
Article:
Marsili, N. orcid.org/0000-0001-7853-7359 (2016) Lying by Promising. International Review of Pragmatics, 8 (2). pp. 271-313. ISSN 1877-3095

https://doi.org/10.1163/18773109-00802005

Lying by promising.
A study on insincere illocutionary acts

Neri Marsili
University of Sheffield, UK
n.marsili@sheffield.ac.uk

This paper is divided into two parts. In the first part, I extend the traditional definition of lying to illocutionary acts executed by means of explicit performatives, focusing on promising. This is achieved in two steps. First, I discuss how the utterance of a sentence containing an explicit performative such as "I promise that *Φ*" can count as an assertion of its content *Φ*. Second, I develop a general account of insincerity meant to explain under which conditions a given illocutionary act can be insincere, and show how this applies to promises. I conclude that a promise to *Φ* is insincere (and consequently a lie) only if the speaker does not intend to *Φ*, or believes that he will not *Φ*, or both.
In the second part, I test the proposed definition of lying by promising against the intuitions of ordinary language speakers. The results show that, unlike alternative accounts, the proposed definition makes the correct predictions in the cases tested. Furthermore, these results challenge the following necessary conditions for telling a lie with content *p*: that you have to assert *p* directly; that you have to believe that *p* be false; that *p* must be false; that you must aim to deceive the addressee into believing that *p*.

Keywords: Speech act theory, Insincerity, Lying, Promising, Intentions

Part A: A theory of lying by promising

1. The definition of lying

According to a standard definition, lying is to state what you believe to be false with the intention of deceiving your audience into believing what you said[2]. More formally:

Standard definition

S lies to A about *p* iff:
i) S states *p*
ii) S believes ~p
iii) S intends A to believe *p*

In recent times, several philosophers have challenged this view, presenting some compelling counterexamples to the intention to deceive condition: lies under coercion (Siegler 1966:129, Carson 1988:511), lies to "go on the record", bald-faced lies (Carson

[2] See Mannison (1969:132), Kupfer (1982:134), Simpson (1992), Williams (2002: 96), Faulkner (2007; 2013), Meibauer (2005; 2014).

et al. 1982:17, Carson 2006:290, Sorensen 2007, Arico & Fallis 2013), and knowledge lies (Sorensen 2009). This has prompted many authors to drop (iii).

However, rejecting (iii) comes at a price: a definition only featuring (i-ii) runs the risk of incorrectly ruling in fictional, ironical and metaphorical utterances. These are believed-false statements (*i.e.* they meet both condition (i) and (ii)), but are clearly not lies Condition (iii) correctly prevents the definition from counting these statements as lies. A putative alternative definition that rejects (iii) needs some alternative condition to play its role in ruling out these cases.

A common way to respond to this challenge is to endorse an 'assertion-based' account of lying. The intuition behind this view is that the 'statement condition' of the standard definition can be narrowed down, to require that the speaker *genuinely assert* the proposition that he believes to be false. Formally:

Assertion-based definition

S lies to A about *p* iff:
i*) S asserts *p*
ii) S believes ~p

The assertion condition (i*) correctly rules out cases of fiction, irony and the like, that are not genuinely asserted. This definition still allows (and its proponents generally recognise) that intending to deceive is a feature that most lies have (or a 'prototypical feature' of lies, as shown in Coleman & Kay 1981). This alternative view traces back to St. Thomas: "The desire to deceive belongs to the perfection of lying, but not to its species, as neither does any effect belong to the species of its cause", and has attracted a strong consensus in the last ten years[3].

An assertion-based definition can be further strengthened by requiring that also (iii) is satisfied, and some proponents of the intention to deceive condition endorse an assertion-based version of the standard view, in which both (iii) and (i*)[4] are required

[3] Carson (2006; 2010), Sorensen (2007; 2009), Fallis (2009; 2012; 2014), Saul (2012), Stokke (2013). Except for Sorensen, all these authors spell out the assertion condition differently to (1*), specifying under which conditions a speaker's statement is a genuine assertion.

[4] Chisholm and Feehan (1977), Simpson (1992), Williams (2002), Faulkner (2007), Meibauer (2005; 2014). Most of these authors identify the *assertion condition* with the intention to deceive condition – they believe that asserting *p* requires intending that your audience believes that *p*. For them, meeting (i*) entails meeting (iii), so that calling these views 'assertion-based' is taxonomically misleading, as they hardly depart from the standard view. Moreover, these views are highly controversial, as the idea that asserting *p* requires intending your audience to

for lying. Overall, there is a general consensus that condition (i*) and (ii) are necessary for lying, the core of the disagreement being on whether (iii) should be added as a further requirement. The present paper does not take a side on this dispute: it will focus instead on how conditions (i/i*) and (ii) apply to illocutionary acts other than assertion. This is an under-investigated problem in the literature. According to the assertion-condition (i*), only assertions can be lies; however, it seems that some illocutionary acts that are not assertions, under some conditions, can be lies. For instance, it seems that (1) and (2) can be lies if the speaker is insincere; by contrast, sentence (3) can be infelicitous in many ways, but under no circumstances it can be appropriately called a lie.

(1) I warn you that Gervasio has a gun
(2) I promise that I will clean the toilet
(3) I order you to clean the toilet

The problem of which kinds of illocutionary acts other than assertions can count as lies is an open research problem for the definition of lying, that has been almost ignored in the literature (but cf. Leonard 1959, Falkenberg 1988, Meibauer 2014): this paper will take on the challenge, and extend the definition to other illocutionary acts. A theory of how the assertion-based definition applies to illocutionary acts other than assertion needs to show not only in which sense utterance like (1) and (2) satisfy the 'assertion condition' (i*), but also how the 'insincerity condition' (ii) applies to them, as it is not clear under which conditions speech acts like warning or promising can be insincere. Some theorists reject the possibility of lying by promising, or by uttering illocutionary acts other than assertion. For instance, Meibauer (2014:76) claims that insincere promises and warnings should not be dubbed lies. Similarly, in the first of his celebrated *William James Lectures,* Austin (1962/1975:256) claims that promising to do something that you do not intend to do is "perhaps misleading, probably deceitful and doubtless wrong, but it is not a lie or a misstatement" [6]. This is because in uttering a sentence

believe that *p* (or to believe that you believe that *p*) is nowadays regarded as indefensible (Searle 1969:42-9, Davis 1999, Alston 2000: 44-50, Green 2007: 75-82, Macfarlane 2011).
[6] In the following lectures, Austin abandons his *constative-performative* distinction, endorsing a different view that would count instead an insincere utterance of (EP) as a lie (cf. section 2.1).

containing an explicit performative[7] like (EP), one also always asserts something true, namely that one is promising that *Φ*.

(EP) I (hereby) promise that *Φ*

Mahon (2010:50'') points out that while the prediction that promises cannot be lies is counterintuitive, it is entailed by all definitions in the literature, since it follows from endorsing either (i) or (i*). Let us call any version of the claim that only assertions are eligible to meet (i*) the 'only-assertion' view.

ONLY-ASSERTION: Only assertions, and no other illocutionary acts, can count as lying

In the next section, I will attempt to do justice to the opposite intuition that not only assertions, but also other illocutionary acts (such as promising) can be used to lie. I will then move on to discuss how the assertion-based definition can be revised to capture these cases, focusing on promising in particular.

2. Meeting the assertion-condition

To provide a theory of how promises can satisfy the assertion condition, the first step is to explain how sentences containing explicit performatives like (EP) can satisfy the assertion condition in a meaningful way.

In (EP), *Φ* indicates a propositional content (generally a future action of the speaker), and "promise" is an explicit performative verb. By focusing on sentences of the form (EP), I am restricting my attention to a particular subset of promises, namely explicit, direct promises. This is because indirect promises, and in particular promises that *Φ* performed indirectly by asserting that one will *Φ* (as in (IP)), can be easily treated within the traditional account of lying, and do not require any further explanation: they clearly satisfy (any version of) the assertion condition, so that they count as lies if the other conditions obtain.

(IP) I will *Φ*

[7] The standard form of a *performative sentence* is "I (hereby) [Performative Verb] that *Φ*". Performative sentences can take the first person plural, impersonal or passive forms, or can be embedded under complex construction; for our purposes, however, we can ignore these complications.

2.1. *Explicit Performatives*

Broadly, one can identify two strands of theories of explicit performatives, *descriptivist* and *non-descriptivist* theories. On a *descriptivist* account[8], performative sentences like (4) are direct assertions that a promise is being made, and under relevant circumstances they can license the inference that a further illocutionary act (in this case a promise with content (4*)) has been made.

(4) I promise that (4*) [I will feed the walrus]

Non-descriptivist accounts[9] maintain instead that a sentence like (4) is uniquely a direct promise with content (4*), and that (4) is not an assertion, as it cannot be properly described as being true or false.

At first glance, the *non-descriptivists* are in a better position than *descriptivists* to allow for lying with explicit performatives. Consider a sentence like (5), where the difference between the two accounts is even more blatant:

(5) I assert that (5*) [I love pizza]

Intuitively, a sentence like (5) can be used to lie: if I utter (5) and I don't love pizza (I believe (5*) to be false), I am clearly lying. This is exactly what the non-descriptivists account predicts: (5*) is the *content* of the primary speech act performed in uttering (5), and the explicit performative "I assert that" merely serves the purpose of making the assertoric force explicit.

According to the descriptivists, by contrast, in uttering (5) I simply assert (5): I assert that "I assert that I love pizza". Since by uttering (5) *I make it the case* that (5) is true, I cannot utter (5) insincerely, and hence I cannot lie in uttering it. Searle (1989) talks in this sense of a 'self-guaranteeing' nature of performative sentences: "performative

[8] Hedenius (1963), Lewis (1970), Bach (1975), Ginet (1979) and Bach & Harnish (1979:§10; 1992), Green (2005).

[9] Austin (1962/1975), Harris (1977), Searle & Vanderveken (1985), Searle (1989), Reimer (1995), Jary (2000).

utterances are self-guaranteeing in the sense that the speaker cannot be lying, insincere, or mistaken about the type of act being performed" [10].

To allow for lying with explicit performatives, the descriptivists need to show that they can treat *both* (5) *and* (5*) as asserted. Some versions of descriptivism yield exactly this prediction: since the utterance of (5) generally licenses the audience to infer that also an assertion that (5*) has been made, in asserting (5) you generally succeed in making your audience realise your intention to also assert that (5*), so that you eventually assert both (5) and (5*) (cf. Bach & Harnish 1979: 208).

However, on this interpretation (5*) is not explicitly or directly said, but rather indirectly asserted, *i.e.* implicated (some would describe it is a 'generalised conversational implicature', as in Grice 1989). In other words, it is something that the hearer has to infer from the speaker's utterance. But if (5*) is merely implicated it cannot qualify as lying, because lying requires explicitly (*i.e*. directly) *saying* something that you believe to be false, rather than *implicating* it[11]. Since the descriptivist view counts proposition like (5*) as implicated, it incorrectly characterises utterances like (5) as misleading rather than lying.

All in all, the non-descriptivist view seems better suited to account for lying with explicit performatives. This is not to deny that there might be ways to cash out the descriptivist view so that it counts both (5) and (5*) are literal assertions[12], overcoming the difficulties considered in this section. In what follows, however, I will maintain a neutral position on this latter respect, and adopt a non-descriptivist framework to characterise lying with performative utterances.

2.3. Asserting by promising

The non-descriptivist account is able to treat performative utterances like (5) (*explicit assertions*) as direct assertions of (5*). However, to account for lying by promising

[10] The argument from the self-guaranteeing nature of explicit performatives is a classic argument against descriptivism. To be sure, in some cases a speaker S can utter (5) even if it is false that "S asserted that S loves pizza", and S believes it to be false; for instance, if S utters (5) alone, just to practice English language. However, if lying requires *meaning* what you say, this kind of examples cannot help. Even if descriptivists can defend that (5) can be believed to be false, it is more difficult for them defend that its utterance can count as lying.

[11] See Adler (1997), Saul (2011; 2012), Stokke (2013); an exception to this view being Meibauer (2005; 2014)

[12] An example is García-Carpintero's (2013) proposal, which is descriptivist in spirit, but counts (5) as a mere 'locution' and (5*) as a direct assertion.

within an assertion-based view of lying, one also needs to show that the utterance of an *explicit promise* like (4) can count as an assertion of (4*).

In this section, I want to show that Searle & Vanderveken's (1985) *non-descriptivist* framework that can meet this challenge, and clarify how it allows for some illocutionary acts, including promises, to count as indirect assertions[13].

According to Searle & Vanderveken, the standard way for a speaker to successfully perform an illocutionary act F(p) is to utter a sentence that expresses literally that force and content: this is called a *literal performance* of that act. In this sense, (EP) is a literal performance of a promise. Literal performance is to be contrasted with two ways of performing a speech act *non-literally*.

First, there are cases of non-literal performance determined by the fact that the performance of an illocutionary act *entails* the performance of another (1985:129-30). The performance of an illocutionary act $F_1(p)$ entails the performance of another illocutionary act $F_2(p)$ iff in the context of the utterance it is not possible for S to perform $F_1(p)$ without performing $F_2(p)$ – so that if S performs $F_1(p)$, S also performs $F_2(p)$. In this sense, the explicit promise (4) illocutionarily entails the assertion of (4*), since I cannot promise that I will feed the walrus without asserting that I will feed the walrus:

(4) I promise that (4*) [I will feed the walrus]

Second, there are non-literal performances of $F_2(p)$, that the speaker performs by way of performing another illocutionary act $F_1(p)$ that does entail $F_2(p)$, but in the performance of which he makes salient his *reflexive intention* (cf. Grice 1989, Bach & Harnish 1979) to make the hearer understand that he intends to perform $F_2(p)$ (either instead of, or in addition to, $F_1(p)$). If the speaker does so and the conditions for the successful performance of $F_2(p)$ are met, then the speaker has indirectly performed $F_2(p)$. In this sense, uttering (6) generally amounts to implicating (6a):

[13] Some assertion-based accounts, like Carson's (2006; 2010) and Stokke's (2013; 2014), could also allow for promises to count as lies if paired with a non-descriptivist account of explicit performatives. I have independent reasons to prefer a speech-act theoretic account, that I aim to explore in another paper. Shortly put, a speech-act theoretic view generalises more easily to other illocutionary acts, and incorporates a criterion to determine how the sincerity condition changes depending on the speech act performed (cf. section 3).

(6) I ate three bananas
(6a) I did not eat more than three bananas

This distinction identifies two different ways of indirectly asserting: *entailment,* that is compatible with satisfying the assertion-condition, and *implicature*, that is not compatible with it. Entailment does not rely on expecting some additional hearer inference about the speaker's communicative intention, and cannot be retracted. By contrast, an implicature relies on expecting some additional hearer inference about the speaker's communicative intention, and generally it can be retracted. I cannot promise (4) without asserting (4*), so that if I believe that I will not feed the walrus I am lying. By contrast, I can assert (6) without intending to imply (6a), so that if I actually ate six bananas, my utterance is perhaps deceptive and misleading, given that (6a) is false, but it is not a lie, because what I literally said, namely (6), is true.
These examples clarify why the indirect performance via entailment is compatible with lying, while implicature is only compatible with misleading. Example (4-4*) also illustrates the sense in which a promise can *entail* an assertion in the sense specified, and consequently meet the 'assertion condition' (i*). It should be noted that this view could be generalised to other illocutionary acts, to determine which ones (under normal conditions) entertain this relation of illocutionary entailment with assertion, and which do not. However, I will need to leave this task aside for the moment[14], and move on to discussing how the insincerity condition for lying apply to promises.

3. Insincerity condition

Under which conditions is a promise a lie? According to 'assertion-based' definitions, a speech act that satisfies the 'assertion-condition' (i*) is a lie only if it also meets the insincerity condition (ii), *i.e.* if its content (identified in a *non-descriptivist* fashion) is

[14] Searle & Vanderveken (1985:182-92) offer a more detailed explanation of what determines entailment, and list of the assertives that entail an assertion (cf. also Marsili 2015:§3). Surely *commissives* (the class including both promising and swearing) can entail assertions, and clearly no *directives* (the class including commands and requests) can. The case of *expressives* and *directives* is more complicated, and I aim to discuss them in another paper.

believed to be false. Call this view, that is endorsed by virtually every author in the literature[15], the 'only-belief' view of the insincerity condition for lying.

ONLY-BELIEF: an utterance can be a lie only if its content is *believed to be false*

In this section, I will challenge the only-belief view, and argue that the insincerity condition for lying needs to be expanded to other propositional attitudes (e.g. intentions, desires) if it is to account for illocutionary acts in general, and for promises in particular. In order to do so, I will first discuss what it is, in general, for a speech act to be sincere or insincere, and then apply this account to promises.

3.1 Expressing attitudes and (in)sincerity

It is a standard view in speech act theory that each illocutionary act expresses a distinctive propositional attitude (Searle 1969, Bach & Harnish 1979). The distinct attitude expressed by a given illocutionary act is part of what identifies it as opposed to others, and it is generally taken to define the point or purpose of the actions that we perform in uttering it. In this sense we say that an assertion expresses a *belief*, that a promise expresses an *intention,* that asking someone for something expresses a *desire*. Philosophers and linguists have presented different taxonomies of illocutionary acts based, amongst other things, on the different psychological attitudes expressed by different (kinds of) illocutionary acts. Most authors would agree that the following characterisation is broadly correct:

A-EX: If S asserts that *p,* S expresses the belief that *p*
R-EX: If S asks for *p,* S expresses a desire for *p*
P-EX: If S promises that *p,* S expresses the intention to do *p*

There are several ways to flesh out what is meant by 'expressing' a psychological state. I will assume that *expressing* an attitude does not entail having that attitude, so that you can insincerely express an attitude that you do not have (but cf. Davis 2003: 25, Green 2007:70-83). Following a hint by Davidson (1985:88, cf. Marušić 2012:13, Fallis 2013),

[15] There is agreement that the content must be believed to be false, but not that it must be identified in a non-descriptivist fashion. See Marsili (2014:156-162) for a review of accounts that depart from the only-belief view, imposing different constraints on belief.

one could say that for a speaker to express a psychological state is for the speaker to *represent himself* as being in that psychological state.

On an orthodox account of sincerity[16], the sincerity condition for uttering a given illocutionary act is that the speaker has the psychological attitude expressed by that act:

SIN-SA: The performance of an illocutionary act F(p) that expresses the psychological state Ψ(p) is *sincere* IFF in uttering F(p), S is in Ψ(p)

From the orthodox account of sincerity, a simple account of insincerity clearly follows: a speaker is insincere whenever he is not in the psychological state that is expressed by the illocutionary act performed:

SA-INS: The performance of an illocutionary act F(p) that expresses the psychological state Ψ(p) is *insincere* IFF in uttering F(p), S *is not* in Ψ(p) [17]

This gives us the following insincerity conditions for the three illocutionary acts we are considering as examples:

A-INS: S asserts that *p* insincerely only if S does not believe that *p*
R-INS: S asks for *p* insincerely only if S does not desire *p*
P-INS: S promises that *p* insincerely only if S does not intend to *p*

This account, however, fails to deal properly with graded attitudes. For instance, sometimes a speaker is only partially confident that the proposition that he has asserted is true: we say in this case that the speaker has a partial, rather than outright, belief in *p*. Now, suppose I utter (7) and I think that (7) is *probably* true:

(7) J.L Borges is the author of *The Invention of Morel*

[16] This view has been defended, under different guises, by Hare (1952:13,19–20,168–99), Searle (1969:60, 64–8); Wright (1992:14), Moran (2005), Green (2007:70-83). The most prominent alternative account takes a speech act to be sincere if it is not aimed at deceiving, and insincere if it is (Williams 2002:74). Since this would conflate the insincerity condition with the intention to deceive condition, and lying with misleading, it is uninteresting for refining the definition of lying. Cf. Eriksson (2011) and Stokke (2014) for an overview.

[17] Classic objections to the orthodox account of insincerity are cases self-deception (in which the speaker is mistaken about his own mental states, cf. Moran 2005, Eriksson 2011, Chan & Kakane 2011, Stokke 2014) and malapropism (in which the speaker is mistaken about what he said, cf. Reimer 2004, Sorensen 2011, Saul 2012:15-9). However, these challenges can be dealt with just by specifying that the speaker has to satisfy (any version of) the insincerity conditions for lying *advertently* (or by endorsing qualified definitions like SA-INS2). "Advertently" in this case indicates that the speaker is aware of the divergence between the psychological attitude he expressed and the one he instantiates.

According to A-INS, in uttering (7) I am insincere, because I lack the full belief that I am expressing in asserting (7), and hence I am lying. However, asserting (7) while believing that it is probably true is in some relevant sense sincere, and clearly it does not amount to lying (but cf. Marušić 2012:8). To account for these cases, one can require that the speaker is more confident in *p* than he is in its negation (Marsili 2014; forth.):

A-INS2: S asserts that *p* insincerely only if S is more confident in *p* than he is in not *p*

This revision hardly applies to illocutionary acts that express psychological states other than beliefs, given that attitudes like intentions do not come in degrees. However, for any act involving gradable attitudes, SA-INS can be generalised just by specifying that insincerity with respect to psychological state Ψ(p) requires the speaker to be in Ψ(~p) more than he is in Ψ(p).

SA-INS2: The performance of F(p) is *insincere* IFF in uttering F(p) S takes himself to express that he is Ψ(p), but S takes himself to be in Ψ(~p) more than he is in Ψ(p)

This gives us a general framework to understand how the insincerity conditions for lying might apply to illocutionary acts in general, and to promising in particular.

3.2 Insincerity conditions for promising

So far, we have encountered two competing ways of accounting for the insincerity conditions for lying by promising. The *first* (section 1) is the approach traditionally used to define lying: the *only-belief* account. This approach applies indifferently the same insincerity conditions to every speech act that satisfies the 'statement condition', and counts them as insincere whenever their content is believed to be false.
The *second* is the speech act theoretic account considered in the previous section: (i) a promise expresses an intention to *Φ*, and consequently (ii) the insincerity condition for promising is intending not to *Φ*. Assumption (i) can be traced back to Hume's view (*THN:* 517-19) that a promise always *expresses* (and communicates) an intention to perform the promised act[18]. Assumption (ii) is found in foundational works of speech

[18] This view is extremely influential in the philosophical literature on the nature of social obligations. Many authors (sometimes referred to as 'information-interest' theorist) take promising's main function to be informing the promisee of what the promisor is going to do

act theory, like those of Austin (1962/1975:50, 135-6) and Searle (1965:243, 1969:60-2). I will refer to this account as the *only-intention* account. On this view a promise is insincere only if, *at the time of the utterance*, the speaker does not intend to perform.
A *third* account of the insincerity conditions for promising can be derived from the relation of entailment between asserting and promising that outlined in 2.3. If one cannot promise to *Φ* without also performing an assertion that one will *Φ*, every time one promises to *Φ* one also asserts that one will *Φ*. And if an assertion is thus performed in addition to the promise, the sincerity conditions for asserting need to be satisfied too, *i.e.* performing the act sincerely also requires one to believe that one will *Φ*. This yields a novel account of the insincerity conditions for promising:

Entailed-Insincerity condition:
A promise is insincere if speaker intends not to *Φ*, or if the speaker believes that he will not *Φ*, or both

More specifically, promise is insincere *qua promise* if the speaker intends not to *Φ*, and insincere *qua assertion* if the speaker believes that he will not *Φ*. This view (be it correct or not) can clearly be generalised: whenever there is illocutionary entailment, and two illocutionary acts are performed, the sincerity conditions of both acts apply.
To sum up, the three candidate insincerity conditions for lying by promising are:

(BIC): *Belief insincerity condition*: S believes that S will not *Φ*
(IIC): *Intention insincerity condition*: S intends not to *Φ*
(EIC): *Entailed insincerity condition:* BIC ∨ IIC

Which account is preferable? The Moorean test for insincerity provides some linguistic data that *prima facie* favours the *entailed-insincerity* view over the other two. It is well known that assertions followed by the negation of their sincerity condition give rise to Moorean absurdities (Moore 1993:210): asserting "p and I don't believe it" is incoherent in some distinctive way. One way of explaining this incoherence is that in uttering these sentences, the speaker performs a speech act and then blatantly violates one condition for its felicitous (in this case, sincere) performance, eventually failing to assert that *p* (Vanderveken 1980, Searle and Vanderveken 1985:150,152). If both BIC

(Sidgwick 1981:442–44, Anscombe 1981:18, Rawls 1981:345, Fried 1981:16, Foot 2011:45). This view has been opposed by Owens (2008:747-51). His arguments seem successful in establishing that in promising to *Φ* one does not necessarily *communicates* an intention to *Φ*, but it is less clear that they demonstrate that in promising to *Φ* one does not necessarily *expresses* an intention to *Φ,* or that sincerely promising does not require intending to *Φ*.

and IIC are insincerity conditions for promising, it seems that they should both give rise to the same kind of unsuccessful incoherence (cf. Marušić 2012:14). As a matter of fact, sentences like (8~B) and (8~I) display this kind of absurdity:

(8~B) I promise that I will pick you up at 6, but I don't believe I will pick you up at 6 #
(8~I) I promise that I will pick you up at 6, but I don't intend to pick you up at 6 #

In both cases, the utterance strikes one as incoherent, and in both cases, it is difficult to imagine that the speaker will be taken to have promised to pick his interlocutor up at 6. This linguistic data supports EIC, and cannot be easily explained by BIC or IIC taken separately.

One might still wonder whether endorsing EIC actually makes a difference, rather than simply introducing needless complications. If in the vast majority of cases one's intention to *Φ* is accompanied by the belief that one will *Φ* (and vice versa), requiring both is redundant. However, as Marušić (2013:293) correctly points out, we utter promises where intentions and beliefs come apart all the time. We often promise "to exercise more, to get out of debt, to quit smoking, to learn a foreign language, or to spend the rest of our life with our spouse" in situations in which we both intend to stick to the promise and *believe* that we will probably end up violating it.

To see how this relates to lying, consider the following example: a mechanic utters (9) and both *intends* to do his best to repair the car and *believes* that the damage is so serious he will almost certainly fail:

(9) I promise that I will repair your car by next week

The mechanic says what he believes to be almost certainly false: intuitively, (9) is a lie (cf. A-INS2). The *only-intention* view is not able to track this intuition; the *entailed-insincerity* view, by contrast, correctly predicts that (9) is a lie.

Arguably, cases like (9) are not straightforward, or prototypical, cases of lying. But this is also a prediction of the *entailed-insincerity* account: (9) is insincere *qua assertion*, but not insincere *qua promise.* The mechanic intends to fulfil the promise, following one sincerity condition, but he believes he will almost surely fail, violating the other. In Part B, I consider a complementary case (a promise satisfying BIC but not IIC), and report

empirical evidence showing that native English speakers judge that both kinds of violation are lies.

All in all, it seems that there are solid reasons to prefer the *entailed-insincerity* account of promising. So far, only Marušić (2013) has defended a similar view (Austin 1961/2003:239 merely hints at this idea). However, Marušić also suggests that a promise that violates BIC but not IIC might be better described as *irrational* rather than *insincere*. This is true if one endorses a 'cognitivist view' (cf. Appendix II), contending that it is irrational to intend to do something you believe that you will not do, so that only utterances satisfying both BIC and IIC (or neither of them) are rational, while utterances like (9) are *irrational* rather than *insincere*.

Marušić's observation points out a possible problem for the proposed account: if a rational intention to *Φ* requires believing that one will *Φ*, then there is no need to require both IIC and BIC, as the satisfaction of the first entails the satisfaction of the second in every case in which the speaker is rational. If readers have this sort of worry, they can refer to Appendix II: there I argue, against the cognitivist view, that (as long as you take it as a live possibility that you will *Φ* – *i.e.* as long as you are not certain that you will not *Φ*) you can rationally intend to *Φ* and believe that you will very likely not *Φ* (or vice versa). On this *weak cognitivist* view, Marušić's observation is not a worry: utterances like (9) are *insincere* rather than *irrational*, and EIC is preferable to the other accounts exactly because it successfully captures these peculiar forms of insincerity.

3.3 Conclusions

According to the proposed account, promising entails asserting, and the utterance of a promise is insincere when the *belief insincerity condition* (BIC) or the *intention insincerity condition* (IIC) or both are satisfied. Violations of both are standard cases of insincerity, while violations of a single one result in less severe cases of insincerity. Plugging this view into the assertion-based definition of lying, we obtain the following definition:

Definition of lying by promising

In promising that *p,* S lies to A iff:

1. S successfully promises that p
2. BIC ∨ IIC

This account can be extended to any illocutionary act, thanks to the account of the assertion-condition and of the insincerity-conditions developed in the previous sections.

Definition of lying by F(p)-ing

Where F(p) is an illocutionary act with content *p* that expresses a psychological state Ψ(p) and illocutionarily entails an assertion that *p*, S lies to A about *p* iff:

1. S successfully performs F(p)
2. BIC ∨ SΨ(~p)

In order to develop and defend my account, I have often appealed to intuitions, claiming that its predictions are consistent with intuitions about lying, and attacking alternative views as failing to track such intuitions. In the next section, I present an experimental study that tests this hypothesis. It shows that the intuitions of native English speakers indeed support the proposed account, and that they are incompatible with all the alternative ones: the *only-assertion, only-belief,* and *only-intention* accounts, and the *cognitivist* view of intentions and beliefs.

Part B – An experimental investigation about lying by promising[20]

4.1 Testing folk intuitions about lying

In the philosophical literature, it is generally agreed that a good definition of lying should track the ordinary usage of the term (Carson 2006:285; Fallis 2009:32): most debaters are after a characterisation of lying that is in line with the linguistic practice of competent speakers. A good account of lying should predict which usages of the term are correct and incorrect according to competent speakers. A corollary of this way of thinking is that if an account of lying makes predictions that are inconsistent with ordinary people's intuitions, that account fails to meet one important *desideratum* of a theory of lying. With this in mind, philosophers and linguists have started to accumulate data about ordinary speakers' intuitions about the correct usage of the term. These

[20] The experimental part of this study has been conducted in collaboration with Miklos Kurthy of the Experimental Philosophy Group at the University of Sheffield. Kurthy is responsible for all the statistical analysis, and in general for providing support in designing the experiment and in collecting and analysing the data.

studies have the potential of giving us insight into what lying is, or at the very least into what lying is perceived to be within a community of speakers[21].

Numerous and diverse empirical studies, stemming from different theoretical backgrounds and motivated by different explanatory aims, have attempted to explore folk intuitions about lying[22]. Among the ones explicitly investigating the intuitions of competent speakers about the concept of lying, the most important strand comes from the framework of prototype semantics. Following the lead of Coleman & Kay (1981), these studies attempt to identify the features of a prototypical lie, and to outline the differences of these prototypes across different cultures (Sweetster 1987, Cole 1996, Hardin 2010, Rong, Chunmei & Lin 2013). The present study addresses similar worries, but in a slightly different framework, namely that of experimental philosophy. Here, the aim of the analysis is to identify the necessary and sufficient conditions for an utterance to be a lie, rather than the prototypical features that make up the concept. Only three studies on lying have been conducted within this framework so far (Arico & Fallis 2013, Turri & Turri 2015, Wiegmann et al. 2016), the first one attempting to test if the intention to deceive is necessary for lying, and the second two if actual falsity is. The present experimental study will instead try to establish which conditions are necessary for lying by promising.

4.2 Aim of the study

This experimental study aims to test the theories developed in the first part of the paper against the intuitions of native English speakers. This means that it will attempt to determine:

(a) If ordinary people rate some insincere promises as lies, and
(b) Which *insincerity conditions* have to be satisfied for them to do so

As a reminder, the candidate insincerity conditions are four combinations of BIC and IIC: only BIC, only IIC, (BIC & IIC), and (BIC ∨ IIC).

[21] For a more detailed defence of the importance of tracking ordinary intuitions for a definition of lying, see Fallis (2009) and Arico & Fallis (2012:794-7)
[22] Here I am only considering studies on competent speakers. For a broader a review, including studies in developmental psychology, see Hardin (forth.).

IIC: the speaker does not intend to *Φ*

BIC: the speaker does not believe that he will *Φ*

Testing (a) is relatively simple: it is sufficient to create a story in which it seems that a character lies by promising, and ask participants whether the character has lied. Testing (b) is a slightly more complex matter. Here we need different stories in which different combinations of insincerity conditions are violated, and for each story test whether the participants believe that the speaker has lied. Given our candidate insincerity conditions, we will need to consider three scenarios in which a character promises something insincerely. In the *straightforward scenario*, the character's (S) utterance satisfies both BIC and IIC. In the *no-intention* scenario, it satisfies IIC but not BIC. In the *no-belief* scenario, it satisfies BIC but not IIC. I will refer to the latter two cases as the *crucial* conditions, as opposed to the *control* (straightforward) conditions.

- *Straightforward scenario*: BIC & IIC [CONTROL]
- *No-intention scenario*: IIC & ~BIC [CRUCIAL]
- *No-belief scenario:* BIC & ~IIC [CRUCIAL]

4.3 The predictions of existing theories

In the first section of the paper, I have mentioned five approaches to lying by explicit promising. Each accounts gives different predictions about which of the three scenarios will be rated as a lie.

1. According to the *only-assertion paradigm* (∅), lying requires a direct assertion, therefore explicit promises cannot count as lies. Giving a negative response to (a), this account simply dodges question (b): there is no 'insincerity condition' for lying by promising, because one cannot lie by explicitly promising. Its prediction is that respondents will claim that in no scenario the character is lying.
2. The *only-belief paradigm* (BIC) rigidly assumes that you lie only if you *believe* that the propositional content of your speech act is false. This view expects positive responses in the *straightforward* and in the *no-belief* condition, but negative responses in the *no-intention* condition.
3. The *only-intention paradigm* (IIC) maintains that a promise is insincere iff the speaker does not intend to fulfil his promise. Applied to lying, this view predicts that a promise

is a lie iff the speaker does not intend to fulfil it. The *straightforward* and *no-intention* cases should then be rated as lies, but not the *no-belief* case.

4. According to a *cognitivist interpretation* (BIC & IIC), the *no-intention* and *no-belief* should be described as cases of irrational thinking rather than lying, so that only the *straightforward* case should be rated as a lie.
5. According to the *entailed-assertion* paradigm (BIC ∨ IIC), a promise is a lie either if the speaker does not intend to fulfil his obligation, or if he believes that he will fail to fulfil it, or both. The account also predicts that all scenarios will be rated as lies, but expects the *straightforward* one to receive slightly higher ratings than the *crucial* ones.

	Straightforward	NO INT	NO BEL
Only-assertion	NO	NO	NO
Only-belief	YES	NO	YES
Only-intention	YES	YES	NO
Cognitivist	YES	NO (IRR)	NO (IRR)
Entailed-assertion	YES	YES (-)	YES (-)

Table 1: the predictions of the five different accounts of lying by promising

4.4 Experiment 1

4.4.1 Method

Participants: Participants were recruited using Amazon Mechanical Turk and tested using Qualtrics. They were compensated $0.2 for taking the survey. Repeated participation was prevented. Overall, 166 U.S. residents were tested (85 females; mean age (SD) = 36.6 years (12.9); range: 18–72; 100% reporting English as a native language). To prevent participants from taking the test negligently, a minimum response time (25 seconds) and a control question were set. Data from three participants who failed to meet these conditions were excluded, but including them would not affect the results.

Design: Each participant was randomly assigned to one of four conditions. Each condition features Coco and Baba, and in each condition Coco promises something to Baba. The first two conditions [(1) *straightforward lie*; (2) *no intention*] belong to the 'drink' story, and the second two conditions [(3) *straightforward lie*; (4) *no belief*] belong to the 'repair' story. In the 'drink' story, Coco promises not to drink; in the 'repair' story, Coco promises to repair Baba's car. For each pair, in the *straightforward* condition Coco lacks both intention and belief, and in the crucial case he lacks one attitude (intention in 2, belief in 4) but not the other.

Having been assigned to one of the conditions and having read the relevant story, the participants were posed two questions, always in the same order. The first asked whether Coco told a lie ("Did Coco tell a lie"? Y/N). The second allowed participants to report whether they felt uneasy in answering the preceding question ("Did you find it easy to make a decision"? Y/N); those answering "no" were solicited to motivate their uneasiness via a simple feedback form. The second question was especially designed to rule out the possibility that participants would have preferred not to give a dichotomic yes-no response, but also to collect qualitative data about the strength of the participants' intuitions.

Some peculiarities of the design are due to consistency constraints on the rationality of "intending without believing", and vice versa (cf. Appendix II). The first peculiarity is that in all conditions Coco has a partial rather than outright belief in whether he will fulfil his promise. This is due to the adoption of *weak cognitivism* in this research, according to which intending to Φ entails not being certain that you will fail to Φ. To grant uniformity between all conditions, Coco has a partial belief both in the crucial cases (where it could not be otherwise) and in the straightforward ones. Similar consistency constraints (also discussed in Appendix II) motivated one asymmetry in the experimental design, *i.e.* the fact that the no-intention and no-belief cases were not tested within the same story. The reason is that an uncontroversial *no-intention* case demands a promise about *refraining from acting,* while an uncontroversial *no-belief* case demands a promise about *actively performing one action*.

4.4.2 Results and discussion

Virtually every respondent rated the *straightforward cases* (in which Coco lacks both belief and intention to perform) as lies: 95% of the participants claimed that Coco lied in the *drink-straightforward* condition (38 of 40) and 95% in the *repair-straightforward*

condition (39 of 41). All except one participant (in the drink scenario) declared that the question was easy to answer. The results for the straightforward cases support the view that it is possible to lie by explicit promising, refuting the *only assertion* hypothesis. They also back the stronger claim that insincere promises can be regarded as *prototypical* cases of lying; and that, more generally, a prototypical lie can be performed by uttering a speech act other than assertion.

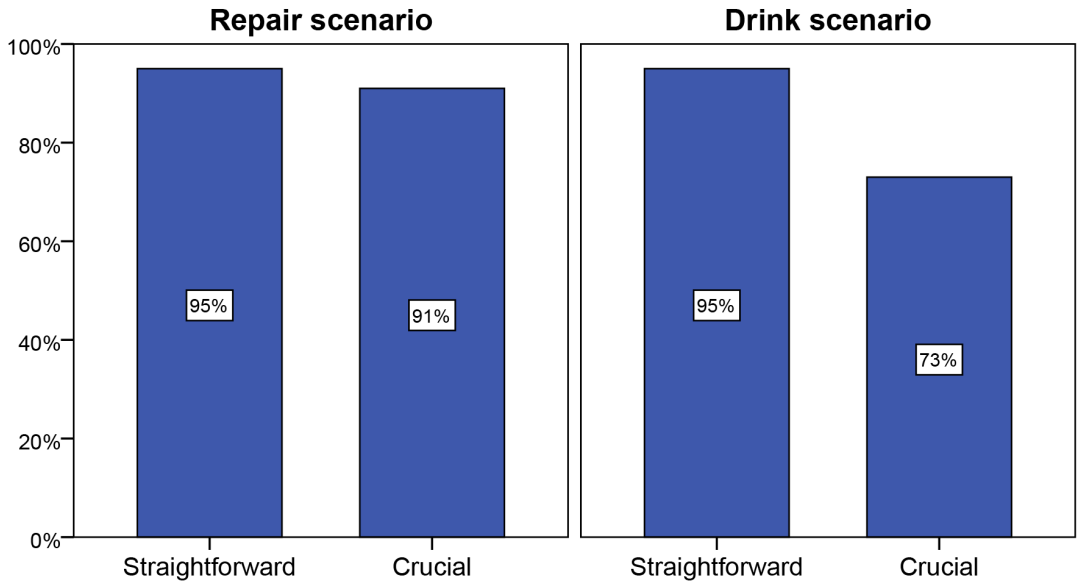


Figure 1: Percentage of respondents rating the protagonist's utterance as a lie in each condition.

In the *no-intention* condition, 90% of the participants (40 of 44) rated the promise as a lie. This refutes the predictions of the *only-belief* and of the *cognitivist* accounts, as respondents classed the promise as a lie even if Coco sincerely believes that he will (almost certainly) fulfil his promise. The difference between this condition and the corresponding *drink-straightforward* case was not significant, $\chi^2(1, N = 84) = .53$, $p = .467$; nonetheless, as much as 14% of the participants (6 of 44) declared that the question was not easy to answer, which is significantly more than in the *straightforward* conditions (Fisher's Exact Test, $N = 125$, $p = .008$, two tails)**.** This shows that intuitions in this condition are not as strong as in the *straightforward* cases, and suggests that the *no-intention* condition is regarded by some participants as a non-paradigmatic instance of lying.

The *no-belief* condition was rated as a lie by 73% of the participants (30 of 41). A binomial test showed that this result is significantly different to a chance distribution of the responses ($p = .004$**,** two tails); in other words, participants were more likely to say

that the promise was a lie than the opposite (OR = 2.73), which logically entails that the *only-intention* account (according to which this case does not qualify as lying) can also be rejected. In this condition, 15% of the participants (6 of 41) declared that the question was difficult to answer: Fisher's Exact Test revealed that this was significantly different to the straightforward cases (p = .006, two tails, OR = 13.71), suggesting that at least some participants did not see this as a paradigmatic case of lying. That the case might not be seen as prototypical is also confirmed by the lower percentage of ratings of the promise as a lie: this case differs significantly from all the rest of the cases jointly, $\chi^2(1, N = 166) = 12.71$, p = .001, as well as from the no-intention case separately, $\chi^2(1, N = 85) = 4.6$, p = .032.

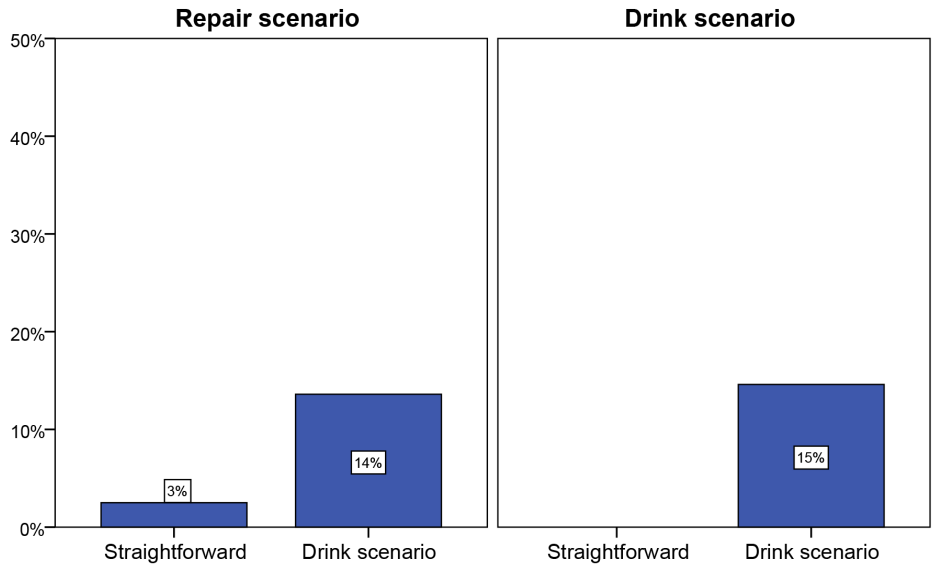


Figure 2*:* Percentage of respondents rating the question as difficult to answer (scale from 0% to 50%).

Interestingly, none of the feedback forms filled in by the participants who found it difficult to answer the question mentions the contrast between intention and belief as problematic, suggesting that the *cognitivist account* of intentions does not reflect the participants' intuition. Logistic regression was used to test whether information about gender or age could be used to predict the responses, revealing that neither gender (p = .592) nor age (p = .808) was a significant predictor.

The results provide strong support for the definition of lying by promising elaborated in part A. However, one can still worry that the dichotomous yes/no design led some participants to polarize their responses in an unnatural way (cf. Xu, Lu, Fuo & Lee 2009:318, Turri & Turri 2015:163). In section 3 we have seen that beliefs, and hence

insincerity, can come in degrees, and that contrasting parameters can influence our evaluation of a statement as insincere. Perhaps participants see lying as a scalar or vague concept (cf. Bazzanella 2009, Marsili 2014; forth.), and would have preferred to give a graded answer. In fact, 14% of the participants in the crucial cases (12 of 85) declared that the question was difficult to answer. Maybe they gave a positive answer to sanction that the speaker misbehaved in some important way – but they would not have classed the utterance as a lie if an intermediate alternative was given. If this line of reasoning is correct, and the probes forced the participants to polarise their responses, the proposed interpretation might be mistaken. To test this hypothesis, the test was rerun with a slightly modified design that allowed for graded rather than dichotomous responses.

4.5 Experiment 2

4.5.1 Method

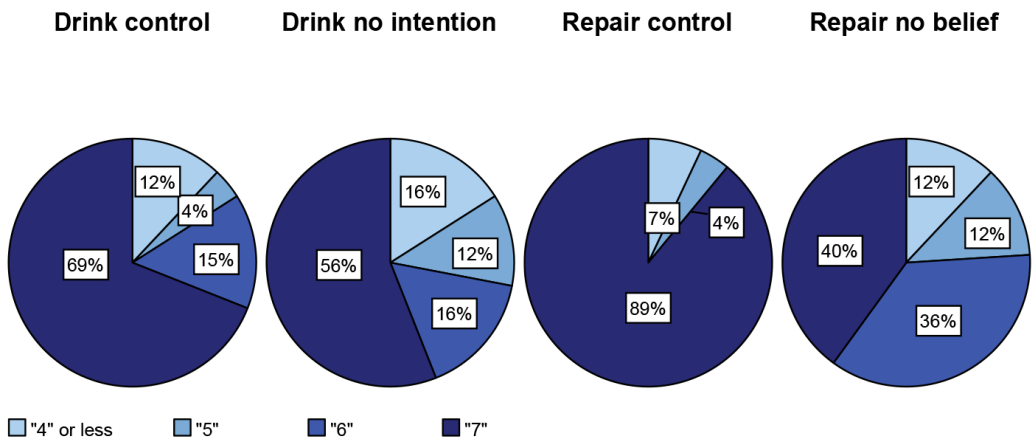
Participants: Participants were 104 adults (57 females; mean age (SD) = 37.1 years (11.8); range: 19-64; 100% reporting English as a native language, recruited and tested online as in Experiment 1). One participant failing to meet the minimum response time was excluded, but including him would have not affected the results.

Design: The design was the same as Experiment 1, but instead of using a dichotomous yes/no measure, participants were asked to rate their agreement with the statement "Coco told Baba a lie" on a 7-point Likert scale (1: "definitely agree", 4: "neither agree nor disagree", 7: "definitely disagree"). If people had the intuition that the crucial conditions do not fully qualify as cases of lying, their ratings should concentrate around the midpoint (or below); if instead they shared the intuition that Coco is actually lying, the ratings should instead concentrate on the 'lie' end of the scale (5 or above).

4.5.2 Results and discussion

In the *no-intention condition*, participants overwhelmingly agreed that Coco did lie: 84% (21 out of 25) of them rated it as "5" or above, and only 16% (4 participants) as "4" or below (mean = 6.0; mode = "7"). Similarly, in the *no-belief condition*, participants overwhelmingly agreed that Coco did lie, 88% (22 out of 25) of them rating it as "5" or above and the remaining 12% (3 participants) rating it below "4" (mean = 5.8; mode = "7"). Like in the first experiment, the mean score in the 'repair belief'

condition was slightly lower than in the 'drink intention' condition, but the difference (unlike in Experiment 1) was not significant, $t(48) = 0.64$, $p = .523$. A comparison of the crucial and the straightforward cases revealed a significant difference between the two 'repair' scenarios, $t(34.3) = 2.52$, $p = .017$ (equal variances not assumed), but no significant difference between the two 'drink' scenarios, $t(49) = 0.58$, $p = .562$[23]. Overall, these results are well in accordance with those obtained with the dichotomous design, strengthening the case for the *entailed-assertion* account.



Tab. 2: Pie charts showing the participants' ratings in each of the four scenarios.

A last worry to address is that participants might have agreed that Coco lied only because they were not allowed to categorise his statement in some alternative way that they found more adequate. Perhaps most had the intuition that in the crucial conditions Coco was deceptive, or insincere, and were led to describe him as lying only because no other category of assessment was offered. They agreed that Coco was lying because it was the only available option to express that he misbehaved, but they would have denied it if some alternative category more adequate to describe the situation, like being deceptive, was also available. To address this worry, a second study was conducted that provided participants with an opportunity to describe the speaker along two different categories: as being deceptive and as being lying[24].

[23] Note, however, that this comparison may have been somewhat distorted by the fact that the control drink scenario got lower scores than expected from a straightforward case.
[24] What if the participants preferred to describe the utterance as insincere rather than deceptive? There is a reason why "deceptive" was preferred to "insincere". On all plausible understandings of these terms, being insincere entails being deceptive, while the opposite is not true. The "deception" option is thus preferable, as it allows all participants to acknowledge that the

4.6 Experiment 3

4.6.1 *Method*

Participants: Fifty-five new participants were tested (31 females, mean age (SD) = 34.6 years (12.5); range: 19–66; 100% reporting English as a native language, recruited and tested online as in the other experiments). Data from six participants failing to meet the minimum response time and/or the control question was excluded, but including them would have not affected the results.

Design: Participants were randomly assigned to one of the two crucial conditions (*no-belief, no-intention*) from Experiment 1. They were then asked to answer two questions that appeared (in randomized order) on the same screen:

- Did Coco say something deceptive? [Y/N]
- Did Coco tell Baba a lie? [Y/N]

Participants were then asked to answer the same control and demographic questions as in Experiment 1.

In both of the vignettes used in the experiment, there is no question that Coco's statement is deceptive: in the *no-belief* condition he pretends to have a belief he does not have; in the *no-intention* one, he pretends to have an intention that he does not have. Participants who had the intuition that Coco was *not lying* thus had the opportunity to deny that Coco is lying while being able to describe Coco as misbehaving, namely as being deceptive.

Given that in both vignettes it is uncontroversial that Coco's statement is deceptive, it is only the second question (about lying) that is of interest here. If the scores obtained in this experiment are significantly lower than those obtained in the same scenarios of Experiment 1, then the proposed interpretation of the results might be unwarranted. By contrast, if participants continue to describe Coco as lying, there is even stronger experimental evidence in favour of the proposed view.

4.6.2. Results

---

protagonist misbehaved also if they think that "insincere" is a more accurate description (agreeing that the protagonist is insincere entails agreeing that he is deceptive).

The overwhelming majority of participants marked both the *no-intention* case (92%, 24 out of 26) and the *no-belief* case (90%, 26 out of 29) as a lie. The scores obtained in the crucial conditions in Experiment 3 are even higher than those obtained in Experiment 1, but neither of the across-experiment differences was significant (*no-intention*: $\chi^2(1, N = 70) = 2.89$, $p = .089$.; *no-belief*: $\chi^2(1, N = 70) = .04$, $p = .84$). These results are consistent with the previous ones, and strengthen the case for the *entailed-assertion* view. Somewhat surprisingly, the difference between the *no-belief* and the *no-intention* condition found in the previous experiments has disappeared, $\chi^2(1, N = 55) = .12$, $p = .73$.

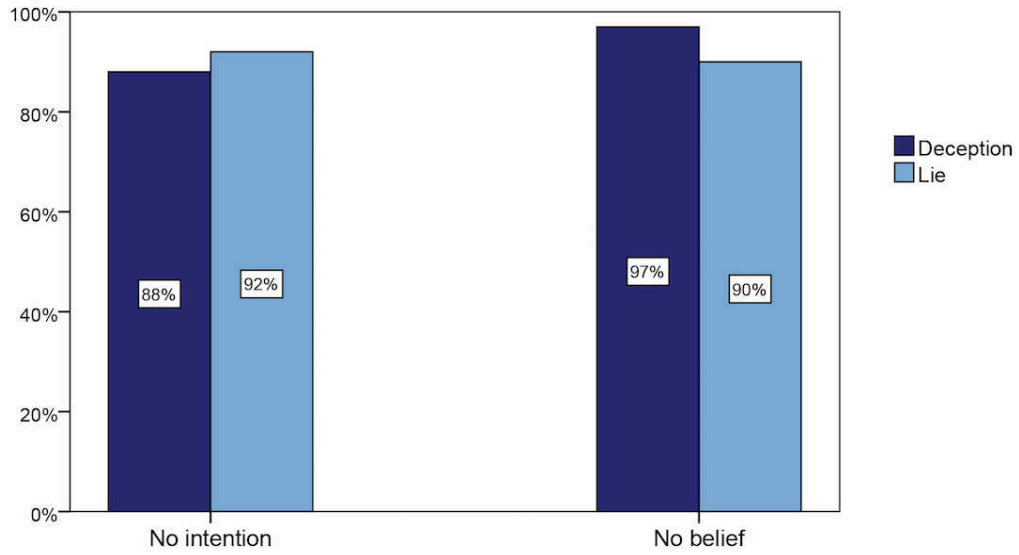


Figure 3: Percentage of respondents rating the protagonist's utterance as a lie and as deceptive.

4.7 General discussion

The results of all experiments strongly support the *entailed-assertion* account. They consistently show that it offers the best predictions of people's intuitions about lying by promising. A promise is a lie iff the speaker lacks belief or intention to fulfil the promise, or both. Furthermore, these results undermine all alternative views: the *only-assertion* view that only assertions can count as lies, the *only-belief* view that believing that what you say is false is necessary for lying, and the *only-intention* view that a promise is a lie only if the speaker does not intend to perform.

In Experiment 1 and 2, the crucial conditions obtained lower scores than the control conditions (Experiment 3 had no control conditions), and in Experiment 1 around one in six participants declared that they found it difficult to decide whether to classify them as

lies. In line with the prediction of the *entailed-assertion* view, this suggests that the consensus is less pronounced when only one of the two sincerity conditions for promising is violated, and that a portion of the the population sees them as non-paradigmatic instances of lying.

5 More on the results

5.1 Intentions vs Beliefs

The crucial conditions got different results in Experiment 1: the *no-intention* condition got significantly higher rates than the *no-belief* one. A slighter difference between them was also found in Experiment 2. This kind of difference, however, was not found in Experiment 3, or in the second question of Experiment 1 (about the participants' uneasiness to define Coco's utterance as a lie). Perhaps the difference is due to random fluctuations in the subjects' intuitions, and does not require an explanation. But if an explanation must be given, it can be given in terms of *moral judgements*.

Several authors have suggested that lying is a morally loaded term (Bok 1978:14, Williams 1985:140), and some of them even contended that white lies are not lies (Margolis 1962, Donagan 1977:89, Grotius, RWP:1212-8). A plausible view is that moral judgements might affect whether one finds a particular case of lying more or less prototypical. Experimental studies have also shown that judgements about intentions (Knobe, 2003) and causation (Alicke 2014) can be influenced by judgements about culpability (*i.e.* moral judgements). If moral judgements can affect folk intuitions about whether a particular case is a lie, a moral asymmetry between the scenarios might have influenced the results. In fact, there is such an asymmetry between the *no-intention* and the *no-belief* cases. In the first case, Coco is fully responsible for not fulfilling the promise: it is in his power to fulfil it, but he willingly decides to infringe it. By contrast, in the no-belief case Coco intends to do what is in his power to fulfil the promise, but it is not fully in his power to do so. In other words, in the first case he is fully responsible for the infraction, while in the second he is only responsible for having set the stakes too high.

While the no-intention case is by definition in 'bad faith', the no-belief case is by definition in 'good faith'. This could explain the slight asymmetry in the results – an asymmetry also expected in any replication of the test, given that it is built into the

difference between violating the 'belief sincerity condition' and violating the 'intention sincerity condition'.

5.2 The intent to deceive condition

Proponents of the *intention to deceive condition* (IDC) for lying (cf. section 1) might be worried that these results have been influenced by the fact that it has been left unspecified whether Coco intends to deceive Baba. For instance, the *no-belief* case might have received lower results in Experiment 1 because it is not clear whether Coco intends to deceive Baba. In this section, I will show not only that this worry is unfounded, but that the results undermine the very idea that ordinary speakers take the IDC to be necessary for lying.

On a standard interpretation, the IDC has to be the content of the statement: if the content of the statement is *p,* the speaker has to intend to make the hearer believe that *p.* On a weaker version of the IDC, the relevant intention is just to make the hearer believe that *the speaker believes* that *p.* In other words, where *p* is the *believed-false* content of the speaker's statement, these are the two versions of IDC:

IDC1: S intends A to believe: (*p*)
IDC2: S intends A to believe: (S believes that *p*)

In the *no-belief* condition, the content of the promise is that Coco will repair the car, a proposition that Coco believes to be very likely false. Now, unless Coco intends his promise not to be accepted, or not to be acted upon (*i.e.* if Coco's promise is a normal promise), Coco clearly intends Baba to believe that he will repair the car, so the participants have no reason to believe that IDC does not obtain.

In the *no-intention* condition, by contrast, the relevant content is that Coco will not drink. But here Coco believes the proposition to be probably true, so that neither IDC1 nor IDC2 can obtain, and the participants cannot think that they obtain. Two consequences can be drawn from this observation. The first one is that intuitions about the intention to deceive condition did not alter the result of the experiment: if they had some weight, they would have favoured the *no-belief* case over the *no-intention* one; instead, it was the latter that obtained significantly higher results. The second is that the high results from the *no-intention* case (90% lies, showing no significant difference from the straightforward cases) strongly suggest that neither IDC1 nor IDC2 is perceived as a necessary condition for lying.

Here is a reply. Even if neither IDC1 nor IDC2 are satisfied in the *no-intention* case, Coco is still aiming to deceive, since he clearly intends Baba to believe that he intends not to drink. The problem with this response is that it relies on a problematic definition of 'intention to deceive', according to which there are no constraints on what the deception is about:

IDC3: S intends to deceive A

However, IDC3 is untenable, because it counts any deceptive believed-false statement as a lie, even if deception has nothing to do with the content of the statement. It seems that the IDC should instead only capture deceptive intents that are somehow related to what is said by the liar. Fallis (2010:6) offers a convincing example of the counterintuitive consequences of endorsing IDC3: "Suppose that I say in a theatrical tone "I am the Prince of Denmark" in order to convince my new acquaintance that I am an actor (rather than that I am royalty). Although I have said something that I believe to be false with the intent to deceive, I have not lied".
To be able to defend the claim that Coco is lying in the *no-belief* case, the proponent of the IDC has to provide a different version of the IDC – perhaps one that is sensitive to the different attitudes expressed by a speech act, such as IDC4:

IDC4: S intends A to believe that (S $\Psi(p)$),

where Ψ is the propositional attitude expressed by the illocutionary act with content *p* that S has performed.
To conclude, the experiment represents a challenge to the existing versions of the IDC, and suggests a further challenge for their proponents: that of constraining the content of attempted deception in a way that generalises across different illocutionary acts.

5.2 Falsity condition

In the literature on lying, virtually every author accepts a '*subjective'* account of lying, according to which asserting an *objectively false* proposition is not necessary for lying, as long as the speaker *believes* that proposition to be false[25]. However, a few authors

[25] Augustine (DM, 3.3), Aquinas (ST, II-II, q.110, a1), Kant (1797), Leonard (1959:182), Isenberg (1964:466; 1974), Lindley (1971), Mannison (1969:138), Chisholm & Feehan (1977), Kupfer (1982:104), Adler (1997), Williams (2002), Mahon (2008), Fallis (2009). Other authors endorse weaker positions. Carson (1982:16; 2006:284; 2010:39) and Saul (2012) provide a

(Grotius *RWP*:1209, Benton forth) have suggested that falsity of the statement, in addition to belief in its falsity, is required for lying – call this the *objective* view. Recently, Turri & Turri (2015) claimed to have found experimental evidence that most laypeople endorse the objective view. However, their study is far from convincing, and has recently been dismissed by Wiegmann et al. (2016). An interesting aspect of the present study is that it puts further pressure on the objective view of lying, and on Turri & Turri's claims on the issue.

Turri & Turri conducted three experiments (A, B, C). Experiment A actually supports the *subjective* view; Turri & Turri challenge its results as possibly primed by perspective-taking or normative judgement. Experiment B and C support the *objective* view, but the probes used strike as disputable even to the eye of the non-expert. Experiment B tests the participants' intuitions about 'successfully lying', and hence can hardly be informative about the necessary conditions for lying: by definition, *successful lying* requires more than just *lying*, and falsity might just be one of such additional requirements (Kupfer 1982:104). Similarly, Experiment C forces participants to pick between "he tried to lie and actually did lie" and "he tried to lie but only thinks he lied" – thus presupposing the falsity of the subjective view, according to which such distinction is meaningless (on the subjective view, if you think you lied, you have lied). Wiegmann et al. (2016) offer a more systematic discussion of what is wrong with the experiment, and back up their critique with convincing experimental evidence. They argue that the results obtained in the second two experiments are due to "pragmatic re-interpretations [of the questions,] induced by the framing of the response options", and demonstrate that the preference for the subjective view disappears once such misleading framing is eliminated. All their experiments consistently show that the subjective view better tracks laypeople's intuitions about lying, thus dismissing Turri & Turri's discussion of their results and providing new evidence favouring the subjective view.

How does the present study relate to this debate? A first suggestion is that if participants regarded falsity as necessary for lying, they would not have rated a promise as a lie unless the promise was actually infringed in the story. However, in all conditions of all experiments participants agreed that the protagonist lied, even if it is always left

---

definition that does not require falsity, but both suggest that the definition could be strengthened to include such a requirement. Siegler (1966:132) suggests that falsity is necessary for telling a lie but not for lying; less controversially, Coleman & Kay (1981:28) argue that falsity is necessary for *prototypical* lying.

unspecified whether the promise was fulfilled or not, *i.e.* if it was objectively false[26]. This seems incompatible with the fact that all respondents in the straightforward conditions, and the majority in the others, indicated that Coco lied. Moreover, in the straightforward cases, 99% of the participants reported that they found it easy to respond. How could this be, if they did not know whether the falsity condition obtained?

One easy response is that in all four cases the respondents predicted that, given the information available, the promise would eventually be infringed in the story: they took it that it was 'implicit' that the falsity condition would obtain. This is clearly plausible for the *straightforward* conditions: since Coco intends not to do what he promised, and believes that very probably he will succeed in not doing it, the falsity condition will almost surely be met. A similar inference is plausible in the *no-belief* condition: even if Coco intends to repair the car, he believes that very likely he will not succeed, and this clearly suggests that the car will not be repaired.

The real problem for proponents of the falsity condition is the *no-intention* case: here, Coco intends to drink against his promise, but he believes that he will very likely fail to do so, because he will not be able to. The information provided in the scenario cannot license the inference to the conclusion that he will drink; as a matter of fact, it only licenses the opposite inference. In other words, not only it is not specified if the falsity condition obtains, but the scenario clearly suggests that it will not obtain. In all experiments, the *no-intention* condition was consistently rated as a lie (obtaining even higher scores than the other crucial condition), and in no experiment its results were significantly different to the straightforward scenarios: this strongly suggests that participants did not take falsity of the promise to be necessary for lying.

---

[26] What do we mean exactly by saying that such a promise *about a future state of affairs* can be false? No straightforward response can be offered in our case, because promises are about future contingents (at the moment in which a promise is uttered, it is still indeterminate whether the promisor will fulfil it: in some possible futures he does, in others he does not) and semanticists disagree about how to determine the truth-conditions of statements about future contingents. The fact that we are considering a promise rather than an assertion further complicates the issue. However, for the purposes of this paper it is sufficient to point out that no plausible account is able to predict that the falsity condition is met in the *no-intention* case. Having noted that, it is worth offering a sketch of how a plausible characterisation of the falsity condition for promising could look like. The following is broadly inspired on Belnap's (2000) account of the truth conditions for promising that Φ:

> *The falsity condition for promising that Φ is met at the moment of the utterance m_U and at the relevant moment m_R ($m_R \geq m_U$) iff it is a settled matter at m_R that Φ was false at m_U.*

I am grateful to Francesco Gallina for having helped me to develop this point.

The experimental data collected thus strengthens the case for the subjective account. As already shown by Wiegmann et al., people think that lying does not require falsity after all[27].

Conclusions

In this paper, I have developed a speech-act theoretic framework to outline the necessary and sufficient conditions to lie by promising, and sketched a way to extend this account to other illocutionary acts performed by uttering a sentence containing an explicit performative. Assuming that an 'assertion-based' definition of lying is broadly correct, the proposed account takes a promise to be a lie under the following conditions:

In uttering a promise with content p, S lies to A about p iff:
1. S successfully promises that p
2. Either S believes that not *p,* or S does not intend to *p,* or both

More generally, the account predicts that the utterance of a given speech act with content *p* is a lie iff that act illocutionarily entails an assertion and the speaker either satisfies the insincerity condition of that act, or that of the entailed assertion, or both.
In presenting this view, I have put into question several ideas that are found in the literature: the view that only direct assertions can be lies; the view that the propositional content of a lie has to be believed to be false; the view a promise is insincere only if the speaker does not intend to perform; the view that falsity is required for lying. The experimental study presented in the second part of the work has shown that with respect to promises the proposed view, but none of the alternative ones, gives predictions that are consistent with folk intuitions about lying.
Overall, it seems that the proposed theory of lying offers a promising background to develop a general theory of the conditions under which the performance of an insincere speech act counts as a lying.

Appendix I – the scenarios

Drink story - Refraining from action

[27] It should be stressed, however, that the experimental design of this study was not conceived specifically to test the falsity condition hypothesis. Even if the data collected supports the *objective* view, it provides weaker support than studies explicitly conceived to test this hypothesis, *i.e.* studies explicitly stated that the falsity condition did not obtain.

Baba and Coco are married. Coco is going to a party, but Baba does not like Coco to drink. So Baba asks him: "Will you be drinking alcohol at the party?". Coco replies: "Do not worry Baba: I promise that I will not drink alcohol at the party."

Scenario 1 - Straightforward:
In fact, Coco intends to drink alcohol at the party, and he is almost certain that he will find something to drink there.

Scenario 2 - No intention (to refrain as promised):
Coco actually intends to drink alcohol at the party and he will attempt to, but he is almost certain that he won't succeed, since he believes that the hosts do not offer alcoholic drinks at their parties.

Repair story - Positive action
Baba has broken her car, but she needs it to visit her family next week. For this reason, Baba has called Coco the mechanic to repair it. Coco the mechanic checks the car and tells Baba: "Do not worry Baba: I promise that I will repair your car by next week."

Scenario 3 - Straightforward:
Coco has no intention whatsoever to repair the car, and he is almost certain that he will not repair it.

Scenario 4 - No belief:
Coco intends to repair the car and he will attempt to do it, but he is almost certain that he won't manage to repair it in the end, because the damage is too serious.

Appendix II – Intentions and beliefs: good intentions and reasonable expectations

This appendix aims to clarify and justify two assumptions that I made in the paper. The first (section 3.2) is that it can be rational to intend to do something you believe you will probably fail to do, and conversely that it can be rational to believe you will probably end up doing something you do not intend to do. The second (section 4.4) is that the rationality constraints applying to mental states about *doing something* are different to the constraints applying to mental states about *refraining from doing something* – an asymmetry that is reflected in the experimental design.

The first assumption challenges the *cognitivist view* that *believing that one will successfully Φ* (not merely believing that one will *try to Φ*) is necessary for *rationally* intending to *Φ*:

COG: S rationally intends to *Φ* only if S believes that S will successfully *Φ*

If COG is right, the *entailed-assertion* account of the insincerity conditions for promising somehow boils down to the *only-intention* account (cf. section 3.2), as they would provide an

equivalent criterion for identifying sincere promises in all cases in which the speaker is rational[28].

There is a vast literature on the nature of intentions, and no consensus on whether COG is true. Rationally enough, defendants of COG are dubbed *cognitivists*, its opponents *non-cognitivist*. Cognitivists[29] endorse COG, and claim that it is irrational to intend to do something if you do not believe that you will do it: for instance, it is irrational for you to intend to arrive to work on time if you believe that you will fail to arrive to work on time. Non-cognitivists[30] deny COG, and hold that it can be rational to intend to do something even if you believe that there are significant chances that you will not do it: under some conditions, it might be rational for me to intend to arrive to work on time even if I do not believe that I will succeed.

This paper endorses a weak form of cognitivism. The aim of this Appendix is to clarify what this means, and to spell out the main assumptions of this view – by contrast, solving the long-lasting dispute between cognitivist and non-cognitivism falls well beyond its aim.

As anticipated in 3.2, it seems possible to intend to *Φ* even if you believe that you will probably fail in *Φ*ing, if your belief is *partial* rather than *outright*. Cognitivism is undeniable only as long as we move within a full-fledged belief framework (Holton 2008, 2009). Once credences (degrees of belief) are considered, rather than a binary opposition between believing and not believing it is possible to appreciate a number of intermediate cases: you may be *certain* that you will do what you intend to do, be *uncertain* that you will do it, believe it *probable* that you will do it, etc. In this framework, it is easy to spell out a plausible version of non-cognitivism.

I take it to be an uncontroversial 'cognitivist' assumption that one cannot rationally intend to do something if one is *certain* that one will not do it. I cannot rationally intend to arrive to work on time if I am certain that I will arrive late. So at least the following version of cognitivism seems undeniable:

W-COG: S rationally intends to *Φ* only if S is not certain that S will not successfully *Φ*

[28] It should be noted, however, that COG departs significantly from the rationality constraints traditionally imposed on the *only-intention* condition for promising. Writing on promises, Searle explicitly states that an intention to *Φ* requires believing that *Φ is possible* (Searle 1965:243), but does not mention any further requirements. One can adopt COG and interpret the *only-intention* account as indirectly requiring a sincere belief, but such an account would be significantly different to (and arguably at odds with) traditional speech-act theoretic views.

[29] Hampshire & Hart (1958:1-12), Grice (1971:263-79), Audi (1973); Harman (1976:461-463), (1989:§8), Davis (1984), Velleman (1985:33-61; 1989), Setiya (2008:388-409), Ross 2009:238-282).

[30] Davidson (1971:50; 1978:91-4; 1980:83-102), Bratman (1987:19-20,37-39; 2009: §2), Mele (1992: §8), Holton (2008:51-55; 2009: §2), Hieronymi (2009:201-20), Paul (2009:1-24), Sinhababu (2013).

This form of is cognitivism only specifies the incompatibility between certainty that you will not succeed to *Φ* and intention to *Φ*. If one accepts W-COG alone as a criterion for rationally intending, one obtains W-COG*:

W-COG*: S can rationally intend to *Φ* only if S has at least a partial belief that S will successfully *Φ*

A similar view is defended by Holton (2008:51-55, 2009:49), who claims that you rationally intend to *Φ* only if you *take as a live possibility* that you will *Φ* successfully. To put our example in his words, it is rational for me to intend to arrive to work on time as long as I take it as a live possibility that I will arrive to work on time. Similarly, in the 'Repair scenario' employed in the experiments, Coco takes it as a live possibility that he will repair the car (he is not certain that he will fail to repair it), so his intention can be rational. Adopting a rationality constraint like W-COG* allows you to rationally promise what you intend to do, even if you do not fully believe that you will succeed.

Conversely, are there intention-based constraints on believing that one will *Φ*? Is it rational to believe that you will *Φ* if you do not intend to *Φ*? If the answer is negative, maybe the *no-belief* account can be redeemed, as it would give the same predictions of the *no-intention* paradigm in all cases in which the speaker is rational.

At first glance, believing that one will *Φ* does not require intending to *Φ*. For instance, I may believe that I will go to jail, while intending not to go to jail, if I think that the police will transport me there in a police van and lock me up. However, the example is misleading, as *Φ* here does not indicate an intentional action that is willingly brought about by the speaker, and this is not what we want to capture with the expression "intending to *Φ"*. We are interested in the rationality of intending to *act*, and acting requires *actively* doing something, whereas being passively subject to someone else's action does not qualify as acting.

When I am escorted to jail, I am not actively performing the action of going to jail: I am rather passively being brought there. By contrast, if I go to jail against my will because I fear the legal consequences of not doing so, I am actively performing the action of going to jail: here I lack the desire to go to jail, but not the intention to do it. The jail examples fail to show that you can rationally believe that you will *Φ* without intending to *Φ*, as long as *Φ* is taken to be an intentional action. They suggest that believing that one will *Φ* does not require *desiring* to *Φ,* but does indeed require *intending* to *Φ.* This results in the converse of the cognitivist view, applied to beliefs:

COG2: S rationally believes that [S will intentionally] *Φ* only if S intends to *Φ*

Putative counterexamples to COG2 are cases in which *Φ* stands for a verb that indicates that S refrains from acting – what I will call a *negative action*. Suppose I am on an only-meat diet, but I love vegetables. I might believe that I will stick to my only-meat diet and refrain from eating

vegetables today, while intending not to refrain from eating vegetables, because I am almost certain that I will have no way to get some vegetables to eat. Here the case is analogous to the examples supporting *weak cognitivism* for intentions: it has to do with an intention that cannot be achieved because of some conditions (vegetables are not available) that are very likely to occur independently of the agent, and that is rational only insofar as the agent is not certain that these circumstances will occur (cf. Sinhababu 2013).

There is a reason for this analogy, and more generally a reason why a 'negative action' can violate COG2 while a 'positive action' cannot. The reason is that believing that you will refrain from Φ even if you intend not to refrain from Φ can be redescribed as believing that you will not Φ while intending to Φ. That is, a case of *believing without intending* that you will perform a *negative action* can be redescribed a case of *intending without believing* that you will perform a positive action. To see that this is the case, consider again the example:

(Br) I believe that I will refrain from eating vegetables
is equivalent to
(B~p) I believe that I will not eat vegetables

(I~r) I intend not to refrain from eating vegetables
is equivalent to
(Ip) I intend to eat vegetables

Or, formally, where ('r' = 'refrain from p'):
B(r) = B(~p)
I(~r) = Ip

(Here (Br)/(I~r) and (B~p)/(Ip) are the pairs of mental states whose rationality is under discussion)

Now, if the conversion from intending without believing into believing without intending is based on equivalence, it can also go the other way round, and convert the description of an intention-belief pair with a negative-action content into a pair that has a positive-action content. This would mean that also positive action cases can be described as cases of intending without believing and treated under W-COG*, so that W-COG* can also apply to cases of believing without intending. The problem is that once we consider intention-belief pairs with negative content, it turns out that W-COG* does not apply to these pairs – in other words, that negative action cases are a counterexample to W-COG*. To see this, consider the opposite example:

(2Bp) I believe that I will almost certainly eat
is equivalent to
(2B~r) I believe that I will almost certainly not refrain from eating

(2I~p) I intend to not eat
is equivalent to
(2Ir) I intend to refrain from eating

(Here (2Bp)/(2I~p) and (2B~r)/(2Ir) are the pairs of mental states whose rationality is under discussion)

That holding both mental states ((2Bp) and (2I~p)) would involve irrationality is quite obvious, as predicted by COG2: whenever I intend to refrain from willingly doing something, it is entirely up to me whether I will not do that thing, so that I cannot believe that I will not refrain if I intend to refrain.

The moral of this story is that, for all intention-belief pairs, for all intentional actions Φ, *weak cognitivism* (W-COG*) constrains the rationality of one's intention to perform a positive action (or one's belief that one will perform a negative action), and *strong cognitivism* (COG-COG2) *) constrains the rationality of one's intention to perform a negative action (or one's belief that one will perform a positive action). In other words, if Φ is a *positive action*, it is rational for you to intend to Φ even if you are not certain that you will Φ, but not to believe (to some degree) that you will Φ if you intend not to Φ. By contrast, if Φ is a *negative action*, it is irrational for you to intend to Φ even if you are not certain that you will Φ, but it is rational to believe that you will (less than certainly) Φ if you intend not to Φ.

This provides a unified picture of what it is to rationally intend without believing, and to rationally believe without intending. From this picture, it follows that a rational speaker can insincerely promise to actively do something by violating only the *belief sincerity condition*, and insincerely promise to refrain from doing something by violating only the *intention sincerity condition*. This is a fundamental assumption for this paper, as it explains both the importance of the distinction between the *belief-insincerity* and the *intention-insincerity* conditions (sections 3.2) and the asymmetry in the experimental design (section 4) [31].

[31] I am grateful to Miklos Kurthy, Jennifer Saul, Manuel García-Carpintero, Mark Jary, Francesco Gallina, and two anonymous reviewers for their helpful comments.

References

Adler, Jonathan E. 1997. Lying, Deceiving, or Falsely Implicating. *Journal of Philosophy.* 94: 435-452.

Alicke, Mark, David Rose & Dori Bloom. 2014. Causation, norm violation, and culpable control. In J. Knobe & S. Nichols (eds.). *Experimental Philosophy* (vol. 2), 229-250. New York: Oxford University Press.

Alston, William P. 2000. *Illocutionary Acts and Sentence Meaning.* Ithaca: Cornell University Press.

Anscombe, Elizabeth. 1981. *Ethics, Religion and Politics*. Oxford: Blackwell.

Aquinas. [ST]. *Summa Theologiae.*

Arico, Adam J., & Fallis, Don. 2013. Lies, Damned Lies, and Statistics: An Empirical Investigation of the Concept of Lying. *Philosophical Psychology* 26 (6), 790-816.

Audi, Robert. 1973. Intending. *Journal of Philosophy* 70: 387–403.

Augustine. [DM]. *De Mendacio*. (*On lying)*.

Austin, John L. 1961/2003. Performative Utterances. In J.L. Austin, *Philosophical Papers*. Oxford: Oxford University Press, 3rd edn.

Austin, John L. 1962/1975. *How to Do Things with Words,* Oxford: Oxford University Press, 2nd edn.

Bach, Kent and Robert M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, Mass.: MIT Press.

Bach, Kent. 1975. Performatives are statements too. *Philosophical Studies* 28: 229-236.

Bazzanella, Carla. 2009. Approssimazioni pragmatiche e testuali alla menzogna. In F. Venier (ed.) *Tra Pragmatica e Linguistica testuale. Ricordando Maria-Elisabeth Conte,* 67–90. Alessandria: Dell'Orso.

Belnap, Nuel D. 2000. Double Time References: Speech-act Reports as Modalities in an Indeterminist Setting. *Advances in modal logic* 3: 37-58.

Benton, Matthew. forthcoming. Lying, Belief, and Knowledge. In J. Meibauer, *The Oxford Handbook on Lying*. Oxford: Oxford University Press.

Bok, Sissela. 1978. *Lying: Moral Choice in Public and Private Life*. New York NY: Random House.

Bratman, Michael E. 1987. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press

Bratman, Michael E. 2009. Intention, Belief, Practical, Theoretical. In S. Robertson (ed.) *Spheres of Reason: New Essays in the Philosophy of Normativity*. Oxford: Oxford University Press.

Carson, Thomas L. 1988. On the Definition of Lying: A reply to Jones and revisions. *Journal of Business Ethics* 7:509-514.

Carson, Thomas L. 2006. The definition of lying. *Nous* 2: 284–306.

Carson, Thomas L. 2010. *Lying and Deception*. Oxford: Oxford University Press.

Carson, Thomas L., Richard E Wokutch and Kent F Murrmann. 1982. Bluffing in Labor Negotiations: Legal and Ethical Issues. *Journal of Business Ethics* 1(1): 13–22.

Chan, Timothy and Guy Kahane. 2011. The trouble with being sincere. *Canadian Journal of Philosophy* 41(2): 215–234.

Chen, Rong, Chunmei Hu and Lin He. 2013. Lying between English and Chinese: An Intercultural Comparative Study. *Intercultural Pragmatics* 10.3: 375-401.

Chisholm, Roderick M. and Thomas D. Feehan. 1977. The intent to deceive. *Journal of Philosophy* 74(3): 143–159.

Cole, Shirley A. N. 1996. Semantic Prototypes and the Pragmatics of *Lie* Across Cultures. *The LACUS Forum* 23:475-83.

Coleman, Linda and Paul Kay. 1981. Prototype Semantics: The English Verb 'lie'. *Language* 57: 26-44.

Davidson, Donald. 1971. Agency. Reprinted in *Essays on Actions and Events*, 43–61. Oxford: Oxford University Press.

Davidson, Donald. 1985. Deception and Division. In J. Elster (ed.), *The Multiple Self*. Cambridge: Cambridge University Press.

Davidson, Donald. 1978. Intending. Reprinted in *Essays on Actions and Events*, 83–102. Oxford: Oxford University Press.

Davis, Wayne A. 1984. A Causal Theory of Intending. *American Philosophical Quarterly* 21: 43–54.

Davis, Wayne A. 1999. Communicating, Telling and Informing. *Philosophical Inquiry* 21(1): 21–43.

Davis, Wayne A. 2003. *Meaning, Expression and Thought*. Cambridge: Cambridge University Press.

Donagan, Alan. 1977. *A Theory of Morality*. Chicago: Chicago University Press.

Eriksson, John. 2011. Straight Talk: Conceptions of Sincerity in Speech. *Philosophical Studies* 153 (2): 213–234.

Falkenberg, Gabriel. 1988. Insincerity and disloyalty. *Argumentation* 2.1: 89-97.

Fallis, Don. 2009. What is lying? *The Journal of Philosophy* 106 (1): 29–56.

Fallis, Don. 2010. Lying and Deception. *Philosophers' Imprint* 10 (11).

Fallis, Don. 2012. Lying as a Violation of Grice's First Maxim of Quality. *Dialectica* 66 (4): 563–581.

Fallis, Don 2013. Davidson was Almost Right about Lying. *Australasian Journal of Philosophy* 91 (2): 337–353.

Fallis, Don. 2014. Are Bald-Faced Lies Deceptive after All? *Ratio* 28 (1): 81-96.

Faulkner, Paul. 2007. What is Wrong with Lying? *Philosophy and Phenomenological Research* 75: 524-547.

Foot, Philippa. 2001. *Natural Goodness*. Oxford: Oxford University Press.

Fried, Charles. 1981. *Contract as Promise*. Cambridge, MA: Harvard University Press.

Garcìa-Carpintero, Manuel. 2013. Explicit Performatives Revisited. *Journal of Pragmatics* 49: 1–17.

Ginet, Carl, 1979. Performativity. *Linguistics & Philosophy* 3: 245-265.

Green, Mitchell S. 2000. Illocutionary Force and Semantic Content. *Linguistics and Philosophy* 23: 435–473.

Green, Mitchell S. 2007. *Self-Expression*. Oxford: Oxford University Press.

Grice, H. Paul. 1971. Intention and Uncertainty. *Proceedings of the British Academy* 57: 263–79.

Grice, H. Paul. 1989. *Studies in the Way of Words.* Cambridge: Harvard University Press.

Grotius, Hugo. [RWP]. *The Rights of War and Peace*. F. W. Kelsey (trans.), Indianapolis: Bobbs-Merrill.

Hampshire, Stuart and Herbert L. A. Hart. 1958. Decision, Intention and Certainty. *Mind* 67.

Hare, Richard. 1952. *The language of morals*. Oxford: Oxford University Press.

Hardin, Karol J. forthcoming. Linguistic Approaches to Lying and Deception. In Meibauer, J. (ed.) *The Oxford Handbook on Lying*. Oxford: Oxford University Press.

Harman, Gilbert. 1976. Practical Reasoning. *Review of Metaphysics* 29: 431–463.

Harman, Gilbert. 1986. *Change in View: Principles of Reasoning.* Cambridge, MA: The MIT Press.

Harris, Roy. 1978. The descriptive interpretation of performative utterances. *Journal of Linguistics* 14(2): 309-310.

Hedenius, Ingemar. 1963. Performatives. *Theoria* 29: 115-36.

Hieronymi, Pamela. 2009. The Will as Reason. *Philosophical Perspectives* 23: 201–220.

Holton, Richard. 2008. Partial Belief, Partial Intention. *Mind* 117: 27–58.

Holton, Richard. 2009. *Willing, Wanting, Waiting.* Oxford: Oxford University Press.

Hume, David. [THN]. *A treatise on human nature.* NY: Oxford University Press.

Isenberg, Arnold. 1964. Deontology and the Ethics of Lying. *Philosophy and Phenomenological Research* 24: 463-480.

Jary, Mark. 2007. Are explicit performatives assertions? *Linguistics and Philosophy* 30(2): 207 - 234.

Kant, Immanuel. 1797. On a supposed right to lie from philanthropy. In Gregor, Mary J. (ed.) - *The Cambridge Edition of the Works of Immanuel Kant- Practical Philosophy*, 605-616. Cambridge: Cambridge University Press.

Knobe, Joshua. 2003. Intentional action and side effects in ordinary language. *Analysis* 279: 190–194.

Kupfer, Joseph. 1982. The moral presumption against lying. *The Review of Metaphysics* 36(1): 103–126.

Lemmon, J.E. 1962. Sentences verifiable by their use. *Analysis* 12: 86-89.

Leonard, Henry S. 1959. Interrogatives, Imperatives, Truth, Falsity and Lies. *Philosophy of Science* 26: 172-186.

Lewis, David. 1970. General Semantics. *Synthese* 22: 18-67

Lindley, T. Foster. 1971. Lying and Falsity. *Australasian Journal of Philosophy* 49: 152–157.

MacFarlane, John. 2011. What is Assertion? In J. Brown and H. Cappelen (eds.) *Assertion: New Philosophical Essays*, 79-96. Oxford: Oxford University Press.

MacIntyre, Alasdair. 1995. Truthfulness, Lies, and Moral Philosophers: What Can We Learn from Mill and Kant?, in *The Tanner Lectures on Human Values*, Salt Lake City: University of Utah Press, 16: 307–361.

Mahon, James E. 2008. Two Definitions of Lying. *International Journal of Applied Philosophy*, 211–230.

Mahon, James E. 2010. The definition of lying, and why it matters. Roundtable at the School of Info Resources and Library Sciences, https://itunesu.itunes.apple.com/feed/id413143120.

Mannison, Don S. 1969. Lying and lies. *Australasian Journal of Philosophy* 2(47): 132–144.

Margolis, Joseph. 1962. "Lying Is Wrong" and "Lying Is Not Always Wrong". *Philosophy and Phenomenological Research* 23: 414–418.

Marsili, Neri. 2014. Lying as a Scalar Phenomenon: Insincerity along the Certainty-Uncertainty Continuum. In S. Cantarini, W. Abraham, and E. Leiss (eds.). *Certainty-Uncertainty – and the Attitudinal Space in between*, 153–173. Amsterdam: John Benjamins Publishing Company.

Marsili, Neri. 2015. Normative accounts of assertion: from Peirce to Williamson, and back again. *Rivista Italiana di Filosofia del Linguaggio* (2).

Marsili, Neri. forthcoming. Lying and Certainty. In J.Meibauer, *The Oxford Handbook on Lying*. Oxford: Oxford University Press.

Marušić, Berislav. 2012. Belief and difficult action. *Philosopher's Imprint 12*(18).

Marušić, Berislav. 2013. Promising against the Evidence. *Ethics* 123(2): 292–317.

Meibauer, Jörg . 2005. Lying and Falsely Implicating. *Journal of Pragmatics* 37(9): 1373–1399.

Meibauer, Jörg. 2014. *Lying at the semantic-pragmatic interface*. Berlin: De Gruyter.

Mele, Alfred. 1992. *Springs of Action.* Oxford: Oxford University Press.

Moore, George. 1993. *Selected Essays*. London: Routledge.

Moran, Richard. 2005. Problems with sincerity. *Proceedings of the Aristotelian Society* 105: 325–345

Owens, David. 2008. Promising without Intending. *The Journal of Philosophy* 105(12): 737–755.

Paul, Sarah K. 2009. How We Know What We're Doing. *Philosophers' Imprint* 9.

Rawls, John. 1981. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Reimer, Marga. 1995. Performative utterances: a reply to Bach and Harnish. *Linguistics and Philosophy* 18: 655-675.

Reimer, Marga. 2004. What Malapropisms Mean: A Reply to Donald Davidson. *Erkenntnis* 60 (3).

Ridge, Michael. 2006. Sincerity and expressivism. *Philosophical Studies* 131: 487–510.

Ross, Jacob. 2009. How to be a Cognitivist about Practical Reason. In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics*. 4: 243–282.

Saul, Jennifer M. 2011. Just go ahead and lie. *Analysis* 72(1): 3–9.

Saul, Jennifer M. 2012. *Lying, Misleading, and what is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford: Oxford University Press.

Searle, John R. 1965. What is a speech act? In Black, M.(ed.) *Philosophy in America*, 221-39. London, Allen & Unwin.

Searle, John R. 1969. *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, John R. 1976. A classification of illocutionary acts. *Language in Society* 5: 1-23.

Searle, John R. 1989. How performatives work. *Linguistics & Philosophy* 12:535-558.

Searle, John R and Daniel Vanderveken. 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Setiya, Kieran. 2008. Practical Knowledge. *Ethics* 118: 388–409.

Sidgwick, Henry. 1981 [1874]. *The Methods of Ethics*. Indianapolis. IN: Hackett.

Siegler, Frederick A. 1966. Lying. *American Philosophical Quarterly* 3: 128-136.

Sinhababu, Neil. 2013. The Desire-Belief Account of Intention Explains Everything. *Noûs* 47(4): 680–696.

Simpson, David. 1992. Lying, Liars and Language, in *Philosophy and Phenomenological Research* 52: 623-639.

Sorensen, Roy. 2010. Knowledge-Lies. *Analysis* 70(4): 608–615..

Sorensen, Roy. 2011. What lies behind misspeaking. *American Philosophical Quarterly* 48(4): 399-409.

Stalnaker, Robert. 1978. Assertion. In R. Stalnaker, *Context and content,* 78–95. Oxford: Oxford University Press.

Stalnaker, Robert. 1998. On the representation of context. In R. Stalnaker, *Context and content,* 96–114. Oxford: Oxford University Press.

Stokke, Andreas. 2013. Lying and asserting. *Journal of Philosophy* 110(1): 33–60.

Stokke, Andreas. 2014. Insincerity. *Noûs* 48(3): 496–520.

Sweetser, Eve. 1987. The definition of lie: an examination of the folk models underlying a semantic prototype, in D. Holland and Q. Naomi (eds.), *Cultural Models in Language and Thought*, 3-66. Cambridge: Cambridge University Press.

Turri, Angelo and John Turri. 2015. The Truth about Lying *Cognition.* 138: 161–168.

Vanderveken, Daniel. 1980. Illocutionary logic and self-defeating speech acts. In J.R. Searle, F. Kiefer and M. Bierwisch, *Speech act theory and pragmatics,* 247-272. Amsterdam: Springer.

Velleman, J. David. 1985. Practical Reflection. *The Philosophical Review* 94:33–61.

Velleman, J. David. 1989. *Practical Reflection*. Stanford: CSLI Publications.

Watson, Gary. 2003. Asserting and Promising. *Philosophical Studies* 117(1): 57–77.

Wiegmann, Alex, Jana Samland & Michael R. Waldmann, 2016. Lying despite telling the truth. *Cognition* 150: 37-42.

Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Williams, Bernard. 2002. *Truth and Truthfulness: An Essay in Genealogy*. Princeton NJ: Princeton University Press.

Wright, Crispin. 1992. *Truth and objectivity*. Cambridge, MA: Harvard University Press.

Xu, Fen, Yang C. Luo, Genyue Fu, and Kang Lee. 2009. Children's and adults' conceptualization and evaluation of lying and truth-telling. *Infant and Child Development* 18(4): 307–322.

Bio Note

Neri Marsili (MA, Turin) is Phd candidate and Teaching Fellow at the University of Sheffield. His main research interests are in pragmatics, epistemology of testimony and aesthetics. He has published in Italian and English on assertion and lying.