Morality and the Folk Concept of Lying

Anna Martinelli-Parker

Advisors: Joshua Knobe and Laurence Horn

Submitted to the faculty of the Department of Linguistics
in partial fulfillment of the requirements for the degree of Bachelor of Arts

Department of Linguistics
Yale University

April 26th, 2024

Abstract

Is moral negativity intrinsic to the concept of a lie? This seems intuitively correct. However, we still often treat morally positive deceptions as lies. In this essay, I investigate whether and how moral valence impacts lie judgments, and thus whether moral negativity forms part of the definition of a lie. I conducted a study with a 2x2 variable format, with moral valence (positive/negative) and inference type (false assertion/implicature) as variables. Participants read a vignette involving a deceptive utterance with a combination of the variables and then gave their lie judgments for the utterance. Using a linear mixed effects model, I found that moral valence has a large effect on lie judgments. A morally negative deceptive utterance is thus much more likely to be perceived as a lie than an otherwise identical positive one. However, I did not find an interaction effect between moral valence and inference type. Additionally, participant responses to morally positive deceptive utterances had a large variance, indicating a lot of confusion about how to treat these. These seemingly contradictory results lead to the conclusion that lies likely function as dual character concepts, with moral negativity being a component of the secondary level of the definition, centered on fulfillment of abstract values, and not of the first level of the definition, centered on concrete features. On the other hand, untruthfulness is also part of the definition, but it is part of the first, concrete level.

Table of contents

Acknowledgments

I extend a huge thank you to my advisors, Josh Knobe and Larry Horn, for their extensive support throughout the writing of this thesis. I'm very grateful for the patience and kindness that characterized their approach to my learning.

I'm also grateful to Claire Bowern, Raffaella Zanuttini, and my fellow students in Ling 491, who provided both emotional and writing support when it was sorely needed.

I would like to thank my mom, dad, and brother for being both a central source of inspiration and my long-distance cheering squad.

I dedicate this essay to my grandfather, Bob, who passed away on April 21st. Love you, Grandpa.

Introduction

Imagine the following scenario:

(1) Chris wants his friend Derek to come with him to his apartment. To convince him to come, he tells Derek, "I want to go to my apartment because I need to pick up my jacket." Chris actually wants to go to his apartment because he's throwing a surprise party for Derek there.

Now imagine an alternative version of the same scenario:

(2) Liam wants his friend Jack to come with him to his apartment. To convince him to come, he tells Jack, "I want to go to my apartment because I need to pick up my jacket." Liam actually wants to go to his apartment because he's planning to ambush Jack there in order to steal his money.

In each of these scenarios, the statement made by the speaker is the same. Both statement are untruthful, and spoken with the intent to deceive the speaker's addressee. In short, we could see both as straightforward cases of lying. But Chris arguably didn't do anything *wrong*, when Liam certainly did. Is it justifiable to treat Chris's utterance as a lie just like Liam's?

They key question at issue here is whether moral status has a role in defining what a lie is, and what statements count as lies. On the one hand, we often seem to treat lies as having an inherently negative moral status. To call someone a liar is a very negatively charged

insult (Saul 2012). When someone lies with a good reason for doing so, we describe the lie as 'justified', implying that lies arise from a position of needing to be morally justified in order for their use to be acceptable. Much of the literature on lies assumes that lies are inherently morally negative.

On the other hand, if we look at scenarios (1) and (2), we could argue that the moral status of the two utterances is incidental to the question of whether they constitute lies in the first place. Perhaps some lies are just morally good, and some lies are just morally bad. We still often seem to perceive well-intentioned lies, or white lies, as *lies*. A classic example of what is often considered a morally good lie is a parent telling their child that Santa Claus is real. Perhaps the association of lies with moral negativity arises not because negativity is inherent in the concept of a lie, but rather by association— lies are forms of deception, and deceptive actions tend to be morally negative.

The relationship between morality and the definition of a lie is thus not a straightforward one. The purpose of this senior essay is to clarify this relationship by investigating the role of morality in shaping the folk concept of lying.

1.1 The folk concept of lying

This essay takes an experimental approach to understanding the definition of a lie. This approach is guided by the agreement in the literature on lies that a good definition of a lie needs to capture the folk concept of lying. The folk, or ordinary, concept of lying is the definition of lying held intuitively by ordinary people and used in ordinary situations.

Conceptions of lying arise primarily from people's actual social experiences rather from an academic or intellectual understanding of lies (Arico and Fallis 2013-01-15; Thomas 2010). An experimental approach is thus particularly well-positioned to provide insights into the topic of lying, as it focuses by design on capturing ordinary intuitions. This approach to the definition of a lie serves in contrast to definitions arrived at primarily by philosophical

reasoning, which was the case for most of the literature on the definition of a lie up until 1981, when the first experimental study focused on capturing the definition of a lie was published.

1.2 Central questions

The primary question of this senior essay is whether and how the moral valence of an utterance impacts the folk concept of lying. The investigation involves an application of experimental methods to examine whether the perceived moral valence of deceptive statements (specifically of the sort common to everyday life) has an impact on lie judgments, these being judgments about whether the statement is a lie or not. Are people more likely to judge assertions to be lies when they are morally negative? And are they less likely to judge them to be lies when they are morally positive? To what degree is moral negativity something associated with lies rather than something that's inherent to them?

As a secondary focus, this essay also explores whether moral negativity increases participant lie judgments for deceptive implicatures, as opposed to outright false assertions. The question of whether it is possible to lie without telling a falsehood, and thus whether deceptive implicatures can be lies, is a hotly debated one in the literature on lies. Results of previous studies are very split. Weissman and Terkourafi (2018), for example, found that study participants did not regard deceptive implicatures as lies, but Antomo et al.(2018) found that, though lie judgments were smaller for deceptive implicatures than they were for false assertions, participants did rate deceptive implicatures as above the midpoint on the scale they provided for lie-judgments.

One interpretation of these differing results is that some deceptive implicatures are more lie-like than others, although what exactly it means to be 'more lie-like' isn't itself clear. Theoretically, if moral negativity is intrinsic in the concept of a lie, then moral negativity should increase participants' lie judgments (by making them 'more lie-like') until there is

less of a separation between lie judgments for false assertions and for deceptive implicatures.

Thus, if there is an interaction effect between the moral valence and inference type (whether an utterance is false or true with deceptive implicatures) of deceptive statements, that is an indication that moral negativity is intrinsic in the concept of a lie. The second central question of this study is thus whether moral negativity makes deceptive implicatures more likely to be perceived as lies.

This study thus aims to simultaneously measure how moral valence (whether a deceptive utterance is morally good or morally bad) impacts lie judgments, how inference type (whether the utterance involves a false assertion or a deceptive implicature) and whether there is an interaction effect between these two variables.

1.3 Lying and morality

Although there is an extensive literature that looks into the relation between lying and morality, almost all of it approaches the question from the normative or ethical perspective that asks whether and when it is morally justified to lie. This question necessarily assumes that the definition of a lie is something which has already been established. This senior essay seeks to reverse the question to make it definitional rather than normative, asking whether morality might itself be a component of the definition of a lie. This essay seeks to fill the gap in the literature that exists at this intersection of the definition of a lie and morality.

In addition to its academic importance for linguistic literature, the question of whether morality is inherent in the definition of a lie, as well as the definition of a lie more broadly, has great relevance to our daily lives. We consistently interact with lies and deception as part of our social existence. They play an integral role in human communication. Establishing a stronger understanding of the relationship between lies and morality, and of what makes a statement a lie, is essential to understanding when someone can (and should) be held morally responsible for lying or deceiving another person.

The question is also significant for topics like lying and deception in the court of law. A better understanding of the folk concept of a lie can help prevent it from interfering with applications of the legal concept of perjury. This is especially true if moral considerations do in fact influence when an utterance is considered a lie in the folk concept of lying. Questions about whether someone committed perjury when testifying should be kept separate from moral judgments of that person's character. A better understanding of the definition of a lie is thus essential to help prevent misapplications of the law.

Literature Review

2.1 The definition of a lie

This senior essay centers on the question of the definition of a lie. This is a topic that has been discussed in philosophical and linguistic literature for hundreds of years without a consensus being reached. However, most traditional views about the definition of a lie share the same conceptual structure: they view lying as the fulfillment of a series of specific conditions, with the central question (to which responses differ) being exactly which conditions are necessary or sufficient in order for an utterance to be defined as a lie.

2.1.1 Condition-based accounts of lying

Untruthfulness Condition

The first of these potential conditions is the falsity of the utterance, or the untruthfulness condition. There are two variations of this condition, which are sometimes treated as separate conditions entirely: objective untruthfulness and subjective untruthfulness. Utterances are objectively untruthful when the truth conditions of the utterance fail to uptake. Put simply, it is when the utterance is actually false. On the other hand, utterances are subjectively untruthful when the speaker of the utterance *believes* that it is false, although it may or may not actually be objectively false (Faulkner 2013; Horn 2017; Wiegmann and Meibauer

2019-07-31).

Although objective and subjective falsity are often aligned, that isn't always the case—there are situations where the speaker might think their utterance is true when it is actually false, or think that their utterance is false when it is actually true.

(3) Lucy asks Natalie what the weather's like. Natalie checks her weather app, which says it is raining, and tells Lucy, "It's raining." As it turns out, the weather app was glitching, and it is actually sunny outside.

If subjective falsity is a necessary condition of lying, Lucy's utterance in (3) is not an instance of a lie, because Lucy did not believe that she was stating something untrue. However, if objective falsity is a necessary condition of lying, then Lucy's utterance is an instance of a lie, because she said something that was actually false.

There is general (although notably not universal) agreement in the literature that the speaker's belief about the truthfulness of their utterance is more important for determining whether the utterance is a lie than the actual truth conditional status of the utterance.(Faulkner 2013; Mahon 2016-10-10; Wiegmann 2023-08) This would mean that only subjective falsity is relevant for the untruthfulness condition.

An early proponent of this view is St. Thomas Aquinas, in his text *Summa Theologica*, which he composed between 1265 and his death in 1274. Aquinas writes that "[I]f one says what is false, thinking it to be true, it is false materially, but not formally, because the falseness is beside the intention of the speaker so that it is not a perfect lie, since what is beside the speaker's intention is accidental", and if "one utters falsehood formally, through having the will to deceive, even if what one says be true, yet inasmuch as this is a voluntary and moral act, it contains falseness essentially and truth accidentally, and attains the specific nature of a lie." (II.II, q. 110). Aquinas thus argues that there is an 'essential' falseness which emerges when the speaker intends for their claim to be false, and that this essential falseness is necessary for an utterance to have the 'specific nature of a lie'. (Horn 2017)

Mannison (1969) compares a case where a speaker intends to deceive with a false statement, with another case where a speaker intends to deceive with a statement that they believe to be false but is not actually so. He argues, on the basis of this comparison that the actual falsity of an utterance should not be a condition for lying because the two cases are not sufficiently different acts. In both cases, 1) the speaker can be held responsible for not speaking candidly, 2) the intended outcome is exactly the same, and 3) whether the deception is successful is irrelevant, and so the "existence of an act of lying is not dependent upon the production of a particular response or state in an addressee" (135). Mannison also argues that requiring objective falsehood as a condition of lying produces unintuitive consequences, wherein any statement about the future, like "I will marry her", is not a lie until the speaker is dead, even if they made the statement with no intention to follow through.

Coleman and Kay (1981-03) found empirical support for subjective falsity as a condition of lying over objective falsity. They conducted a study examining whether certain conditions were more important than others in establishing an utterance to be a lie. They presented participants with multiple scenarios involving an utterance that held a combination of three conditions: objective falsity, subjective falsity, and intent to deceive, and then asked participants to give their lie judgments (the degree to which they agreed with the claim that the utterance was a lie) on a scale.

Coleman and Kay found that, in general, utterances with a higher amount of the three conditions had higher mean lie-judgments than those with less. However, for the scenarios where only a few of the conditions were present, the conditions affected participants' lie-judgments to different degrees. When only one of the three conditions was present, subjective falsity had the the highest mean lie-judgment (4.61), intent to deceive was in the middle (3.48), and objective falsity had the lowest mean lie-judgments (2.97). Additionally, the scenario which only had the condition of subjective falsity had a higher lie-judgments over the one which had both objective falsity and intent to deceive (3.66). On the other hand, the scenario with intent to deceive combined with subjective falsity had a mean lie-judgment

of 5.61.

Some studies, however, do argue that objective falsity is more important. In 2015, Turri and Turri published a rebuttal to Coleman and Kay, arguing that Coleman and Kay's methodology was flawed, and respondents were being led to answer in ways that were not actually reflective of their true conceptions of lying (Turri and Turri 2015-05). Their concern was that Coleman and Kay's methodology led respondents to answer as if they were taking the perspective of the speaker in the story, and not a neutral, third-person perspective about whether the claim was a lie or not. They conducted a study based on their revisions, where they asked participants whether a lie failed or whether it was successful, and, separately, whether the speaker thought the lie failed or whether it was successful. Turri and Turri found that the actual falsity of the claim was most, not least important for whether an assertion was a lie.

In turn, in 2016, Wiegmann et al. published a rebuttal to Turri and Turri, where they argued that Turri and Turri's study also had methodological problems which undermined the quality of their results and that their finding that the actual falsity of the assertion was very important was unjustified (Wiegmann et al. 2016-05). They specifically pointed out that the distinction that Turri and Turri attempted to draw, between the 'trying to achieve' a situation and the 'achieving' a situation was not a very natural one, which led respondents to answer with regard to whether a claim was objectively false rather than whether it was a lie. In conducting a study which avoided these potential issues, Wiegmann et al. found that their results reinforced the initial ones of Coleman and Kay. They emphasized that respondents of studies that regarded lying were highly sensitive to the pragmatic implications of the way that questions were framed.

This discussion of untruthfulness as a condition of lying has focused on whether it is subjective or objective falsity that is a condition for lying. However, there is a lot of discussion in the literature over whether falsity (subjective or objective) is necessary at all for an utterance to be a lie. The section of the literature review on implicature involves a lengthier

discussion of this topic.

Intent to Deceive Condition

The second of these conditions is the speaker's intent to deceive their addressee. This is the intent to deceive condition. It requires the speaker to intend to produce a false belief in the addressee through their utterance. The implication of this is that irony, jokes, and other untruthful statements are not lies because they are not intended to actually deceive the addressee, despite their falsehood (Mahon 2016-10-10).

It is a matter up for debate whether intent to deceive is necessary for an utterance to be a lie, with the field split into two camps: deceptionists, who believe that this condition is necessary, and non-deceptionists, who believe there are cases where a speaker can lie without intending to deceive the addressee.

Non-deceptionists often point to bald-faced lies as an example of utterances that are not intended to deceive but are still considered lies (Horn 2017; Wiegmann and Meibauer 2019-07-31). Bald-faced lies arise when someone makes an utterance which everyone involved in the conversation knows is false, like someone who says "I'm holding a fork" when everyone can see that they're holding a spoon." Since everyone knows the utterance is false, non-deceptionists argue that the utterance cannot be intended to deceive, but it intuitively still seems like a lie.

Some deceptionists respond by arguing that bald-faced lies are not really lies (Dynel 2015; Wiegmann and Meibauer 2019-07-31). Other deceptions argue that even in cases where the speaker knows that they will never be believed, they are in fact still trying to deceive the addressee— intent to deceive is not reliant on whether deception was actually accomplished (Krstić 2019).

Arico and Fallis (2013-01-15) empirically tested the claim that bald-faced lies are not lies by conducting a study where they presented participants with examples of bald-faced lies and asked them to judge whether these were lies. In nearly all cases, the great majority of

participants agreed that these were lies. Their results thus suggest that bald-faced lies are in fact lies, and thus that intent to deceive is not necessarily required for an utterance to be a lie.

Other Conditions

While intent to deceive and untruthfulness are the conditions that most frequently occur in condition-based accounts of lying, there a handful of potential others as well.

One of these is the statement condition, which requires that a person make a statement for that statement to be a lie. While this condition might seem obvious and already presupposed by the other conditions, articulating it distinguishes lies from situations where one deceives non-linguistically. One example is a man wearing a wedding ring when he is not married. Because there is no statement, this scenario cannot involve a lie (Horn 2017).

The statement condition also clarifies that lies can take many forms, and do not necessarily need to involve a verbal assertion. It is possible to lie in ASL, for example, as well as to lie through smoke signals, because there is still a conventional statement being made through the smoke. The statement condition also implies that it is not possible to lie by omission, because emitting something does not translate to making a positive statement (Mahon 2016-10-10).

Another condition which might define which statements are lies is the addressee condition. This one is often excluded as a potential condition— Mahon (2016-10-10) includes it in his essay summarizing the conditions, but Horn (2017) does not. This condition requires that lies be directed to a specific audience. The addressee condition implies that if a person happens to overhear a speaker lying to an addressee, then the addressee was lied to, but the person who overheard the statement wasn't (Mahon 2016-10-10).

2.1.2 Lies as prototypes

Prior to 1981, the underlying assumption held by those investigating the definition of a lie was that lies were conceptually structured as checklists of necessary or sufficient conditions, conditions which needed to be 'checked off' in order for utterance to be a lie (Sweetser 1987). The first study to challenge this assumption was Coleman and Kay's classic 1981 paper "Prototype Semantics: The English Word Lie." This was also the first empirical investigation into the definition of a lie.

Coleman and Kay (1981-03) approach the question of the definition of a lie in the context of seeing whether the definitions of certain words are more accurately discussed as gradable prototypes than as rigid categories, as an attempt to account for the flexibility of the concept of lying with regard to the presence of the component conditions.

Prototypes are a conceptual category that is defined by there being a gradable degree of belonging to that category. There is usually an 'ideal' member, which is highest on the scale, and where something falls on the scale is defined by how well it aligns with that ideal member of the category (Faulkner 2013; Sweetser 1987). For lies, this would mean that there is such a thing as an ideal lie, but there are also lots of utterances which have some of the qualities of lies but not all of them. These then function as 'half-lies'.

Coleman and Kay find that participants are more likely to judge as lies utterances in which more of the conditions were present, and less likely to judge as lies utterances which have less of the conditions. They conclude that lies are not structured as a checklist of potential conditions. Instead, lies are likely conceptually structured as gradable prototypes, functioning on a scale, with a higher amount of fulfilled conditions corresponding to a higher degree of lie-hood.

Experimental literature on lies has broadly embraced the conception that lies function as gradable prototypes (Faulkner 2013). This is reflected by the convention in the field to solicit participant lie judgments on a 5 or 7 point scale rather than as a forced choice.

2.2 Implicature

A conventional implicature is a meaning conveyed by an utterance that is not logically entailed by that utterance. Implicatures are invoked when a speaker implies one thing by saying something different Davis (2019); Horn (2017). There is thus characterized by a disparity between the speaker's meaning and the sentence's meaning (Grice 1989). Implicatures arise not because of the semantic content of the utterance but because of contextual factors and the understanding of shared conversational conventions— despite not forming part of the semantic contentm implicatures are still inferred and predictable. Conversational implicatures were first substantially examined by the philosopher H. P. Grice (Davis 2019; Grice 1975).

(4) Matthew asks Luisa, "Are you going to Woads today?" Luisa replies, "No, I have to work on my thesis."

(5) Matthew asks Luisa, "Are you going to Woads today?" Luisa replies, "My thesis is due on Friday."

In (4), Luisa is directly expressing the proposition that she is not going to Woads. In (5), however, she does not outright state that this is the case. However, by expressing the proposition that she has a major assignment which is due very soon, she is *implicating* that she is not going to Woads, because she needs to work on that assignment. Thus, the understood meaning of Luisa's sentence is similar in both (4) and (5) despite their semantic differences.

Conversational implicatures should be distinguished from conventional implicatures. While conversational implicatures depend on the context of the utterance or the conversation to be conveyed, conventional implicatures arise directly from the lexical, syntactic, or semantic content of the utterance Davis (2019); Grice (1989).

(6) Maria is an athlete but she's smart.

(6) is an example of a a statement that carries a conventional utterance. The use of 'but' implicates that people who are athletes tend not to be smart, and that Maria is an exception to this trend. By contrast, Luisa's utterance in (5) is an example of a conversational implicature, because the implicature arises through the conversational context of her utterance and not only through the literal contents of the utterance itself. This study focuses primarily on conversational implicature.

There are four main types of conversational implicatures. These types arise from Grice's notion of the cooperative principle, which defines how people can most effectively achieve mutually intelligible communication in social and conversational settings (Grice 1975). Grice argued that these maxims can also operate on a non-literal level. When a speaker appears to flout one of these maxims when making a statement, the addressee should read additional, not-directly-said information into the statement until it is in accordance with the cooperative principle. Observing the maxims on a non-literal level thus leads to the emergence of conversational implicatures (Davis 2019).

The cooperative principle is an umbrella for four communicational maxims: the maxim of quality, the maxim of quantity, the maxim of relevance, and the maxim of manner. There are types of conversational implicatures that arise in response to each of these four maxims.

The maxim of quality guards against the making of statements that are false, or for which the speaker does not have adequate evidence (Grice 1975, 1989).

(7) It is raining outside.

(7) is an example of an utterance that contains a quality implicature. The statement has no claims about the belief of the speaker, but the implicature associated with the making of this utterance in the first place is that the speaker believes and has adequate evidence to support the claim that it is in fact raining outside.

The maxim of quantity says that a speaker should provide as much and not more information than is required for them to make their point (Grice 1975, 1989).

(8) I ate some of the chocolates

One example of a quantity implicatures, or a scalar implicatures, is found in (8). The use of 'some' in the utterance does not logically preclude the speaker from having actually eaten all the chocolates. However, the choice to use 'some' as opposed to 'all' in the statement indicates that the speaker did not in fact eat all the chocolates.

The maxim of relevance is very straightforward; it requires that utterances be relevant to the conversational context (Grice 1975, 1989). An example of a relevance implicature is Luisa's utterance in (5). On first interpretation, her statement doesn't seem relevant to the question that Matthew asked her, and so it does not seem like she is following the maxim of relevance. However, Matthew would still assume that Luisa is following the conversational principle, and so he would infer how exactly Luisa's statement is relevant to his question—that her thesis is preventing her from being able to go to Woads.

Finally, the maxim of manner centers on perspicuity and requires that statements be brief, sensibly ordered, and refrain from ambiguity or obscurity. Manner implicatures generally arise as part of the form of the utterance (Grice 1975, 1989; Rett 2020).

(9) Elisa produced a series of sounds that held a resembled the tune of Sweet Home Alabama.

The maxim of manner requires that utterances be as brief an unambiguous as possible. Referring to the Eliza's singing in (9) as 'producing a series of sounds' seems to violate this requirement. The implicature that arises on the basis of this is that Eliza's singing was likely very bad, but it would be impolite and unsociable to say so directly.

2.2.1 Can deceptive implicatures be lies?

As discussed in the section on whether untruthfulness as a condition for lying, some linguists believe that utterances do not necessarily need to be untrue for them to be lies. This is because they believe that it is possible to lie through a deceptive implicature and not just

through an outright false assertion.

Imagine two alternative versions of scenarios (4) and (5):

(10) Henry asks Jordan, "Are you going to Woads today?" Jordan replies, "No, I have to work on my thesis."

(11) Henry asks Jordan, "Are you going to Woads today?" Jordan replies, "My thesis is due on Friday."

(10) and (11) differ from (4) and (5) in that Jordan actually *is* planning to go to Woads that night. However, she doesn't want Henry to find out, because she knows that if he thinks she's going, he will go as well and try to dance with her, which she doesn't want. Her thesis *is* due on Friday, but she's already all done with it.

In both scenarios, Jordan is deceiving Henry. In (10), she does so by telling him an outright untruth, intending to produce in him the false belief that she is not going to Woads. However, in (11), Jordan is not saying anything which is actually false- again, her thesis is in fact due on Friday. However, she knew that by saying so, Henry would erroneously infer that she was not going to go to Woads. Jordan is thus using an implicature to deceive Henry into having a false belief— this is what is a called a deceptive or a false implicature (Adler 1997-09).

In both (10) and (11), Jordan had the same intention— to deceive Henry. However, only (10) is straightforwardly a lie, because in (10) Jordan's utterance was false, but in (11) it was not. Is this separation between the two scenarios justified? Are Jordan's actions in the two really so different, such that (10) a lie but (11) is not? Is (11) merely misleading? (Adler 1997-09).

Quite a lot of literature on lies is dedicated to answering this question, in exploring whether a deceptive conversational implicature can be considered a lie. Results so far have varied greatly.

On the one hand, untruthfulness does seem like a very central component of most lies.

Coleman and Kay (1981-03) found that (subjective) falsity is the most important condition defining whether or not an utterance is a lie. Weissman and Terkourafi (2018) conducted a study where they investigated 15 different cases of conversational implicatures. In 9 of the cases, the utterance was strongly rated as a not-lies; 2 were strongly rated as lies, and 4 were somewhere in the middle (Weissman and Terkourafi 2018). Thus, although they found that deceptive implicatures were in general not considered by participants to be lies, Weissman and Terkourafi were unable to make very generalizable statements about whether deceptive implicature were or weren't lies.

On the other hand, deceptive implicatures seem like they are very similar to lies in a few key respects. They are both intended to achieve the same outcome, that of deception. In both cases, they take advantage of communicative norms to lead an addressee to false beliefs, and they both violate the cooperative principle in doing so. Wiegmann et al. (2017) conducted a study analyzing deceptive implicatures derived from Grice's maxims and found that particpants were amenable to viewing deceptive implicatures as lies, particularly so in cases where a speaker deliberately did not communicate important information to the addressee in order to deceive them.

Antomo et al. (2018) conducted a study where they asked participants to evaluate cases of false assertions and deceptive implicatures on a 5-point lie judgment scale. As expected, they found that participant lie judgments for the false assertions had a median of 1. However, they also found that lie judgments for the deceptive implicatures had a median of 2, which, though lower than false assertions, was still closer to being a lie than not a lie. They argued that Weissman and Terkourafi had not sufficiently signalled the speaker's intent to deceive their addressee in the vignettes that they created for their study, hence their diverging results.

Ultimately, the question of whether deceptive implicatures can be lies is still up in the air. One hypothesis for the differing results to this question is that some deceptive implicatures are more lie-like than others, and so different deceptive implicatures occupy different places on a lie-judgment scale. However, what it means to be 'more lie-like', and what qualities

make an implicature more lie-like, are still to be determined.

2.3 Morality and lying

As mentioned in the introduction, the majority of the literature on the relation between lies and morality approaches it from an ethical perspective rather than a definitional one, which is the one taken in this senior essay. There are, however, a few primarily theoretical studies that do touch on this approach.

Timmermann and Viebahn (2020-06-15) argued that calling a statement a lie is a morally charged, value-negative statement, and so it would follow that morality would impact lie judgments.

Going back to the question of whether deceptive implicatures can count as lies, Adler (1997-09) looked at whether lies can be distinguished from deceptive implicatures on a moral basis. He conducted a comparative assessment of the moral weight of lying and deceiving through implicature. He acknowledged that there is a general shared intuition that lying is always morally worse. He asked— how morally distinct is for a speaker to assert what they believe is false versus to assert what they believe is true while implicating what they believe is false? When someone deceives through an implicature instead of outright lies, was any moral progress done?

Adler argued that there was a felt difference between a deception done via a false assertion and a false implicature, and that this felt distinction is reflective of a real ethical difference between the two. This distinction is derived from the norms of communicative practice. People expect cooperativity from those who they engage in communication with, and lies are an abuse of that implicit trust. This implies that implicatures cannot be lies because do they do not share the same moral status that lies have, as they do not violate the cooperative principle as much as lies do.

There are very few experimental investigations into the relationship between morality and

the definition of lying. For Wiegmann et al. (2022), it was a secondary component of their investigation. The primary goal of their paper was to see whether the speaker's belief in the falsity of the claim was a necessary component of a lie, such that a deception accomplished via conventional implicature (wherein the claim was not objectively false) could also count as a lie.

To do this, Wiegmann et al. examined how their participants' lie judgments corresponded to their judgments about whether the speaker believed the truth of what they were saying. At the same time, they also tracked the degree to which participants agreed that the speakers had commited to the truth of an utterance. As part of this investigation, they also examined how lie judgments track against morality judgments as a proxy for lies. They find that it is not as strong as the relation between commitment judgments and lie judgments, and they do not examine the question of the impact of morality further. They go on to develop a commitment based account of lies, wherein the injury of a lie is because the speaker commits themselves to the truth of an utterance that is not actually true.

Methodology

This study aims to examine whether and how moral valence impacts participants' lie judgments for deceptive utterances. It also aims to measure whether moral negativity makes people more likely to judge a deceptive implicature to be a lie in comparison with a false assertions. It finally aims to see whether there is an interaciton effect between moral valence and inference type.

3.1 Experiment Design

The study has a 2(moral valence) x 2(inference type) x 6(vignette) factorial design, so there is a total of 24 total conditions. The two independent variables (and fixed factors) are moral valence (this being positive/negative) and inference type (whether deception is accomplished through a false assertion/false implicature). The dependent variables are the study participants' lie judgments.

6 vignettes were presented to study participants to read. The vignettes are the study's random factors, meaning that they were 6 examples chosen to be representative out of a broader range of other potential examples. Each vignette has four versions corresponding to the 2x2 independent variable design (moral valence x inference type): one morally positive with a false assertion, one morally positive with a false implicature, one morally negative with a false assertion, and one morally negative with a false implicature. The full list of

vignettes is in Appendix A.1.

The study consisted of asking participants to read one iteration of a vignette and judge the degree to which they agreed with the claim that the utterance in the vignette was a lie. Study participants were only presented with one of the four variations of one vignette, in order to try and avoid bias influencing the results were participants to realize they were being asked about lies.

It is hard to standardize the moral valence of an utterance across participants in a study because moral intuitions tend to differ widely among different individuals. The study aimed to avoid that problem by including scenarios in the vignettes that were either clearly morally positive or clearly morally negative, such that judgments about the exact moral valence of the situation were of less importance than the great moral contrast between the two. For one example, the morally positive variant of the "Earrings" vignette involved someone deceiving their friend to pleasantly surprise them with their thought-lost earrings, while the morally negative iteration of the vignette involves the friend selling the earrings for money.

Modifications between each iteration of the vignette were kept as minimal as possible in order to avoid additional considerations impacting responses. For example, for vignette "Surprise", the deceptive assertion vignette involves Chris saying "I'd like you to come with me to my apartment because I need to get my jacket," whereas the deceptive implicature iteration involves him saying "I'd like you to come with me to my apartment. I need to get my jacket."

The deceptive implicature iterations of the vignettes also feature a combination of the various types of conversational implicatures, in order to ensure that the results reflect evidence about implicatures overall rather than just one implicature type. For example, vignette "Packing bags" involves a relation implicature, as remodeling is not why the bags are packed. Vignette "School test", on the other hand, involves a scalar implicature, because "some" does not logically exclude "all."

During the question portion of the study, participants were presented with the bolded

statement "X lies to Y" (with X and Y replaced with the relevant names of the specific vignette they read.) Thy were then asked to rate the degree to which they agree with the claim that the utterance was a lie on a 7-point scale, where 1 equaled completely disagree, 4 equaled unsure, and 7 equaled completely agree. Responses were elicited on a 7 point Likert-scale scale because it allowed for greater sensitivity to distinctions in participant judgments. In addition, this is the most common method of capturing lie judgments in the literature. The majority of the recent empirical literature on the subject takes this approach, such as in Wiegmann et al's 2022 study and García-Carpintero's 2023 study.

The study question is followed by a comprehension question. The responses of participants who failed to correctly respond to the comprehension question were discounted from the data.

The study also gathered biographic data about age and gender in order to see whether there is a relation between either of these and participant lie-judgments. Previous studies have examined that gender plays a big role in notions of politeness, and there is significant overlap between politeness and the realm of 'white lies' Gotzner and Mazzarella (2021-09-27).

The study was conducted digitally, in order to gather data quickly from particpants from many English-speaking backgrounds. It was created through Qualtrics and distributed through MTurk, though it was restricted it to survey-takers who had a high amount of surveys completed and a high approval rating in order to control responses for quality. 400 participants were paid to take the survey.

Conducting the study digitally imposed certain limitations, in that there was no capability to observe participants as they were taking the survey, ask individualized follow-up questions, or control for participants who spend very little or a lot of time responding to the question. Ultimately these limitations were worth the benefit of having access to a much larger pool of respondents.

3.2 Participants

For the experiment, participants were recruited on Prolific. They completed an online survey built using Qualtrics. All participants were required to be at least 18 years old and from either the US or the UK. Of the total of 441 survey respondents, 61 failed the comprehension test. The analysis in the following sections of the essay is based on the remaining 380 responses. 39% of these participants were male, 60% were female, and 1% were nonbinary. The mean age group was 35-44. Participants received $0.80 for an estimated three minutes of participation.

Results

The data was analyzed using linear mixed effect models, with moral valence, inference type, and their interaction as fixed effects and vignette as a random effect. All of the analysis was conducted using R; the 'lmer' function from the 'lme4' package was used to fit a linear mixed-effects model to the data. A Type II Wald chi-square test was done using the 'ANOVA' function from the 'car' package to understand the significance of each variable in the model as well as how they interacted with each other. The results are represented in Table 1.

	Chi-square statistic	dF	P-value
Inference type (Main effect)	6.8264	1	0.008982
Moral valence (Main effect)	38.5371	1	5.372e-10
Inference type:Moral valence (Interaction effect)	0.5887	1	0.442927

Table 4.1: Main Effects and Interaction Effect

These results indicate that inference type has a statistically significant effect on lie judgments, and that moral valence also has a highly statistically significant effect on lie judgments (the p-value is very close to zero). However, the interaction between moral valence and inference type is not statistically significant, indicating that the effect of inference type on lie judgments is not impacted by the moral valence and vice versa.

For a more in-depth look at the data, the 'emmeans' function of the 'emmeans' was used

to compute the estimated marginal means of the interaction between inference type and moral valence. This gives the predicted means of the lie-judgments for each combination of levels of inference type and moral valence. Pairwise comparisons were computed for each estimated marginal means, with the results represented in Table 2. Comparisons with a p-value of less than 0.05 are statistically significant.

Contrast	Estimated difference in means	Standard error	dF	T-ratio	P-value
Assertion Good – Assertion Bad	-1.051	0.260	371	-4.035	0.0004
Implicature Good – Implicature Bad	-1.345	0.282	372	-4.773	<.0001

Table 4.2: Emmeans, Within moral valence

Both of these comparison pairs are statistically significant. This indicates that there is a statistically significant difference between the mean of lie judgments between morally positive assertions and morally negative assertions. There is also a statistically significant difference between the mean of lie judgments between morally positive and morally negative implicatures.

Contrast	Estimated difference in means	Standard error	dF	T-ratio	P-value
Assertion Good – Implicature Good	0.645	0.270	371	2.395	0.0798
Assertion Bad – Implicature Bad	0.351	0.273	372	1.288	0.5710

Table 4.3: Emmeans, Within inference type

Neither of these comparison pairs are statistically significant. This indicates that there is no meaningful difference between the mean of lie judgments for morally positive assertions and morally positive implicatures, and also no meaningful difference between the mean of

lie judgments for morally negative assertions and morally negative implicatures.

Contrast	Estimated difference in means	Standard error	dF	T-ratio	P-value
Assertion Good – Implicature Bad	-0.700	0.258	371	-2.713	0.0350
Implicature Good – Assertion Bad	-1.696	0.283	372	-5.986	$<.0001$

Table 4.4: Estimated difference in means

Both of these comparison pairs are statistically significant. This indicates that there is a statistically significant difference between the mean of lie judgments for morally positive assertions and morally negative implicatures, as well as a statistically significant difference between the mean of lie judgments for morally positive implicatures and morally negative assertions.

The overall results are graphically represented on the scatter plot in Figure 1 and on the violin plot in Figure 2, in order to more effectively see the variance of the data. The numbers on the original scale on which participants ranked their lie judgments (where 1 was most lie-like and 7 was least lie-like) were flipped so that a higher number now represents a higher lie judgment.

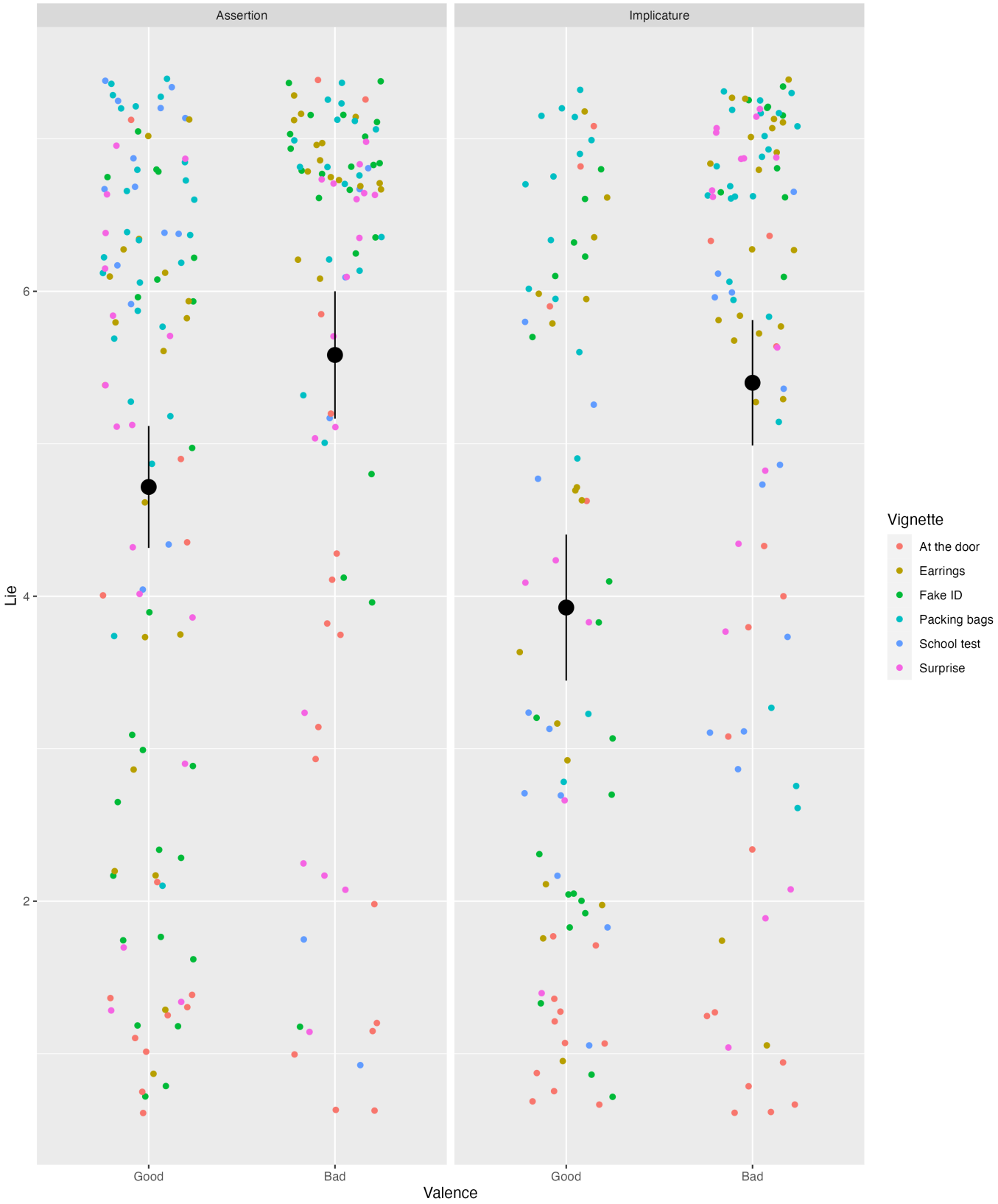

Figure 4.1: Lie Judgments on a Scatter Plot

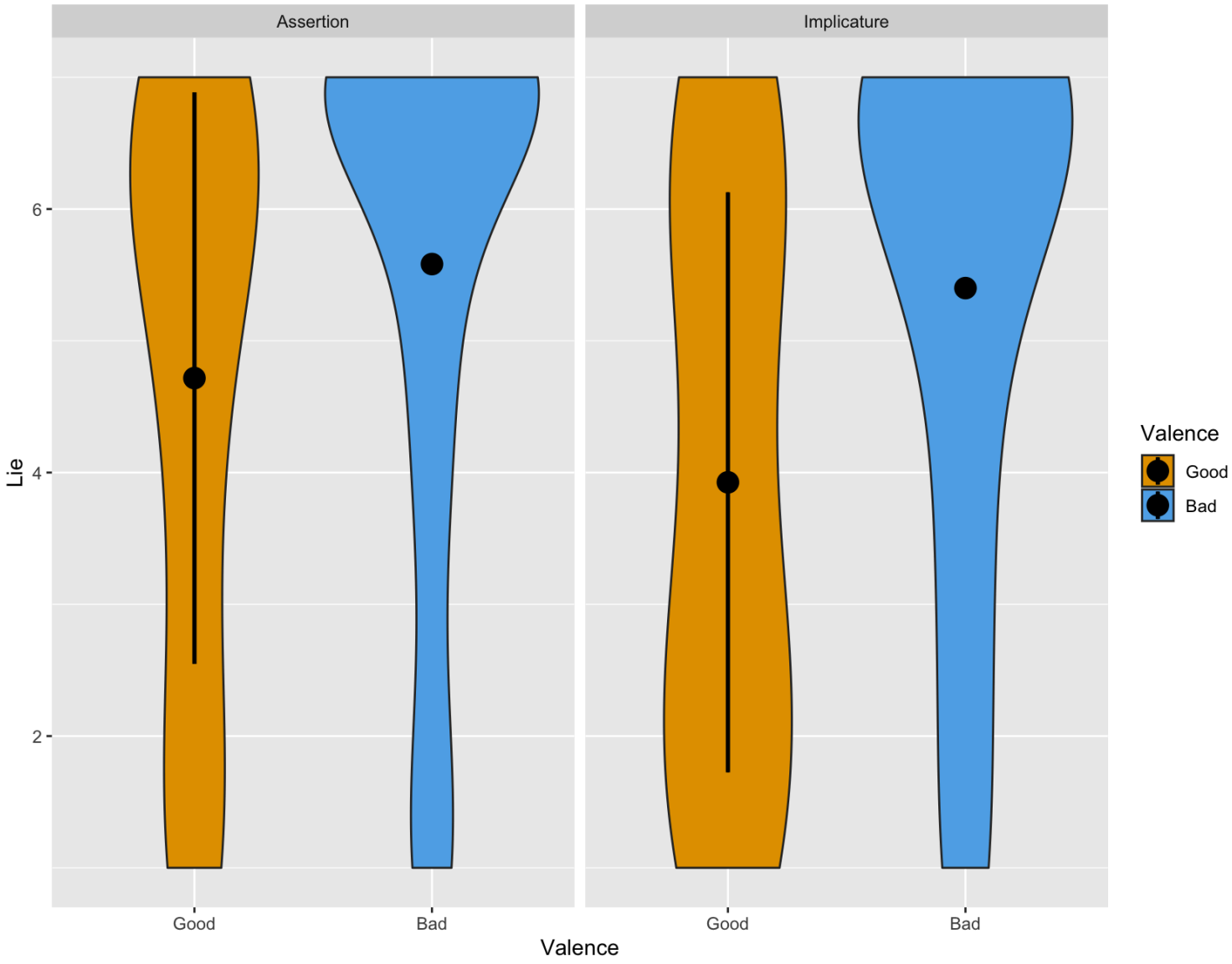

Figure 4.2: Lie Judgments on a Violin Plot

Figure 2 highlights that responses had a low variance for the morally negative vignettes, instead clustering around the top of the scale. The morally positive vignettes, on the other hand, have a very spread out variance across the entire lie judgment scale.

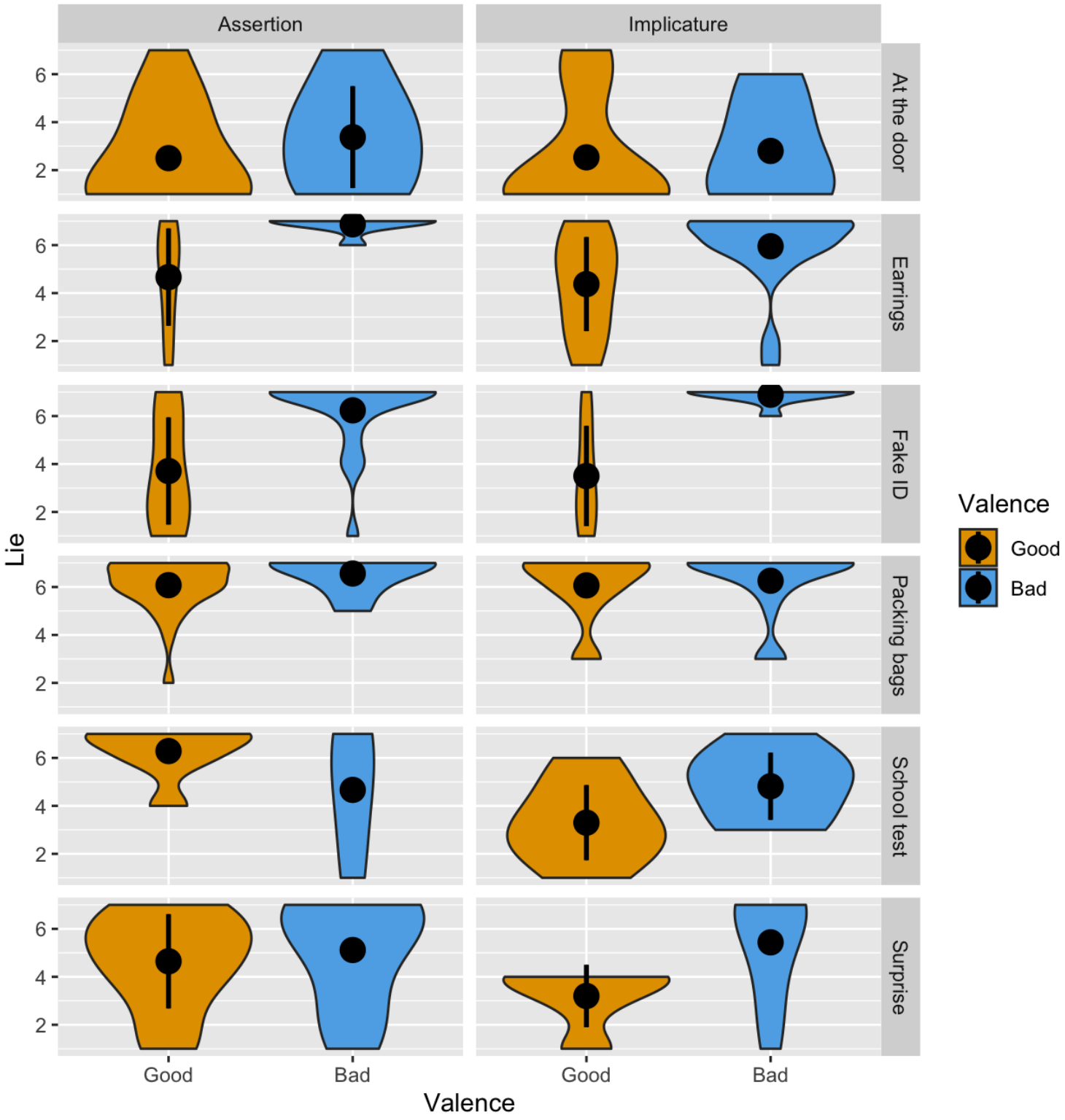

Figure 4.3: Lie judgments for each vignette

In Figure 3, the results are divided by vignette, in order to see whether the implicature type or other considerations which varied between vignettes impacted the results. Results between vignettes vary greatly, and there is no consistent pattern to the variance of lie judgments for individual vignettes.

There was no statistically significant impact of age or gender at all.

Discussion

5.1 Can deceptive implicatures be lies?

The presence of a main effect of inference type indicates that deceptive assertions are more likely to be perceived as lies than deceptive implicatures. This result is unsurprising, given that this pattern is well-established in the literature on lies. The more interesting question that is still at issue is whether people treat deceptive implicatures as lies that are simply less lie-like than deceptive assertions, or whether they're not considered lies at all. This study's results indicate that the former is correct.

Despite the presence of a main effect, there is no statistically significant difference between the estimated difference in the mean lie judgments for morally positive assertions and morally positive implicatures, and for morally negative assertions and morally negative implicatures. The implication of this is that the mean lie judgments for morally positive deceptive utterances, regardless of whether the semantic content of the utterance is false or not, are very similar to each other and that the mean lie judgments for morally negative deceptive utterances behave the same way. This data shows that participants likely intuitively consider some deceptive implicatures to be lies, though they are less lie-like than false assertions.

5.2 Lies and moral negativity

The data shows a significant main effect of moral valence (38.5371 chi-square statistic). This indicates that participants perceive a morally negative deceptive utterance as more of a lie than an otherwise identical morally positive deceptive utterance. Thus, when a deceptive utterance is morally negative, people are more likely to then consider that utterance a lie. The implication of this usage pattern is that that moral negativity should be considered inherent to the definition of a lie.

However, this conclusion is complicated by the lack of interaction effect between moral valence and inference type. If it is true that moral negativity is a component of the definition of a lie, and deceptive implicatures are less likely to be considered lies than false assertions, then greater moral negativity should increase lie judgments for deceptive implicatures so their lie judgments are closer to those of deceptive assertions. The data shows that this is not the case. The implication of this is that moral negativity is not part of the definition of a lie. These results thus seem contradictory.

The lack of interaction effect between inference type and moral valence implies that something is being held consistent between the vignettes where the two inference types (implicature and assertion) are varied. This is what moral negativity is actually affecting to produce the differing lie judgments between morally positive and morally negative utterances. Stated differently, this means that the actual truthfulness or untruthfulness of a deceptive utterance is irrelevant for the association of moral negativity with lies.

The objective falsity of the utterance, and even more so the subjective falsity of the utterance, are in the literature seen as important components to the definition of a lie. Coleman and Kay (1981-03) actually found that the subjective falsity of the utterance (that is, the speaker's belief that the utterance was false) was the single most important thing that made an utterance a lie. It is again unintuitive that moral negativity, if it is central to the concept of a lie, would not interact with subjective falsity at all.

5.3 Approach 1: Morality as incidental to the definition of a lie

It is conceivable that what morality is interacting with to produce the main effect of moral valence is the deceptiveness of the utterance. The speaker's intent to deceive the addressee through their utterance is held consistent in my study between implicatures and assertions, and in each case the aim of the deception is the same. This is also intuitive because deception itself is seen as morally negative.

It could be the case that participants are simply associating greater moral negativity with a worse level of deception or a worse intent to deceive, which in turn makes an utterance more lie-like. The deceptiveness of lies thus could be where our belief that lies are morally negative is derived from in the first place. A fair amount of the literature on lies that has looked into the question of whether intent to deceive is required for an utterance to be a lie has concluded that it is an important component of the definition of a lies, and this potential conclusion would be consistent with that finding as well.

Thus, what this could indicate is that the association of lies with moral negativity is not a question of the actual definition of a lie but rather a question of either the usage of lies to achieve deception or of the intent behind the usage of lies to achieve deception. Morality might thus simply be incidental to the question of the definition of a lie. Stated differently, it might be the case that the actual definition of lies is morally neutral, but because lies used for deception, which is itself perceived as a morally negative thing, we have come to associate lies with moral negativity in a way that makes us more likely to consider morally negative utterances to be lies. Evidence for this also lies in the usage of phrases like, "It may be a lie, but *come on*." The emphasis on 'come on' implies that the the speaker's lie may still be a lie, but it's not the sort of lie that we are worried about when we penalize lies for their moral negativity.

However, the association of lies with moral negativity being simply a product of their deceptiveness seems contrary to the conclusion of many philosophical papers on lies that there is something uniquely morally negative about lies that separates them from other forms of linguistic deception, such as (Timmermann and Viebahn 2020-06-15). Though it's fair to say that the question of whether and why lies are morally worse is also still a controversial one in the literature, it does intuitively seem like we treat lies as if they are uniquely bad in comparison to other forms of linguistic deception.

Ultimately, treating morality as incidental to the definition of a lie is an unsatisfying conclusion, primarily so because it fails to explain to strong and direct link between moral valence and participant lie judgments in the study.

5.4 The conceptual structure of lies

As discussed in the literature review, a lot of the experimental work on lies has approached the question of the definition of a lie from the perspective of treating lies, on a conceptual level, as prototypes. This means that lies are gradable concepts and there's a specific combination of necessary or sufficient conditions which defines the ideal lie (these are most often subjective falsity and intent to deceive). Utterances which have some but not all of these conditions operate on a scale from most lie-like to least-like and then to not at all lie-like.

If moral negativity is one of the properties that defines a lie, and lies do conceptually function as prototypes,then participants' lie judgments for the morally negative utterances should theoretically cluster around the top of the scale. Lie judgments for the morally positive utterances, on the other hand, should cluster around the middle of the scale, as they have some but not all the properties of the ideal lie. While this prediction turns out to be accurate in the case of the morally negative utterances, it is not the case for the morally positive ones— they do not cluster around the middle of the lie judgment scale. Although they do average out to the middle, participants' lie judgments for morally positive utterances

have a much larger variance than those for the morally negative utterances. This is most clearly visible in Figure 2, where the morally negative graph is very top-heavy, indicating that most responses are similar in being high on the scale, while the morally positive side is distributed evenly through the lie-judgment scale, indicating that the number of responses for every number on the scale were similar.

The implication of this high variance is that morally positive lies are not being treated as half-lies, or non-ideal lies. Instead, it seems more likely that participants are very confused and very divided about how exactly to classify morally positive deceptive utterances. Thus, if it is in fact the case that moral negativity is part of the concept of lies, then this study's findings do not support the claim that lies are conceptually structured as prototypes.

One alternative conceptual structure, which could explain the seeming contradictions in this study's results, is that lies are functioning as dual character concepts.

5.5 Approach 2: Lies as dual character concepts

Dual character concepts have definitions that function on two levels; the first level is characterized by concrete features, and the second level by abstract values that are a realization of those concrete features. They were first described in 2013 by Knobe, Prasada, and Newman in their paper "Dual character concepts and the normative dimension of conceptual representation." The dual character of these concepts means that there are two criteria for belonging to the concept and two bases for evaluating whether something is a member of the concept.

Based on the results of the study conducted in this senior essay, lies are a good fit for the criteria of dual character concepts. In one of the studies they conducted for the paper, Knobe et al. (2013) found that when something had the concrete features but did not fulfill the abstract values associated with the concept, participants judged that thing to be part of the concept in one sense but not in another. This provides a satisfying explanation for the

confusion, embodied by the high variance, that participants seemed to have about whether the morally positive deceptive utterances were lies or not— because they are lies in one sense, but not lies in another.

This implies that if lies do function as dual character concepts, moral negativity *is* part of their definition, on the level of being one of the abstract values that lies realize. This explains why the results of this study indicated that moral negativity increases participant lie judgments— because moral valence is actually part of the definition of a lie, just in a very particular sense. When an utterance fulfills the concrete criteria for lies but does not fulfill the value of moral negativity, that utterance only realizes the criteria for one level of the definition of a lie but not the other. Thus, when someone tells a white lie and is accused of lying, an appropriate response to this accusation might be something like 'But *c'mon*' (or, if one belongs to Gen-Z, "Be so for real").

This also explains why false assertions had higher lie judgments than deceptive implicatures. Untruthfulness is also part of the definition of a lie, but unlike moral negativity, it is part of first level of the definition, that of concrete features, rather than the second level of fulfilled abstract values. Thus, deceptive implicatures may share the abstract values that untruthful lies have, but they don't share the concrete features, and so they are like lies in ethos but not in another more technical sense— hence, when someone misleads another person through a deceptive implicature that is not actually untrue, they might say of their statement, "Well, it wasn't *technically* a lie."

The dual nature of the definition of a lie would also explain why this study found no interaction effect between moral valence and inference type despite finding that each had a main effect. If untruthfulness is part of the definition of lies on the concrete feature-level, then whether or not it is fulfilled has no bearing on whether moral negativity, on the abstract value-level, is fulfilled. The reverse is also the case.

It resolves the seeming contradiction, discussed earlier, between the strong main effect of moral valence, which implies that moral negativity is part of the definition of a lie, and the

lack of interaction effect between moral valence and inference type, which implies that moral negativity isn't part of the definition. Moral negativity is part of the definition, but it's part of another component of the definition than inference type, hence the lack of interaction. Lies being dual character concepts offers a very comprehensive explanation for many of the complexities in this study's data.

If lies do function as dual character concepts, this finding has very significant implications for the field and for many of the questions about the definition of a lie which are being debated in it. It would seem to explain why none of the conditions which seem important for the definition of a lie actually seem outright *necessary* in order for an utterance to be considered a lie. For each one of the conditions, it is possible to visualize a lie that does not contain that condition. This is because the presence of the conditions is relevant to whether one of the two levels of the definition of lying is fulfilled, but not both, and so utterances which do not have conditions essential to the concrete features of lies are sometimes considered lies if they have strong enough abstract values, and vice-versa. Thus a dual character structure would explain why deceptive implicatures are occasionally seen as lies even when they are not outright untruthful.

A central question for further consideration should thus be whether lies actually do fit the description of dual character concepts. Determining that moral negativity is part of the abstract definition of lies does not help narrow down further what the concrete definition of lies is, and this should be the object of further study as well. Examining which components of the definition have interaction effects with each other might be a good way of narrowing down which components of the definition are part of the concrete level and which ones are part of the definitional level of abstract values.

5.6 Other moral and lingusitic considerations

In Figure 3, the variance of the data differs quite a lot between the different vignettes. The sample sizes for the results when divided by vignette are relatively small, and so all of the dissimilarities between the results for each vignette could simply be a matter of noise. On the other hand, though, these distinctions could be a reflection of differences in the vignettes themselves. Perhaps particular moral and linguistic considerations have a strong impact on lie judgments. Identifying these considerations can help further clarify the question of the definition of a lie and provide more informationa about on what bases people make their lie judgments.

One potential factor influencing participants' lie judgments is the weight of the moral stakes involved in the scenario. In the 'At the door' vignette, for example, Brenda's life is at stake were Charlie to find her. In the 'Fake ID' vignette, by contrast, what's at stake is Daniel getting a job with his new government ID. Lie judgments for the morally positive iterations of the 'At the door' vignette are much lower than they are for the positive iterations 'Fake ID' vignette. One reason for this could be that participants felt more comfortable seeing the utterances as more lie-like (this label implicitly being critical) in cases where the reason for the lie is less fundamentally important, as it is in a case where someone's life is at stake.

On the other hand, however, this does not explain why, in the 'school test' vignette, the lie judgments for the morally positive deceptive assertions are actually higher than those for the morally negative deceptive assertions. If it is correct that utterances where the stakes are lower are judged as more lie-like, then this should not be the case. One explanation could be that participants simply did not think that the cause of Rachel's lie (helping her friend avoid the embarrassment of their friends knowing she had a bad grade) justified the lie in the first place, and so perhaps it was simply not a well-chosen scenario. Another potential explanation is that the 'school test' vignette is the only one which features a scalar implicature. Scalar implicatures may be more difficult to conceptually grasp, requiring more

thought than mere linguistic intuition.

Another potentially relevant moral consideration is the epistemic status of the addressee. Epistemic status involves the relation between the addressee who is the intended target of the lie and their access to the truth. One of the reasons why parents are comfortable lying to their children, as in the case of Santa Claus, is because children, as children, do not necessarily require full access to the truth— they are seen as too young for that. It may not be in the best interest of a child to find out that their parents are the ones giving them gifts each year.

Two final potentially relevant moral considerations could be whether the addressee is at any point going to find out about the lie, and whether they would approve of being lied to if so. One example is the case of the surprise party in the morally positive 'surprise' vignette and in example (1). Derek will find out about Chris's lie soon enough, as soon as they get to Chris's apartment. Derek will likely approve of being lied to, because the end goal was a nice surprise party for him. This situation is distinct from the morally positive 'packing bags' vignette, where the abusive husband will also eventually find out about the lie but would not approve of beinf lied to. This could be a potential explanation for why the lie judgments for the morally positive 'packing bags' vignette are much higher than that of the morally positive 'surprise' vignette.

There was no impact of age or gender in the data. It is thus unlikely that the distinctions between vignettes, or other patterns in the data, are the product of changes in the meaning of 'lies' or cultural distinctions. Instead, these are likely a question of participants' moral and linguistic intuitions. These vary greatly from person to person and do not necessarily have a singular rational explanation. This is why there are study participants who responded that the morally bad false assertions as not-lies, despite this seeming clearly to be the case. A further study into the topic of lies and moral negativity could look into the impact of these moral considerations in particular to see if varying them affects participants' lie judgments.

5.7 Potential Limitations

One potential limitation of the study design is that, in asking participants to provide their lie judgments on the 7 point Likert-scale, the study design itself creates the scale of participant responses, rather than this being an actual reflection of lies as definitionaly gradable. In the future, studies should also solicit lie judgments as forced choice,s where participants have to define utterances as either a lie or not a lie, in order to see whether responses are still similar.

One concern when conducting the study was that participants would react to the moral valence of the vignette they were presented with and would rate the moral valence of the utterance in the vignette rather than the lying. The study's results appear to indicate that this concern is unfounded in a few different ways. For example, if participants were simply ranking the moral valence of each vignette, then results for each vignette between the two inference types would be very similar. While this is the case for some of my vignettes, like the 'At the door' and the 'Packing bags' vignette, it is not the case for all of them, and in vignettes like the 'School test' and the 'Fake ID' these distinctions are very pronounced. This is an indication that participants were not simply rating the moral valence of the scenario on the scale.

Additionally, some of the morally negative scenarios in the vignettes are much morally worse than others, and some of the morally positive scenarios are far more positive, (albeit with the caveat that people's judgments about this may differ). One example is the morally negative iteration of the 'school test' vignette when compared with the morally negative iteration of the 'packing bags' vignette. Lying to friends out of embarrassment to cover up a bad score on a school test, although on the morally negative end of the scale, is still nowhere near as bad as not telling a significant other about one's infidelity.

Conclusion

This essay finds that moral negativity and untruthfulness are both part of the definition of a lie. It contends that lies are likely conceptually structured as dual character concepts, so that they have two definitional levels, one characterized by concrete features (of which untruthfulness forms part) and one characterized by abstract values (of which moral negativity forms part). This essay thus helps clarify that morally positive lies are lies in a technical sense but not in the sense of the abstract values that lies realize, one of these being moral negativity.

Appendix

A.1 Vignettes

Vignette "Packing bags":

Positive deception through implicature:

Mariana's boyfriend Thomas has become abusive, and she plans to leave him. When she thinks Thomas has left for work, she starts packing up her belongings. Thomas, however, hasn't actually left, and he walks into the room where Mariana is packing.

Thomas asks Mariana what she's doing. She says, "I'm remodeling the room."

In packing her belongings Mariana is remodeling the room, but she is doing so because she is leaving Thomas.

Positive deception through assertion:

Mariana's boyfriend Thomas has become abusive, and she plans to leave him. When she thinks Thomas has left for work, she starts packing up her belongings. Thomas, however, hasn't actually left, and he walks into the room where Mariana is packing.

Thomas asks Mariana what she's doing. She says, "The bags are there because I'm remodeling the room."

In packing her belongings Mariana is remodeling the room, but she is doing so because she is leaving Thomas.

Negative deception through implicature:

Mariana is cheating on her boyfriend Thomas, and she plans to leave him for her new lover. When she thinks Thomas has left for work, she starts packing up her belongings. Thomas, however, hasn't actually left, and he walks into the room where Mariana is packing.

Thomas asks Mariana what she's doing. She says, "The bags are there because I'm remodeling the room."

In packing her belongings Mariana is remodeling the room, but she is doing so because she is leaving Thomas.

Negative deception through assertion:

Mariana is cheating on her boyfriend Thomas, and she plans to leave him for her new lover. When she thinks Thomas has left for work, she starts packing up her belongings. Thomas, however, hasn't actually left, and he walks into the room where Mariana is packing.

Thomas asks Mariana what she's doing. She says, "I'm remodeling the room."

In packing her belongings Mariana is remodeling the room, but she is doing so because she is leaving Thomas.

Vignette "Surprise":

Positive deception through implicature:

Today is Derek's birthday. Unbeknownst to Derek, Chris planned a surprise party for him. The party will be in Chris's apartment. Chris needs to bring Derek over without Derek finding out about the party.

Chris tells Derek: "I'd like you to come with me to my apartment. I need to get my jacket."

Chris does need to get his jacket from his apartment, but more important to him is the surprise waiting for Derek there.

Positive deception through assertion:

Today is Derek's birthday. Unbeknownst to Derek, Chris planned a surprise party for him. The party will be in Chris's apartment. Chris needs to bring Derek over without Derek finding out about the party.

Chris tells Derek: "I'd like you to come with me to my apartment because I need to get my jacket."

Chris does need to get his jacket from his apartment, but more important to him is the surprise waiting for Derek there.

Negative deception through implicature:

Chris is being initiated into a violent gang. To impress the gang, Chris planned an ambush to kill Derek. The ambush will be in Chris's apartment. Chris needs to bring Derek over without Derek finding out about the ambush.

Chris tells Derek: "I'd like you to come with me to my apartment. I need to get my jacket."

Chris does need to get his jacket from his apartment, but more important to him is the surprise waiting for Derek there.

Negative deception through assertion:

Chris is being initiated into a violent gang. To impress the gang, Chris planned an ambush to kill Derek. The ambush will be in Chris's apartment. Chris needs to bring Derek

over without Derek finding out about the ambush.

Chris tells Derek: “I’d like you to come with me to my apartment because I need to get my jacket.”

Chris does need to get his jacket from his apartment, but more important to him is the surprise waiting for Derek there.

Vignette ”Earrings”:

Positive deception through implicature:

Laura stayed in Miriam’s guest room during the holidays and accidentally left a pair of earrings behind. Miriam finds the earrings and, as a fun surprise, decides to drive them over to Laura’s house.

When Laura calls Miriam and asks if she has seen them, Miriam says, “Last I saw, the earrings weren’t in my guest room.”

It’s true that Miriam doesn’t see the earrings in her guest room now, but that’s because Miriam herself took them out to bring them to Laura.

Positive deception through assertion:

Laura stayed in Miriam’s guest room during the holidays and accidentally left a pair of earrings behind. Miriam finds the earrings and, as a fun surprise, decides to drive them over to Laura’s house.

When Laura calls Miriam and asks if she has seen them, Miriam says, “I haven’t seen the earrings in my guest room.”

It’s true that Miriam doesn’t see the earrings in her guest room now, but that’s because Miriam herself took them out to bring them to Laura.

Negative deception through implicature:

Laura stayed in Miriam's guest room during Christmas and accidentally left a pair of earrings behind. Miriam finds the earrings and, instead of telling Laura, decides to sell them for money.

When Laura calls Miriam and asks if she has seen them, Miriam says, "Last I saw, the earrings weren't in my guest room."

It's true that Miriam doesn't see the earrings in her guest room now, but that's because Miriam herself took them out to sell them.

Negative deception through assertion:

Laura stayed in Miriam's guest room during the holidays and accidentally left a pair of earrings behind. Miriam finds the earrings and, instead of telling Laura, decides to sell them for money.

When Laura calls Miriam and asks if she has seen them, Miriam says, "Last I saw, the earrings weren't in my guest room."

It's true that Miriam doesn't see the earrings in her guest room now, but that's because Miriam herself took them out to sell them.

Vignette "At the door":

Positive deception through implicature:

Ellie finds Charlie knocking at her door on a Thursday evening. Charlie is trying to locate his estranged wife, Brenda, whom he is abusive towards.

Charlie asks Ellie where Brenda is. Ellie says, "Every Thursday at 4:00 she has to teach a class at 245 Temple St."

Brenda does usually teach a class at 245 Temple St. on Thursdays. However, today she

cancelled class. She is actually hiding from Charlie in another room of Ellie's house.

Positive deception through assertion:

Ellie finds Charlie knocking at her door on a Thursday evening. Charlie is trying to locate his estranged wife, Brenda, whom he is abusive towards.

Charlie asks Ellie where Brenda is. Ellie says, "She is teaching her class at 245 Temple St."

Brenda does usually teach a class at 245 Temple St. on Thursdays. However, today she cancelled class. She is actually hiding from Charlie in another room of Ellie's house.

Negative deception through implicature:

Ellie finds Brenda knocking at her door on a Thursday evening. Brenda is trying to escape from her abusive husband Charlie.

Brenda asks Ellie where Charlie is. Ellie says, "Every Thursday at 4:00 he has to teach a class at 245 Temple St."

Charlie does usually teach a class at 245 Temple St. on Thursdays. However, today he cancelled class. He's actually in another room of Ellie's house, waiting for Brenda.

Negative deception through assertion:

Ellie finds Brenda knocking at her door on a Thursday evening. Brenda is trying to escape from her abusive husband Charlie.

Brenda asks Ellie where Charlie is. Ellie says, "He is teaching his class at 245 Temple St."

Charlie does usually teach a class at 245 Temple St. on Thursdays. However, today he cancelled class. He's actually in another room of Ellie's house, waiting for Brenda.

Vignette "School test":

Positive deception through implicature:

Rachel and Pauline had a test on Thursday. Rachel studied very hard, but Pauline didn't study at all. As a result, Rachel performed very well on the test, and Pauline performed very poorly.

When their friends ask her how she did, Rachel, not wanting to embarrass Pauline, says: "I got some of the questions right."

In fact, Rachel got all of the questions right.

Positive deception through assertion:

Rachel and Pauline had a test on Thursday. Rachel studied very hard, but Pauline didn't study at all. As a result, Rachel performed very well on the test, and Pauline performed very poorly.

When their friends ask Rachel how she did, Rachel, not wanting to embarrass Pauline, says: "I got some of the questions wrong."

In fact, Rachel got all of the questions right.

Negative deception through implicature:

Rachel and Pauline had a test on Thursday. Pauline studied very hard, but Rachel didn't study at all. As a result, Pauline performed very well on the test, and Rachel performed very poorly.

When their friends ask Rachel how she did, Rachel, embarrassed by her score, says: "I got some of the questions wrong."

In fact, Rachel got all of the questions wrong.

Negative deception through assertion:

Rachel and Pauline had a test on Thursday. Pauline studied very hard, but Rachel didn't study at all. As a result, Pauline performed very well on the test, and Rachel performed very poorly.

When their friends ask Rachel how she did, Rachel, embarrassed by her score, says: "I got some of the questions wrong."

In fact, Rachel got all of the questions wrong.

Vignette "Fake ID":

Positive deception through implicature:

Daniel had to enter a witness protection program because he told the police about corruption at his former workplace. As part of the program, he is given a new identity.

When applying for a job under the new identity, his potential employer asks him for his date of birth. Daniel tells him, "Here are my papers showing that I was born on January 3rd, 1989."

This is the birth date associated with his new identity; he was actually born on December 11th, 1989.

Positive deception through assertion:

Daniel had to enter a witness protection program because he told the police about corruption at his former workplace. As part of the program, he is given a new identity.

When applying for a job under the new identity, the potential employer asks him for his date of birth. Daniel tells him, "I was born on January 3rd, 1989."

This is the birth date associated with his new identity; he was actually born on December 11th, 1989.

Negative deception through implicature:

Daniel killed a man and is on the run from the authorities. He pays to have a new identity developed.

When applying for a job under the new identity, his potential employer asks him for his date of birth. Daniel tells him, "Here are my papers showing that I was born on January 3rd, 1989."

This is the birth date associated with his new identity; he was actually born on December 11th, 1989.

Negative deception through assertion:

Daniel killed a man and is on the run from the authorities. He pays to have a new identity developed.

When applying for a job under the new identity, his potential employer asks him for his date of birth. Daniel tells him, "I was born on January 3rd, 1989."

This is the birth date associated with his new identity; he was actually born on December 11th, 1989.

A.2 Informed Consent Form

Purpose: We are conducting a research study to examine people's ordinary intuitions.

Procedures: Participation in this study will involve filling out a brief questionnaire. We anticipate that your involvement will require approximately three minutes. You will be compensated financially for participating. Participants must be at least 18 years of age to

participate.

Risks and Benefits: You will be paid for participation through Amazon's Mechanical Turk. Further, we hope that our results will add to scientific knowledge about people's ordinary social intuitions.

Confidentiality: All of your responses will be anonymous. Only the researchers involved in this study and those responsible for research oversight will have access to the information you provide.

Voluntary Participation: Participation in this study is completely voluntary. You are free to decline to participate, to end participation at any time for any reason, or to refuse to answer any individual question without penalty.

Questions: If you have any questions about this study, you may contact the investigator, Anna Martinelli-Parker at anna.martinelli-parker@yale.edu. If you would like to talk with someone other than the researchers to discuss problems or concerns, to discuss situations in the event that a member of the research team is not available, or to discuss your rights as a research participant, you may contact the Yale University Human Subjects Committee, Box 208252, New Haven, CT 06520-8252, 203-436-3650, human.subjects@yale.edu. Additional information is available at http://www.yale.edu/hsc/Subject/subjectsrights.html.

Do you voluntarily consent to participate in this study?

Bibliography

Adler, Jonathan. 1997-09. Lying, deceiving, or falsely implicating. *The Journal of Philosophy* 94(9). 435–452. doi:https://doi.org/10.2307/2564617.

Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations*. University of Chicago Press.

Antomo, Mailin, Susanne Müller, Katharina Paul, Markus Paluch and Maik Thalmann. 2018. When children aren't more logical than adults: An empirical investigation of lying by falsely implicating. *Journal of Pragmatics* 138. 135–148. doi:https://doi.org/10.1016/j.pragma.2018.09.010. URL `https://www.sciencedirect.com/science/article/pii/S0378216618302479`.

Aquinas, St. Thomas. c. 1270. *Summa theologica*. Project Gutenberg.

Arico, Adam and Don Fallis. 2013-01-15. Lies, damned lies, and statistics: An empirical investigation of the concept of lying. *Philosophical Psychology* 26(6). 790–816. doi:https://doi.org/10.1080/09515089.2012.725977.

Coleman, Linda and Paul Kay. 1981-03. Prototype semantics: The english word lie. *Language* 57(1). 26–44. doi:doi.org/10.1353/lan.1981.0002.

Davis, Wayne. 2019. Implicature. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University, Fall 2019 edn.

Dynel, Marta. 2015. Intention to deceive, bald-faced lies, and deceptive implicature: Insights into lying at the semantics-pragmatics interface. *Intercultural Pragmatics* 12(3). 309–332. doi:doi:10.1515/ip-2015-0016. URL `https://doi.org/10.1515/ip-2015-0016`.

Faulkner, Paul. 2013. *Lying and deceit*. John Wiley Sons, Ltd. doi:https://doi.org/10.1002/9781444367072.wbiee482. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444367072.wbiee482`.

García-Carpintero, Manuel. 2023-01-21. Lying versus misleading, with language and pictures: the adverbial account. *Linguistics and Philosophy* 46(3). 509–532. doi:https://doi.org/10.1007/s10988-022-09355-0.

Gotzner, Nicole and Diane Mazzarella. 2021-09-27. Face management and negative strengthening: The role of power relations, social distance, and gender. *Frontiers in Psychology* 12. doi:doi.org/10.3389/fpsyg.2021.602977.

Grice, H. P. 1975. Logic and conversation. In Donald Davidson and Gilbert Harman (eds.), *The logic of grammar*, 64–75.

Grice, Herbert Paul. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.

Horn, Laurence. 2017. *Telling it slant: Toward a taxonomy of deception*. De Gruyter Mouton.

Johnson, Robert and Adam Cureton. 2022. Kant's Moral Philosophy. In Edward N. Zalta and Uri Nodelman (eds.), *The Stanford encyclopedia of philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edn.

Knobe, Joshua, Sandeep Prasada and George E. Newman. 2013. Dual character concepts and the normative dimension of conceptual representation. *Cognition* 127(2). 242–257. doi:https://doi.org/10.1016/j.cognition.2013.01.005. URL https://www.sciencedirect.com/science/article/pii/S0010027713000139.

Krstić, Vladimir. 2019. Can you lie without intending to deceive? *Pacific Philosophical Quarterly* 100(2). 642–660.

Mahon, James. 2016-10-10. The definition of lying and deception. *Stanford Encyclopedia of Philosophy* URL https://plato.stanford.edu/archives/win2016/entries/lying-definition/.

Mannison, D.S. 1969. Lying and lies. *Australasian Journal of Philosophy* 47(2). 132–144. doi:10.1080/00048406912341141. URL https://doi.org/10.1080/00048406912341141.

Mazzarella, Diana, Robert Reinecke, Ira Noveck and Hugo Mercier. 2018. Saying, presupposing and implicating: How pragmatics modulates commitment. *Journal of Pragmatics* 133. 15–27. doi:https://doi.org/10.1016/j.pragma.2018.05.009. URL https://www.sciencedirect.com/science/article/pii/S0378216617308445.

Rett, Jessica. 2020. Manner implicatures and how to spot them. *International Review of Pragmatics* 12(1). 44 – 79. doi:10.1163/18773109-01201105. URL https://brill.com/view/journals/irp/12/1/article-p44_2.xml.

Saul, Jennifer. 2012. *Lying, misleading, and what is said: An exploration in philosophy of language and ethics*. Oxford University Press.

Strichartz, Abigail F. and Roger V. Burton. 1990. Lies and truth: A study of the development of the concept. *Child Development* 61(1). 211–220. URL http://www.jstor.org/stable/1131060.

Sweetser, Eve E. 1987. *The definition of lie: An examination of the folk models underlying a semantic prototype*, 43–66. Cambridge University Press.

Thomas, Carson. 2010. *Lying and deception: Theory and practice*. Oxford University Press.

Timmermann, Felix and Emanuel Viebahn. 2020-06-15. To lie or to mislead? *Philosophical Studies* 178(5). 1481–1501. doi:doi.org/10.1007/s11098-020-01492-1.

Turri, Angelo and John Turri. 2015-05. The truth about lying. *Cognition* 138. 161–168. doi:doi.org/10.1016/j.cognition.2015.01.007.

Weissman, Benjamin and Marina Terkourafi. 2018. Are false implicatures lies? an empirical investigation. *Mind & Language* 34(2). 221–246. doi:https://doi.org/10.1111/mila.12212. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/mila.12212.

Wiegmann, A, S Willemsen, G Gunzelmann, A Howes, T Tenbrink and EJ Davelaar. 2017. Proceedings of the 39th annual conference of the cognitive science society. Cognitive Science Society.

Wiegmann, Alexander. 2023-08. Does lying require objective falsity? *Synthese* 202. doi:https://doi.org/10.1007/s11229-023-04291-3.

Wiegmann, Alexander and Jörg Meibauer. 2019-07-31. The folk concept of lying. *Philosophy Compass* 14(8). doi:doi.org/10.1111/phc3.12620.

Wiegmann, Alexander, Jana Samland and Michael Waldmann. 2016-05. Lying despite telling the truth. *Cognition* 150. 37–42. doi:doi.org/10.1016/j.cognition.2016.01.017.

Wiegmann, Alexander, Pascale Willemsen and Jörg Meibauer. 2022. Lying, deceptive implicatures, and commitment. *Ergo* 8(50). 709–740. doi:https://doi.org/10.3998/ergo.2251.