (b) must be abandoned.

This points out to a more general problem: that descriptive accuracy cannot be sacrificed at the expenses of moral import. The schema (a-b-c) generalises to any account of lying, once the content of the square brackets is replaced by an alternative requirement for the definition of lying. Note that claim (c) here represents the negation of the desideratum of descriptive accuracy, whereas condition (a) represents the desideratum of moral import. The schema shows that if descriptive accuracy is not met (that is, if (c) is true), moral import cannot be achieved (that is, (b) must be abandoned)– as long as we accept (a). Generalising, the schema shows that correctly characterising the moral import of lying is conditional on a definition's descriptive accuracy⁴ – whenever a definition fails to

⁴ Note, however, that the 'only if' in (b) only captures the necessity leg of the desideratum of moral import, and (c) only represents a challenge to the necessity leg of the definition. This because my main focus is W-O, and Krauss only claims that W-O is a necessary condition for lying (and for being blameworthy in the distinctive way that liars are). To generalize the schema so that it applies to full definitions, we should replace the 'only if' in (b) with a biconditional, and weaken (c) so as to allow for counterexamples to sufficiency.

capture all lies, it also fails to capture the feature that makes lies blameworthy in the same distinctive way⁵.

We are now in a position to determine what is wrong with assumptions (i) and (ii). The problem with (i), the view that the primary desideratum of a definition is tracking the 'moral import' of lying, is that meeting this desideratum is conditional on meeting 'descriptive accuracy'. As illustrated by the (a)-(b)-(c) scheme, any definition of lying that fails to meet descriptive accuracy also inevitably fails to capture the distinctive blameworthiness of this linguistic phenomenon. Against assumption (ii), W-O neither offers a good necessary condition for defining lies (since it is subject to counterexamples), nor it is able (for the same reason) to track their distinctive blameworthiness.

To be sure, to deny these two assumptions is not to say that it is impossible to identify a criterion that is able to meet both desiderata. Clearly, there could be an alternative way to capture a morally salient feature that all lies have (thereby meeting 'moral import'), and it may turn out that this feature is shared by all lies (thereby meeting 'descriptive accuracy'). But in order to determine whether such a criterion exists, we need to test it against both desiderata (pace to Krauss 2017:2,4,5). In what follows I argue that, at closer inspection,

⁵ One way to resist this conclusion is to argue that definitions should only aim to capture a morally uniform phenomenon, that can differ significantly to our ordinary concept of lying. The aim of this radical 'revisionist' project would be to identify a concept that has moral significance, lying*, that *may or may not* coincide with our *ordinary* concept of lying. If it does coincide, the definition *also* provides an analysis of our ordinary concept. But if it does not, the definition of lying* thus obtained would still be illuminating: for instance, it could be useful for moral theorising, or for law-making purposes. This alternative project is certainly coherent, and possibly worth pursuing. For my purposes, it is sufficient to point out that this project is so radically different from the one currently pursued by scholars working on the definition of lying that it cannot be regarded as a continuation of it. Authors engaging in the current debate explicitly aim to characterise the ordinary concept of lying (as it is employed by laypeople in their reasoning and talking about lying), as opposed to an artificial concept of lying* – within a common effort to engage in conceptual analysis, rather than conceptual engineering (cf. Fallis 2009, Mahon 2015). Lack of concern for ordinary use is thus incompatible with intervening within this pre-existing debate (which is, quite uncontroversially, what Krauss 2017 aims to be doing).

the ‘comparative insincerity condition’ CIC can help to meet this challenge: it identifies a concept that coincides with our ordinary conception of lying, and tracks the distinctive blameworthiness of this phenomenon.

2. Accuracy with Blameworthiness

2.1. *Rescuing the Comparative Insincerity Condition*

I will begin by reconsidering why CIC was rejected in the first place. Krauss’ criticism of CIC relied on the assumptions that we just rejected: that a definition of lying only needs to meet the desideratum of moral import, and that having W-O as a necessary condition is the only way to meet this desideratum. Now that we have established both these premises to be misguided, the counterexample on the ground of which Krauss rejects CIC loses its intuitive pull.

In the counterexample to CIC discussed in §1, a speaker asserts that p while having a credence of 0.51 in p . Unless we assume that a definition only has to track the distinctive blameworthiness of lying, and adopt the worse-off requirement as the criterion to identify this blameworthiness (but we have seen compelling arguments not to), there is no strong reason to think that this assertion is either a lie or not a lie. Intuitions about this sort of cases are not straightforward, and do not warrant a preference for either CIC or W-O. But suppose that we are forced to give a polar verdict, and to establish that either this is a lie or it is not. Intuitions may vary, but the fact that the speaker is more confident in the *truth* of p than in its *falsity* at least hints at the fact that this is *not* a lie. CIC is able to make sense of both these observations. W-O, by contrast, has no resources to acknowledge that the example is a borderline case; nor is it able to acknowledge, like CIC, that the more we move from these cases of uncertainty (around 0.5) to certainty (0 or 1), the sharper will be our intuitions about whether a given utterance is a lie (cf. Marsili 2018:176).

At any rate, if the descriptive accuracy of a given account is best measured against straightforward cases, this example bears limited argumentative weight. Crucially, when we consider straightforward

cases of lying, CIC makes the right verdict in every scenario in which other definitions fail. Unlike W-O, it avoids Benton's counterexamples (as discussed in Benton 2018:3, fn5); and unlike the standard 'belief requirement' (B), it is able to rule in graded-belief lies. In sum, CIC fares better than its rivals in terms of descriptive accuracy.

2.2. *Blameworthiness as Norm-Violation*

Is CIC also able to track the distinctive blameworthiness of lying? I think that this question can be answered positively. It can be argued that such blameworthiness resides in a feature identified by CIC: the (more or less severe) violation of a norm prescribing speakers to be sincere. Like the standard insincerity belief requirement (B), CIC captures a set of acts that have a morally salient feature in common – they all violate a sincerity norm. But improving on (B), CIC acknowledges the intuitive fact that a lie can be 'a more or less severe violation' of such a norm (Marsili 2014, 2018): one thing is to present as true a proposition that you are certain to be false, another to present as true a proposition that you merely believe more likely to be false than true. In other words, CIC represents an improvement on the tradition both in terms of 'descriptive accuracy' and in terms of 'moral import'.

Reinterpreted as characterising the wrongness of lying in terms of norm-violation, CIC places itself in a 'classic' tradition in the literature on the morality of lying. What is wrong with lying is here understood in *deontological* terms – there is a rule that lying infringes, and its blameworthiness is a function of such infringement. By contrast, Krauss' account identifies the blameworthiness of lying in terms of its expected harmful consequences, falling rather under a *consequentialist* tradition.

This might be seen as a sign that there is something counterintuitive about CIC. Prominent proponents of deontological accounts like Augustine and Kant also subscribed to 'absolutism' about lying: the view, often judged to be counterintuitive, that the impermissibility of lying is exceptionless (cf. Augustine [DM]; [CM], Kant [G]; [LE]; [RL]). But accepting a deontological story about the wrongness of lying does not commit one to absolutism. One can still maintain that

while all lying is wrong *qua violation* of a sincerity norm, any given act of lying can be nonetheless *overall* morally permissible; to put it in slightly different terms, a plausible version of this view is that lying is *prima facie* wrong (cf. Ross 1930). On this view, the fact that you say something you believe more likely to be false than true is a *defeasible* reason to classify your action as morally reprehensible. Furthermore, saying that lying is wrong in virtue of the violation of a sincerity norm does not entail that the wrongness of lying is exhausted by such violation. This view allows for *other* criteria of moral evaluation to be salient when we make moral judgments about lying (cf. Stokke 2017). For instance, the expected deceptiveness of a lie will typically be a salient dimension of evaluation, as hardly anyone denies that lying *typically* involves intended deception.

3.3 A Reconciliation: Lies and Deception

We can conclude by devising a partial reconciliation between Marsili's and Krauss' view. When it comes to a descriptively accurate definition of lying, lying is best understood as the act of asserting something insincerely, where insincerity is captured by CIC. All lies are *prima facie* wrong *qua lies*, to the extent (captured by CIC) that they all violate a norm of sincerity— the greater the violation, the more severe the blameworthiness that arises from such violation. Furthermore, *most* lies are *prima facie* wrong *qua attempts to deceive*; arguably, this kind of blameworthiness is a function of the extent to which they meet W-O: the greater the epistemic risk they impose on the hearer, the greater the blameworthiness of the attempt to deceive⁶. This gives us a two-fold account of the blameworthiness of lying: in terms of the credences they *necessarily* misrepresent, and of the inaccurate credences they *typically* aim to induce.

⁶ While there is growing consensus in the literature that lying does not *necessarily* involve an intent to deceive, some philosophers still subscribe to this view (e.g. Faulkner 2007; Lackey 2013). If they are right, then *all* lies are also blameworthy *qua attempts to deceive*. In either case, in light of Benton's (2018) counterexamples, W-O still needs to be refined to successfully track the blameworthiness of each and every act of attempted deception.

Conclusions

This article has proposed a solution to three fundamental questions in the philosophy of lying: (i) establishing the primary desiderata of a definition of lying, (ii) defining what lying is, and (iii) characterising its distinctive blameworthiness.

In relation to the first issue, I argued (against Krauss and a tradition that traces back to Grotius) that offering a definition of lying and characterising its moral import are two independent tasks, and that achieving the latter is conditional on achieving the former. When it comes to evaluating whether a definition of lying is correct, we should thus look at a definition's descriptive accuracy rather than its ability to explain the distinctive way in which lies are blameworthy, since only by meeting the former desideratum we can meet the latter. These considerations have important consequences for the ongoing philosophical debate about what lying is, and what makes it blameworthy. Once Krauss' mistaken assumptions about the desiderata for the definitions of lying are abandoned, it becomes clear that his critique of the 'comparative insincerity condition' is misguided. This in turn means that we can after all offer a fine-grained definition of lying that accounts for partial beliefs – for which no options are otherwise left available, given that neither Krauss' 'worse-off requirement' nor the traditional 'belief requirement' are able to draw the right distinctions about *graded-belief lies*.

Finally, I have argued that CIC has the resources to track the distinctive blameworthiness of lying, in terms of a (more or less severe) violation of a sincerity norm. Endorsing this explanation of the moral import of lying is compatible with acknowledging that also W-O captures a morally salient dimension of blameworthiness: the (higher or lower) epistemic damage that the liar typically aims to impose on the dupe. In advancing this partial reconciliation between two competing views about graded-belief lies, this paper has offered a novel, bipartite characterisation of lying and its distinctive blameworthiness: in terms of the norms that it necessarily violates, and in terms of the epistemic risks that it typically imposes on the hearer.

References

- Augustine. [DM], De Mendacio, 'On Lying,' M. S. Muldowney (trans.) (51–110)
- Augustine. [CM], Contra Mendacio, 'Against Lying,' H. B. Jaffee (trans.) (121–179), in R. J. Deferrari (ed.) *Fathers of the Church* (Volume 16: *Treatises on Various Subjects*), New York: Fathers of the Church, 1952
- Carson, Thomas L. 2010. *Lying and Deception*. Oxford University Press.
- Chisholm, Roderick M., and Thomas D Feehan. 1977. The Intent to Deceive. *Journal of Philosophy* 74 (3): 143–59.
- Fallis, Don. 2009. What Is Lying? *Journal of Philosophy* 106 (1): 29–56. [internal-pdf://0.0.4.12/FALWIL.html](https://www.jstor.org/stable/4012412).
- Faulkner, Paul. 2007. What Is Wrong with Lying? *Philosophy and Phenomenological Research* LXXV (3): 535–57.
- Fuller, Gary. 1976. Other-Deception. *The Southwestern Journal of Philosophy* 7 (1): 21–31.
- Grotius, Hugo. 1625. *The Rights of War and Peace*. Edited by Jean Barbeyrac (2005 edition). Indianapolis: Liberty Fund. <https://oll.libertyfund.org/titles/grotius-the-rights-of-war-and-peace-2005-ed-vol-3-book-iii>.
- Kant, I. [G]. *Groundwork of the Metaphysics of Morals*, in *Immanuel Kant, Practical Philosophy*, edited by A. W. Wood and M. J. Gregor, Cambridge: Cambridge University Press, 1996.
- Kant, I. [RL] *On a supposed right to lie from philanthropy*, in *Immanuel Kant, Practical Philosophy*, Cambridge: Cambridge University Press, 1996.
- Kant, I. [LE] *Lectures on Ethics*, P. Heath (trans.), P. Heath and J. B. Schneewind (eds.), Cambridge: Cambridge University Press, 1997.
- Kemp, Kenneth W., and Thomas Sullivan. 1993. Speaking Falsely and Telling Lies. Edited by Michael Baur. *Proceedings of the American Catholic Philosophical Association* 67: 151–70. doi:10.5840/acpaproc19936712.
- Krauss, Sam Fox. 2017. Lying, Risk and Accuracy. *Analysis* 73: 651–59. doi:10.1093/analys/anx105.
- Lackey, Jennifer. 2013. Lies and Deception: An Unhappy Divorce.

- Analysis* 73 (2): 236–48. doi:10.1093/analys/ant006.
- Mahon, James Edwin. 2007. A Definition of Deceiving. *International Journal of Applied Philosophy* 21 (2).
- Mahon, James Edwin. 2015. The Definition of Lying and Deception. *Stanford Encyclopedia of Philosophy*.
- Margolis, Joseph. 1963. “Lying Is Wrong” and “Lying Is Not Always Wrong.” *Philosophy and Phenomenological Research* 23 (3): 414–18.
- Marsili, Neri. 2014. Lying as a Scalar Phenomenon. In *Certainty-Uncertainty – and the Attitudinal Space in Between*, edited by Sibilla Cantarini, Werner Abraham, and Elisabeth Leiss, 153–73. Amsterdam: John Benjamins Publishing Company. doi:10.1075/slcs.165.09mar.
- Marsili, Neri. 2016. Lying by Promising. *International Review of Pragmatics* 8 (2). Brill: 271–313. doi: 10.1163/18773109-00802005.
- Marsili, Neri. 2018. Lying and Certainty. In *The Oxford Handbook of Lying*, edited by Jörg Meibauer, 169–82. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198736578.013.12.
- Ross, W. D. 1930. *The Right and the Good*. doi: 10.1093/0199252653.003.0001.
- Stokke, Andreas. 2017. Lies, Harm, And Practical Interests. *Philosophy and Phenomenological Research*, no. 2007: 1–17. doi: 10.1111/phpr.12439.
- Williams, Bernard Arthur Owen. 1985. *Ethics and the Limits of Philosophy*. Routledge.