

On Deniability

Alexander Dinges & Julia Zakkou

April 4, 2021

Strategic speakers often convey their messages by insinuation or innuendo, or by using so-called code words or dogwhistles. Why? Many hold that retaining deniability is one key factor here: people often use these indirect forms of communication because they want to retain deniability. In this paper, we shed light on various questions and puzzles that surround the notion of deniability at issue by offering an account of deniability. On our account, deniability is an epistemic phenomenon. A speaker has deniability if she can make it *epistemically* irrational for her audience to *reason* in certain ways. To avoid predictable confusion, we distinguish deniability from a practical correlate we call untouchability. Roughly, a speaker has untouchability if she can make it *practically* irrational for her audience to *act* in certain ways.

1 Introduction

Strategic speakers often convey their messages by insinuation or innuendo, or by using so-called code words or dogwhistles. One important question is how these practices work. For instance, how do hearers retrieve merely insinuated contents, and how does the use of code words trigger cognitive effects in the audience? An equally important question is why speakers engage in these practices to begin with. Why insinuate or use code words when you could say it outright? Many people hold that retaining deniability is one key factor here: people often use indirect forms of communication such as insinuation or code words because they want to retain deniability.¹

Various questions and puzzles surround the notion of deniability at issue. First, speakers who have deniability leave their audience in a ‘frustrating position’ (Camp, 2018, p. 46), because the audience cannot pin the speakers down to a communicated content and criticize them on this basis. This raises the pressing practical question of how we can avoid crediting our interlocutors with more deniability than necessary. Second, many people hold that there are cases of ‘implausible deniability’ alongside cases of ‘plausible deniability’. The former cases are puzzling because, in these cases, a speaker seemingly has deniability despite the fact that the audience knows what they meant to convey (Camp, 2018, p. 52; Berstler, 2019, pp. 27–28). Finally, there is an apparent tension between the ubiquity of deniability and the widespread availability of testimonial knowledge (Fricker, 2012; Peet, 2018; Davies, 2019).

In this paper, we offer an account of deniability that sheds light on these issues. We suggest that deniability is an epistemic notion. A speaker has deniability if she can make it *epistemically* irrational for her audience to *reason* in certain ways. To address putative cases of implausible deniability, we distinguish this epistemic notion from a practical correlate we call untouchability. Roughly, a speaker has untouchability if she can make it *practically* irrational for her audience to *act* in certain ways. Both accounts are theoretically neutral

¹ See e.g. Walton, 1996; Pinker, Nowak, and Lee, 2008; Lee and Pinker, 2010; Fricker, 2012, p. 25; Camp, 2018, p. 44 on insinuation and e.g. Jamieson, 1992, pp. 84–93; Mendelberg, 2001; Stanley, 2015, p. 156; Khoo, 2017, p. 47; Saul, 2018, p. 365 on code words.

and can be taken on board by scholars from different camps. For instance, the suggested account of deniability is compatible with different views of the epistemic norms of reasons, and it is impartial towards the question of whether knowledge is defeasible. The two accounts still have substantial theoretical consequences in that they help to resolve the questions and puzzles above.²

A preliminary methodological remark. The term ‘deniability’ may have different uses inside and outside of philosophy. We focus here on only one usage, the one which features in the above explanation of why strategic speakers prefer some means of communication over others. Since this usage is at least semi-technical, we cannot exclusively rely on ordinary intuition. So, instead, we focus on the theoretical role that the notion of deniability is supposed to play. More specifically, our primary goal is to capture general desiderata that one must respect if one wants to uphold the indicated explanatory relation between deniability and strategic speech. We are optimistic that, besides this theoretical role, our notion of deniability can play further theoretical roles in adjacent debates. For instance, van Elswyk (2020, p. 1168) suggests to use deniability as a criterion to distinguish lying from misleading, Berstler (2019, p. 27) and Pepp (2018, p. 842) float competing ideas on how deniability grounds moral differences between lying and misleading, Reins and Wiegmann (2021) use deniability as a measure of commitment, and according to Saul (2017, p. 109) figleaves can be used to provide deniability. We leave a thorough discussion of these further theoretical roles for another occasion.³

Our paper is structured as follows. We begin with a couple of sample cases in which a speaker’s desire to retain deniability drives their use of indirect means of communication. On this basis, we specify the indicated theoretical role of the notion of deniability (§2). We uncover what one might call the logical form of deniability statements (§3) and state desiderata on adequate accounts of deniability (§4). We briefly discuss accounts of deniability that we find wanting (§5) and then offer our own account (§6). We suggest a generalization of our account (§7) before we address cases of implausible deniability (§8). We conclude with some tentative lessons on how to limit the scope of deniability we grant to our interlocutors (§9).

2 The Theoretical Role of Deniability

Let us begin with four sample cases taken from the literature, where the desire to retain deniability supposedly drives indirect speech. We stipulate that, in each of these cases, the indicated denial would be a lie and thus that the speaker has the respective communicative intention. In the first two cases, the speakers indirectly offer a bribe.

Speeding Driver: *A driver stopped for speeding, to the police officer.*

‘I’m in a bit of a hurry. Is there any way we can settle this right now?’

Denial: ‘I didn’t mean to offer a bribe! (I was just wondering if I could pay my fine right away.)’ (See Lee and Pinker, 2010, p. 790.)

Restaurant: *A customer waiting to be seated, to the waiter.*

‘Is there any way to shorten my wait?’

Denial: ‘I didn’t mean to offer a bribe! (I was just wondering if you could e.g. split us up on separate tables.)’ (See Lee and Pinker, 2010, p. 794.)

In the next two cases, the speakers are making implicit suggestions or implicit statements.

²On our account, both deniability and untouchability come out as distinct from the familiar notion of cancellability; see below.

³For a notion of deniability that is clearly distinct from ours, see e.g. Viebahn, 2020, p. 733.

Realtor: *A realtor, to a potential buyer from a different racial background.*

‘Perhaps you would feel more comfortable locating in a more . . . transitional neighborhood, like Ashwood?’

Denial: ‘I didn’t mean to convey that you don’t fit here. I only meant that Ashwood is on an upward trajectory and might be more interesting to you for this reason.’ (See Camp, 2018, p. 43.)

Inner City: *A politician in a press briefing.*

‘I oppose the food stamp program because it primarily benefits inner-city Americans’

Denial: ‘I didn’t mean to convey anything racist. I just think that the urban poor already receive enough governmental assistance.’ (See Khoo, 2017, p. 40.)

In all four cases the speakers could have conveyed their messages directly. Why didn’t they? Why didn’t the speakers of the first two cases directly offer the bribe, by saying for instance ‘Could we settle this with a bribe?’, and why didn’t the speakers of the third and fourth case make explicitly racist remarks?

Various things might play a role. We run with the following assumption, which is both plausible and widely accepted. The speakers speak the way they speak partly because they want to retain deniability regarding the claim that they meant to bribe their interlocutor or that they meant to suggest something racist.⁴ In what follows, we develop an account of deniability that allows deniability to play this explanatory role.

Before we move on, notice a couple of assumptions we do *not* make. First, we do not assume that a desire to retain deniability is the only reason one can have for speaking indirectly (see e.g. Camp, 2018, p. 44 for a brief overview of other possible reasons). We merely assume that deniability is the key factor in cases like the ones above. Second, we do not assume that the direct-indirect distinction maps onto the deniable-undeniable distinction. Indeed, our analysis will reveal that directly communicated contents are sometimes deniable while indirectly communicated contents are sometimes undeniable. Finally, we do not assume that the speakers in the above cases have deniability. For now, we only assume that they are driven by a desire to retain deniability. More on this below.

3 Logical Form

This section uncovers what one might call the logical form of deniability statements. We suggest that deniability is relative to propositions of a specific type, to two temporal indices, to a part of the audience and to a possible world. Drawing these parameters out into the open helps to avoid conceptual confusions, as will become apparent below.

First, deniability is relative to propositions. The speeding driver, for instance, could say ‘I didn’t mean to offer a bribe’ and if this denial sticks in the relevant way, then she has deniability relative to the proposition she denies, i.e., the proposition that she meant to offer a bribe. She might still lack deniability relative to, say, the proposition that she meant to ask whether the police officer is willing to settle things right now. This gives us the following preliminary form.

LF1 S has deniability relative to the proposition p.

For the purposes of this paper, we assume that deniability only applies to propositions about one’s communicative intentions. Thus, we assume that deniability statements have the following, more specific logical form. In this form, ϕ stands for the description of a speech act,

⁴See footnote 1 for references.

broadly construed, such as offering a bribe, asking a question, insulting someone, conveying something, etc.

LF2 S has deniability relative to the proposition that she meant to ϕ .⁵

Let us flag though that deniability might target other propositions too. For instance, it might be possible to have deniability relative to propositions about communicative acts rather than communicative intentions, such as offering a bribe as opposed to meaning to offer a bribe. Arguably, deniability is not even restricted to the communicative realm. For instance, Khoo (2017, p. 38) following Mendelberg (2001) discusses ‘deniable norm violations’. All accounts of deniability we discuss below straightforwardly allow for such extensions, as readers can verify for themselves.⁶

Second, deniability is doubly time-relative. Both the embedded proposition and deniability itself have a temporal index along the following lines.

LF3 S has deniability at time t relative to the proposition that she meant to ϕ at time t' .

Consider the speeding driver again. She might have deniability at t_1 relative to the proposition that she meant to offer a bribe at t_0 , but once she admits, at t_2 , ‘Ok, I did mean to offer a bribe, and I am terribly sorry. But . . .’ she will have lost deniability at t_2 relative to the same proposition.

Third, deniability is relative to parts of the audience along the following lines.

LF4 S has deniability at time t relative to the proposition that she meant to ϕ at time t' and relative to a part a of the audience she has at t .

A politician at a press conference, for instance, has the press as her audience but also the wider public. She may lack deniability relative to the former while having deniability relative to the latter. Depending on one’s favored account of deniability, this might be because, say, press correspondents know much about the conversational context while the wider public is generally less knowledgeable.

Finally, deniability is relative to a possible world. Speakers may lack actual deniability while retaining counterfactual deniability.

LF5 S has deniability at possible world w and time t relative to the proposition that she meant to ϕ at time t' and relative to a part a of the audience she has at w and t .

Consider the speeding driver. She might have no deniability in her actual conversation with the police officer. Arguably, though, she would have deniability if she faced a judge in court should the police officer arrest her. This might be because the judge ‘could not make a bribery charge stick in court by the high standard of proof beyond a reasonable doubt’ (Lee and Pinker, 2010, p. 790) or because the judge ‘lacks full access to the immediate context’ (Camp, 2018, p. 50).

In what follows, we will mostly leave the world and the time indices as well as the relativizations to parts of the audience implicit and work with the condensed form of LF2. All parameters mentioned in this section will occasionally become relevant though.

⁵Camp (2018, pp. 44–45) similarly suggests that deniability targets combinations $M(Q)$ of a ‘proposition, Q , in a mode M : as a contribution of information, a query, a directive, etc.’

⁶Deniability 4 below, the fourth account discussed, is the only exception, and this might be objectionable. We will not push this worry here though.

4 Desiderata

This section presents desiderata on accounts of deniability. We start with three key desiderata. They result from the idea that a desire to retain deniability partly explains why the speakers in the above example cases—henceforth, the *target cases*—choose to speak indirectly in the way they do.

A speaker's desire to retain deniability can explain why she chooses to speak a certain way only if the speaker actually has this desire. And she presumably has this desire only if deniability is *desirable*. This gives us the first desideratum on adequate accounts of deniability.

Desideratum 1 In the target cases, deniability is desirable.

The desire to retain deniability by itself does not explain why speakers speak indirectly rather than directly unless speaking indirectly is a means to fulfil this desire while speaking directly is not.⁷ This gives us the following desiderata.

Desideratum 2 In the target cases, speaking directly isn't a means to retain deniability.

Desideratum 3 In the target cases, speaking indirectly is a means to retain deniability.⁸

Note that all these desiderata specifically concern the target cases. As indicated, we assume neither that speaking directly never grants you deniability nor that speaking indirectly always grants you deniability.

If we respect the previous relativizations, we can state the first three desiderata more precisely as follows:

There is some possible world w , moment in time t and t' , speech act ϕ and part a of the speaker's audience at w and t such that:

Desideratum 1' In the target cases, deniability relative to w , t , ϕ , t' and a is desirable.

Desideratum 2' In the target cases, speaking directly isn't a means to retain deniability relative to w , t , ϕ , t' and a .

Desideratum 3' In the target cases, speaking indirectly is a means to retain deniability relative to w , t , ϕ , t' and a .

To illustrate, the speeding driver might be aware that she cannot retain deniability in her present conversation with the police officer and so she might aim only at future deniability relative to e.g. a judge. This suffices to explain her choice to speak indirectly if indirect but not direct speech allows her to reach this goal. Lee and Pinker (2010, p. 796) similarly suggest that indirect communication may be motivated already when a denial is 'plausible to a virtual audience, even if it is plausible to neither the speaker nor the hearer'. We will consider Desiderata 1' to 3' when precision matters, but for the most part, we will work with the simplified Desiderata 1 to 3.

Note that Desideratum 3' (and similarly Desideratum 3) entails that the speakers in the target cases have deniability relative to some relevant set of parameters. For taking a means

⁷Or if both speaking directly and speaking indirectly are means to retain deniability, speaking indirectly must be a *better* means to retain deniability than speaking directly is. We leave this qualification implicit for simplicity.

⁸More precisely: In the target cases, speaking in the way the respective speaker speaks (i.e. using the respective form of indirect communication) is a means to retain deniability.

to an end normally suffices to achieve this end. And, by speaking indirectly, the speakers in the target cases take a means to the end of retaining deniability, according to Desideratum 3'. For now, we run with the assumption that the speakers in the target cases have relevant deniability, but we qualify this assumption below.

Let us turn to secondary desiderata on accounts of deniability. We start with factors on which deniability notably does not depend and then turn to factors on which deniability notably does depend.

First, consider the following familiar case from Grice (1989).

Petrol: *A towards B, who has run out of gas.*

'There is a garage around the corner'

Denial: 'I didn't mean to suggest that you might get petrol there. I only meant that you could get a newspaper to pass the time.'

Initially, A seems to have deniability vis-à-vis the proposition that she meant to convey that the garage might sell petrol. If A issued the indicated denial right away, B would (and should) believe her and grant that A did not have the respective communicative intention.^{9,10} Meanwhile, if A does not issue the denial, B knows that A meant to convey that the garage might sell petrol; she will at least act as if she knew this, heading off to the garage right away without ascertaining A's communicative intentions.¹¹ It thus seems that a speaker can have deniability about whether she meant to ϕ even though her audience knows that she meant to ϕ . This gives us the following desideratum.

Desideratum 4 Deniability is consistent with the audience knowing the speaker's communicative intentions, in cases where the respective denial is not actually made.

To support this desideratum further, notice that many people hold that deniability is widespread, being a relatively general feature of indirect communication and maybe even of some types of direct communication involving context-sensitive expressions (e.g. Fricker, 2012 and Peet, 2015). Unless we accept the indicated desideratum, this implies radically skeptical consequences for testimonial knowledge. After all, if a hearer does not know that a speaker meant that p, it is hard to see how she can come to know that p based on the speaker's testimony (see also Davies, 2019, pp. 23–24).¹²

Second, 'deniability' and 'plausible deniability' are sometimes used interchangeably. However, almost everybody in the debate agrees that even if a speaker has deniability, her denials need not be plausible. Many authors hold, for instance, that the speakers in Speeding Driver or Realtor from above have deniability despite the fact that the relevant denial would

⁹Of course, B may lose deniability over time, e.g., if she observes A take off with a canister and issues the denial only after A's empty-handed return. But she still has deniability at the outset (recall the temporal indices).

¹⁰Lee and Pinker (2010, p. 791) and Camp (2018, p. 45) sometimes seem to suggest that deniability entails a prior strategic intention to deny the target message under certain circumstances. Since B lacks any such intention, they might claim that she lacks deniability. However, we suggest distinguishing whether a speaker plans to deny her communicative intentions from whether this will relevantly succeed. Deniability, as we understand it, concerns the latter question.

¹¹One may question this verdict depending on one's favored account of knowledge. What ultimately matters to us is that B meets whatever epistemic standard suffices for actionability in the context at hand, be it knowledge or something else. See below.

¹²One might object that testimonial knowledge does not require knowing *exactly* what the speaker meant. To avoid such concerns, one can just restate Desideratum 4 in terms of knowledge of what the speaker *roughly* meant and adjust the rest accordingly. One may object that testimonial knowledge does not even require such rough knowledge (e.g. Peet, 2018). We cannot address this worry here, except to note that this is a minority position (as Peet would certainly agree).

be quite implausible (e.g. Lee and Pinker, 2010, p. 793; Camp, 2018, p. 48; Berstler, 2019, pp. 27–28). We get the following desideratum.

Desideratum 5 Deniability is consistent with the audience finding the denial implausible, in cases where the respective denial is actually made.

By granting this desideratum, we grant that there are cases of ‘implausible deniability’ in one sense of the term. Usually, though, cases of ‘implausible deniability’ are construed as cases where the respective denial is not just implausible but mutually known to be false. For now, we stay neutral on whether there are cases of implausible deniability in this latter sense, but we reject this assumption below.

So much for factors on which deniability does not depend. Let us now turn to some interesting factors on which deniability does depend.

Whether a speaker has deniability depends on the whole range of factors relevant for Gricean calculability, such as the sentence’s conventional meaning, background knowledge and expectations about cooperativeness (Grice, 1989). This is obvious enough. Interestingly, deniability also depends on factors that usually have no place in Gricean calculations.

First, Peet (2015, p. 48) suggests that ‘repeat offenders’ who continuously mislead about their communicative intentions may lose deniability. They presumably lose deniability because we no longer trust them when they deny their communicative intentions. This gives us the following desideratum.

Desideratum 6 Deniability can vary with the trustworthiness of the speaker when it comes to reports about her communicative intentions.

Two clarificatory remarks. First, we can trust speakers in one domain but not another. For instance, I trust the weather-caster when it comes to the weather but not when it comes to politics. On the above desideratum, deniability depends on trust *concerning reports about communicative intentions*. It might not depend on the general trustworthiness of the speaker or her trustworthiness in other domains.¹³ Second, the above desideratum entails that a loss of trust can lead to a loss of deniability. It does not entail, however, that deniability requires even a minimal amount of trust. In some contexts, you may have deniability while being completely untrustworthy, say, because your communicative intentions are obscure independently of whether you deny them.

Second, sometimes (not always!) we have deniability because we can claim that we misspoke or that the audience misheard us (see e.g. García-Carpintero, 2018, pp. 203–204 and Berstler, 2019, p. 28). This strategy becomes more efficient when a speaker e.g. mumbles or chooses a noisy environment (e.g. by standing next to a landing chopper). We get the following desideratum.

Desideratum 7 Deniability can vary with presentational factors (e.g. mumbling) and environmental factors (e.g. noise) bearing on the probability of misspeaking and mishearing.

This concludes our collection of desiderata. We now turn to an assessment of some candidate accounts of deniability. We think they fail because they do not respect the desiderata outlined.

¹³Desideratum 6 is thus consistent with Camp’s (2018, p. 51) observation that ‘a known likelihood of conflict or other motivation for strategic interpretation [...] widens the scope of deniability’ even though an increased ‘likelihood of conflict or other motivation for strategic interpretation’ presumably leads to lesser trust overall. Lesser trust *overall* can widen the scope of deniability even if lesser trust *regarding communicative intentions* narrows down the scope of deniability.

5 Accounts

For starters, consider the following definition of deniability (see e.g. Mazzarella et al., 2018, p. 16).

Deniability 1 S has deniability relative to the proposition that she meant to ϕ iff S can deny that she meant to ϕ .

This definition is clearly problematic unless we specify the relevant sense of ‘can’. After all, the ‘can’ of ability would credit everybody who is able to communicate with universal deniability, because competent communicators generally have the ability to deny their communicative intentions by simply producing utterances of the form ‘I didn’t mean to ϕ .’ This is at odds with Desideratum 2, according to which speaking directly is not a means to achieve deniability in the target cases.

Consider the following restriction of Deniability 1.

Deniability 2 S has deniability relative to the proposition that she meant to ϕ iff S can plausibly deny that she meant to ϕ .

This definition is untenable as well because it directly clashes with Desideratum 5, according to which deniability does not entail plausibility.

Consider then the following alternative restriction of Deniability 1.

Deniability 3 S has deniability relative to the proposition that she meant to ϕ iff S can properly deny that she meant to ϕ .

This definition is problematic because it clashes with Desideratum 3. Lies are improper because they violate the norm of assertion. According to the knowledge norm of assertion, for instance, one can properly assert p only if one knows p (Williamson, 2000), and in cases of lying, this condition is clearly violated. Now, in the target cases, the denials in question are lies because the speaker has the respective communicative intentions (by assumption). Consequently, these denials are improper. Given Deniability 3, the speakers lack deniability. As argued above, however, Desideratum 3 entails that they have deniability. Analogous considerations hold for familiar alternatives to the knowledge norm of assertion, such as justified belief norms (e.g. Lackey, 2007) or truth norms (e.g. Weiner, 2005).¹⁴

We have now considered some immediate attempts to define deniability. Let us turn to a more elaborate and more promising account from the literature, which is extracted from Camp (2018) and adapted to our terminology. Peet (2015, p. 31), following Fricker (2012), offers a similar definition (see also Davies, 2019, p. 19).

Deniability 4 S has deniability relative to the proposition that she meant to ϕ iff it is ‘reasonable to calculate [that the speaker meant to ψ ($\neq \phi$)] on the basis of the uttered sentence’s conventional meaning [...], the commitments undertaken in the conversation to this point, and some set [π]’ of epistemically accessible presuppositions consistent with those commitments [...].’ (Camp, 2018, p. 50)

¹⁴One could propose weaker understandings of propriety to respond. But such notions would have to be spelled out, while bearing in mind Desideratum 2. Berstler (2019, p. 28), for instance, can be taken to suggest a comparative notion of propriety, where deniability entails that the respective denial ‘sounds better’ than a straightforward denial of a literally expressed communicative intention. As she notices herself, however, this would make deniability generally available even in direct communication, through utterances of e.g. ‘I misspoke, I meant to ψ ’, contrary to Desideratum 2. To be sure, Berstler might have other theoretical roles for her notion of deniability in mind than we have, and for these purposes, her notion might be fine.

Very roughly, according to this account, a speaker has deniability when their utterance is open to more than one interpretation, where this openness is a matter of the ‘epistemic accessibility’ of pertinent contextual background assumptions. This definition makes progress because the previously problematic desiderata now seem satisfied. Indirect as opposed to direct communication often makes alternative interpretation epistemically accessible even if it does not make them particularly plausible.

Problems remain. First, this definition focuses exclusively on Gricean context factors. It is thus at odds with Desideratum 6, according to which deniability can vary with the trust we place in the speaker’s denials, and Desideratum 7, according to which factors bearing on mis-speaking and mishearing affect deniability. For, whether a speaker’s denials are trustworthy relevantly affects neither the ‘epistemically accessible’ presuppositions nor the implicatures one can ‘reasonably’ calculate on their basis, and whether a speaker e.g. mumbles or whether it is e.g. noisy does not relevantly affect these factors either.

Further problems arise depending on how we interpret ‘epistemic accessibility’. Here is a first interpretation.

EA1 A set of presuppositions π' is *epistemically accessible* iff it is consistent with what the audience knows that presuppositions π' govern the present context.

On this interpretation of epistemic accessibility, we face the following problem. Whenever I have deniability on the given definition, my audience does not know what I meant. For on the given definition, if I have deniability, my audience does not know whether the context is governed by presuppositions π that make it reasonable to calculate what I meant or by some other presuppositions π' that make an entirely different intention reasonable to calculate. This is at odd with Desideratum 4, according to which hearers may have such knowledge despite deniability.

Consider the following, alternative interpretation of epistemic accessibility.

EA2 A set of presuppositions π' is *epistemically accessible* iff the audience cannot rule out that presuppositions π' govern the present context.

This interpretation irons out the previous problem. For even if the audience cannot rule out that the context is governed by presuppositions π' , she might still know that the context is governed by presuppositions π and thereby know what the speaker meant. On a relevant alternatives account of knowledge, for instance, one has to rule out only the epistemically *relevant* error-possibilities to p in order to know p (e.g. Dretske, 1970). And the possibility that presuppositions π' govern the context might be epistemically irrelevant.

But now we face a different problem. It becomes difficult to explain Desideratum 1, according to which deniability is desirable in the target cases. For if deniability is consistent with the audience *knowing* one’s communicative intentions, then why would deniability be desirable? If you know I had e.g. racist intentions, you can blame me for this, you can tell others about this, and you can generally let your actions be guided by this. So what do I gain from retaining deniability? This is a challenge rather than an objection, but so far, we do not see how the proponent of Deniability 4 can meet this challenge.

6 Our Proposal

We have proposed some desiderata on accounts of deniability, and we have shown that it is difficult to meet them. It is time to turn to our own favored analysis. We begin with a simplified proposal. We use this proposal to illustrate some basic features of our official account, which we present below.

Deniability S has deniability relative to the proposition that she meant to ϕ iff: if S denies that she meant to ϕ , then S's audience does not know that she meant to ϕ .¹⁵

On this account, there are two ways to have deniability. According to the first, you have deniability if your audience does not know what you meant to begin with.¹⁶ According to the second, you have deniability if your denial defeats otherwise present knowledge.¹⁷

The suggested account easily accommodates our secondary desiderata. Desideratum 4 ('knowledge'): On the given definition, deniability only entails that the audience lacks knowledge *if* the speaker issues a denial. If she does not, the audience may know what was meant. Desideratum 5 ('plausibility'): On the given definition, the audience does not know a respective denial is false in cases of deniability. They might still find the denial implausible. After all, we consider many propositions implausible even though we do not know that they are false. Desideratum 6 ('trust'): If you are trustworthy, your denials have more evidential weight and are therefore more likely to undermine my knowledge. This makes it more likely that you have deniability on the given definition. Desideratum 7 ('misspeaking and mishearing'): If you e.g. mumble or speak in a noisy environment, it is more likely that your audience does not know what you said and therefore what you meant. Again, this makes it more likely that you have deniability.

Let us now turn to our key desiderata, which deserve an extended discussion. We begin with Desideratum 1, according to which deniability is desirable in the target cases. Many people assume a knowledge norm of action along the following lines (e.g. Hawthorne and Stanley, 2008).

KN-A It is epistemically proper to treat p as a reason for action iff one knows p .

For instance, I can properly treat the proposition that you will be at the party as a reason to go to the party only if I know that you will be at the party. Given KN-A, it is clear why deniability is often desirable on the given definition. If I have deniability, I can rob my audience of knowledge and therefore of reasons for adversarial actions against me. For instance, I can rob them of reasons to blame me for being a racist.

Naturally, emotional reactions such as resentment, anger, etc. are governed by reasons too.¹⁸ Friends of KN-A should thus be sympathetic to the following principle.

KN-E It is epistemically proper to treat p as a reason for emotional attitudes iff one knows p .¹⁹

For instance, I can properly resent you based on your racist attitudes only if I know you have them. Given KN-E, it is clear why deniability is desirable even in cases where the speaker need not worry about adversarial *actions*. In Restaurant, for instance, the waiter might just politely ignore an attempt to bribe, even if she objects to it. She might still feel resentment,

¹⁵Berstler (2019, p. 27) similarly suggests that 'I have plausible deniability [when] my interlocutors won't know whether my claim 'I didn't mean p ' is true or not.' But, first, this definition deserves much deeper discussion than Berstler aims to provide. Second, it must be modified in certain respects; see below. Third, we think it should be extended to deniability full stop, whether plausible or not; see below.

¹⁶This may be for various reasons. Noteworthy among them may be the use of figleaves in prior discourse; see e.g. Saul, 2017, pp. 103–106 on 'synchronic figleaves'.

¹⁷This suggests that knowledge is defeasible, contra e.g. Lasonen-Aarnio, 2010. We can dispense with this assumption once we get to our official definition below.

¹⁸Maguire (2018) opposes this 'dogma', but even he denies only that reasons for affective attitudes are of the same kind as reasons for actions. This seems consistent with what we have to say.

¹⁹Buchak (2014, p. 299), for instance, considers a knowledge norm for blaming.

and one may want to retain deniability because one wants to be able to eliminate reasons for such emotional attitudes.²⁰

Consider Desideratum 2, according to which speaking directly is not a means to retain deniability in the target cases. We can straightforwardly accommodate this desideratum as follows. If e.g. the speeding driver were to directly offer a bribe, her audience would know she offered a bribe even if she denied this. Thus, she would not retain deniability by speaking directly. Turning to the more precise Desideratum 2', she would not even retain deniability in relevant counterfactual situations. For instance, even a judge would know what she meant based on the police officer's testimony.

Note that our account does not imply that speaking directly is never a means to retain deniability. On the contrary. Think of cases where it is easy to mishear one another. Here, even literally expressed contents can be deniable because the hearers can fail to know what was said. Or think of situations in which we are prone to misspeak or where we can convincingly appeal to irony or metaphor, or where we can convincingly 'feign ignorance' (Stanley, 2015, p. 156) of the literal meaning of a given word or sentence. Here, too, even semantically expressed contents can be deniable, not because the hearer does not know what the speaker said but because she does not know whether the speaker meant what she said.²¹

Consider Desideratum 3, according to which speaking indirectly is a means to obtain deniability in the target cases. This desideratum may seem problematic for our account. It may seem that, in *Speeding Driver* for instance, the police officer knows that the speeding driver meant to offer a bribe and that she would retain this knowledge even if the speeding driver denied that she meant to offer a bribe. On our account, it follows that the speeding driver lacks deniability. Given that the speeding driver speaks indirectly, it then follows that speaking indirectly is not a means to obtain deniability, contrary to Desideratum 3. We offer three responses to this type of concern, focusing on *Speeding Driver* for concreteness.

First, it is not obvious that the speeding driver cannot shake the police officer's knowledge with a relevant denial. Thus, it is not obvious that she lacks deniability. Recall that there are two ways to retain deniability: either your audience does not know what you meant to begin with, or your denial defeats otherwise present knowledge. Now there are two ways to defeat otherwise present knowledge. The first works by undermining the evidential dimension of knowledge and endows you with what we call *evidential deniability*. You have evidential deniability if your audience is no longer *in a position to* know that you meant to ϕ once you deny that you meant to ϕ . This could be because your audience's body of evidence no longer suffices for knowledge that you meant to ϕ once we add to this body of evidence that you denied that you meant to ϕ . The second way to defeat otherwise present knowledge works by undermining the psychological dimension of knowledge and endows you with what we dub *psychological deniability*. You have psychological deniability if your audience no longer believes and therefore no longer knows that you meant to ϕ once you deny that you mean to

²⁰Lee and Pinker (2010, p. 795) similarly suggest that speakers sometimes speak indirectly to avoid 'the risk of awkwardness or shame in the same way that a briber avoids the risk of an arrest.' But they focus only on such emotions as a result of a clash between 'the relational model assumed by the speech act' and 'the model that currently holds between the speaker and hearer'. This seems too narrow. When I convey racist communicative intentions, my worry is not that I shift the 'relational model' in some unappreciated way, but rather that I face an audience who condemns racism and thus acquires a reason to condemn me.

²¹See Boogaart, Jansen, and van Leeuwen, 2020, pp. 9–16 for an array of real-life cases of attempted denials of literal contents and Michaelson, 2021 on 'sneaky reference', which might afford deniability relative to literal contents in yet another way. These considerations suggest that deniability does not entail cancellability (*pace* Lee and Pinker, 2010, p. 791 and Mazzarella et al., 2018, p. 16). After all, the literal meaning of a sentence is normally thought to be non-cancellable. See e.g. Blome-Tillmann, 2008 and Zakkou, 2018 for discussion. Stanley (2015) also suggests that certain not-at-issue contents are deniable (156) but not cancellable (139), but see Saul, 2018, p. 374 for critical discussion of his specific examples.

φ. While psychological and evidential deniability often go together, they can come apart if your audience loses a belief despite the fact that it would amount to knowledge.²²

Returning to Speeding Driver, we can acknowledge that the speeding driver may lack evidential deniability because her audience's evidence is too strong. The speeding driver may still have psychological deniability. Strategic speakers often deny their communicative intentions by offering an alternative interpretation of their utterance (Camp, 2018, p. 45). The speeding driver, for instance, could say, 'I didn't mean to offer a bribe! I was just wondering if I could pay my fine right away.' Now it is widely held that the salience of such alternatives can affect our thinking in irrational ways. According to Gerken (2017, p. 116), for instance, speakers have a psychological tendency to treat salient possibilities as epistemically relevant even when they are not. Relatedly, Hawthorne (2004, pp. 162–166) and Williamson (2005, p. 226) suggest that salience can make an error-possibility appear more likely than it is (see also Nagel, 2010 and Dinges, 2018b for further discussion). Each of these mechanisms may make the police officer lose her belief about what the speeding driver meant even though her evidence is strong enough for knowledge. Each of these mechanisms may therefore credit the speeding driver with psychological deniability even when she lacks evidential deniability.²³ Admittedly, though, this does not cover all cases. Some police officers, for instance, might not fall for these psychological effects and so the speeding driver may lack deniability after all.²⁴

Second, recall that, officially, our third desideratum is Desideratum 3' rather than Desideratum 3. Even if we cannot accommodate the latter desideratum because the speaker lacks deniability, we may still be able to accommodate the former. This desideratum is accommodated already if the speeding driver has deniability relative to a relevant part of her audience in a relevant counterfactual or future situation. Speaking indirectly might not be a means for the speeding driver to retain deniability relative to the police officer, but it might still be a means to retain deniability relative to a judge if the police officer decides to arrest the speeding driver for bribery (see also Lee and Pinker, 2010, p. 790 and Camp, 2018, p. 50 cited above).

Third, even Desideratum 3' is a bit too strong. We have assumed so far that speaking indirectly must be an actual means to retain deniability in the target cases. However, to explain why a speaker speaks indirectly, it suffices that she believes, or maybe only hopes, that speaking indirectly is a means to retain deniability. We should weaken Desideratum 3' accordingly. This opens up further ways to accommodate this desideratum even if we grant that the speeding driver has neither actual nor relevant counterfactual, future or partial deniability. For instance, optimism bias (e.g. Sharot, 2012) may lead the speeding driver to be overly optimistic about the extent of her deniability or she might just be confused about the epistemic standards in court. She might therefore believe she has deniability when in fact she does not.

We think that all target cases, i.e., all cases where a desire to retain deniability supposedly explains indirect speech, fall into one or the other category above. In all these cases, speaking indirectly is a means to retain evidential or psychological, actual or counterfactual, present or

²²We follow the widely accepted assumption that knowledge entails belief. For recent discussion, see e.g. Myers-Schulz and Schwitzgebel, 2013; Rose and Schaffer, 2013.

²³According to many versions of epistemic contextualism (e.g. Lewis, 1996; DeRose, 2009; Blome-Tillmann, 2014) and some versions of anti-intellectualism (e.g. Hawthorne, 2004; Dinges, 2018a), salient error-possibilities may destroy 'knowledge' or knowledge because they tend to raise the epistemic standard. This yields another form of non-evidential deniability potentially available to the speeding driver.

²⁴This suggests that cancelability does not entail deniability. For on many accounts of cancellability (e.g. Blome-Tillmann, 2008 and Zakkou, 2018), the respectively insinuated content is still cancellable. Weiner (2006) suggests an account of cancellability where this is questionable, but we think he confuses cancellability with deniability. Lee and Pinker (2010, p. 791) also argue that cancelability does not entail deniability, but their argument relies on a problematic conception of deniability; see footnote 10.

future, full or partial or real or perceived deniability. In this way, we can accommodate the basic insight behind Desideratum 3.

Some readers will harbor residual doubts about our account of Desideratum 3. Aren't there cases where speakers speak indirectly due to deniability even though every relevant hearer obviously knows what they meant even if they deny this? We address such cases of implausible deniability below. First, however, we turn to a generalization of our account.

7 A Generalization

The suggested account of deniability is available only to friends of the above knowledge norms, for these norms helped to explain our key Desideratum 1. While knowledge norms are popular, we do not want to rely on them. We therefore suggest the following revised definition.

Deniability' S has deniability relative to the proposition that she meant to ϕ iff: if S denies that she meant to ϕ , then it is not epistemically proper for S's audience to treat this proposition as a reason for adversarial reactions (actions or emotions).

Given the knowledge norms, this new definition collapses into our previous definition, but if we assume a different norm, these definitions come apart. Our desiderata should still be satisfied because the epistemic state that replaces knowledge on alternative norms is normally quite similar to knowledge (see Benton, n.d. for an overview of the discussion).²⁵

This definition is still not neutral enough. For the epistemic requirements on reasons may differ for different reactions, say, for emotions as opposed to actions or, within actions, for more or less drastic measures. This would make Deniability' problematic, because one now has to ask *which* reactions matter for deniability. For instance, does the speeding driver have deniability already if the police officer cannot treat the proposition that she meant to offer a bribe as a reason for arresting her? Or does she gain deniability only if the police officer cannot treat this proposition as a reason for resenting her, for scolding her, etc.?

One response would be that one has deniability when one's denial undermines reasons for *some* or for *every* adversarial reaction available to the audience. We favor a more flexible approach that further relativizes deniability to a set of adversarial reactions. The logical form of deniability and, correspondingly, our key Desiderata 1' to 3' should be adjusted accordingly.

Deniability'' S has deniability relative to the proposition that she meant to ϕ and relative to a set of adversarial reactions α available to S's audience iff: if S denies that she meant to ϕ , then it is not epistemically proper for S's audience to treat this proposition as a reason for adversarial reactions α .

For instance, the epistemic requirements on reasons for resentment may be low, and so the speeding driver may not have deniability when it comes to this reaction. Meanwhile, the epistemic requirements for reasons for arresting someone may be high, and so, by speaking

²⁵The above definition no longer presupposes that knowledge is defeasible; see footnote 17. Only the epistemic status required for properly treating something as a reason comes out as defeasible. One might worry that, given the knowledge norms, it still follows that knowledge is defeasible. We can avoid this conclusion by granting that speakers have deniability already if the respective reasons lack primary or secondary (or tertiary) epistemic propriety. This commits us only to the conclusion that knowledge or the respective secondary (or tertiary) epistemic conditions are defeasible. Proponents of indefeasible knowledge seem happy to accept this; see e.g. Williamson, forthcoming, n. 19.

indirectly, the speeding driver may retain deniability relative to this reaction. The latter fact may suffice as a reason to speak indirectly if the speaker cares about not being arrested. For simplicity, we leave the relativizations to adversarial reactions implicit in the following.

8 Implausible Deniability

Both Camp (2018) and Berstler (2019, pp. 27–28) suggest that there are cases of implausible deniability. In these cases, speakers have deniability and yet the relevant audience knows their communicative intention even when they deny it; or as Camp (2018, p. 47) puts it, the respective denials are ‘bald-faced lies’. Cases of implausible deniability are immediate counterexamples to our initial definition of deniability in terms of knowledge. They also clash with our revised, reasons-based definition if we assume that it is epistemically proper to treat known propositions as a reason; in brief, that knowledge suffices for actionability. For assuming that knowledge suffices for actionability, cases of implausible deniability are cases where a speaker has deniability and yet, even in the face of denials, it is epistemically proper for the audience to treat the target proposition as a reason for adversarial reactions. In brief, the argument against our account of deniability rests on the following two premises.

- P1** There are cases of implausible deniability.
P2 Knowledge suffices for actionability.

These premises jointly entail that there are counterexamples to Deniability, Deniability’ and Deniability”.

In what follows, we respond to this concern. Our official response denies P1. However, we also present familiar ways to deny P2. This shows that even die-hard proponents of implausible deniability have work to do before they can reject our account.

Why should we accept P1? Camp (2018, p. 47) points out that speakers often ‘get away’ with their denials and are not ‘held accountable’ even though it is mutually known that they are lying. This seems to motivate the idea that speakers can have deniability despite mutual knowledge of the speaker’s communicative intentions and hence that there are cases of implausible deniability. But we think this motivation is unconvincing for the following reason.

For starters, let us define a practical notion of ‘untouchability’ alongside the epistemic notion of deniability.

Untouchability S has *untouchability* relative to the proposition that she meant to ϕ iff: if S denies that she meant to ϕ , then it is practically irrational for S’s audience to engage in pertinent adversarial reactions.

Untouchability does not entail deniability as defined. Suppose a speaker had racist communicative intentions and that it is epistemically proper for the audience to treat this intention as a reason for e.g. speaking up. The speaker lacks deniability on our definition, but she may still have *untouchability*. This is because the audience may have reasons that speak against speaking up, and these reasons can make it overall irrational to do so. For instance, speaking up may be socially inappropriate in certain contexts, it may draw more attention to a speaker than she deserves, it may be costly if you are e.g. dealing with a superior or an oppressor, it may be practically difficult if you are e.g. part of a larger crowd, etc.²⁶

Now, when speakers ‘get away’ with their denials and are not ‘held accountable’, this means primarily that they have *untouchability*. But as argued, it does not follow that they have deniability. Camp’s argument loses its force once we distinguish these two phenomena.

²⁶We draw here on a similar list of reasons against speaking up from Goldberg, 2020, p. 177.

Absent further arguments for assuming cases of implausible deniability, we are free to reject them based on our definition of deniability.

One could object that what we call untouchability actually is deniability. However, defining deniability as untouchability has odd consequences. A speaker may gain untouchability relative to the proposition that she meant to convey something racist only because I bribe everybody to not call her a racist. Similarly, a speaker may lose untouchability relative to the proposition that she meant to convey something racist only because I bribe everybody to groundlessly accuse her of racism. We find it odd to say that speakers gain or lose deniability in this way. Most authors, us included, think of deniability in much more epistemic terms. The above definition Deniability 4 from the literature, for instance, clearly targets a more epistemic notion. To be sure, ‘deniability’ might have both a practical and an epistemic reading. Practical deniability could then be defined as untouchability. However, our paper should be construed as concerning epistemic deniability only.

These considerations undermine P1, according to which there are cases of implausible deniability. As indicated, we also want to challenge P2, according to which knowledge suffices for actionability. This premise bears on a wide-ranging discussion of the norms of action and emotion, which we cannot cover here. We only want to flag some familiar ways to deny that knowledge always suffices for actionability.

Many authors hold that knowledge suffices for actionability. This follows from the previous knowledge norms, and it also follows from norms that appeal to e.g. truth or justified belief. After all, knowledge entails true, justified belief. Gerken (2017, pp. 130–150), however, defends a context-sensitive alternative to such norms where the epistemic requirements on reasons sometimes rise above the standard of knowledge (see also Brown, 2008; Reed, 2010). Simplified and adjusted to the present terminology, he suggests the following principle.

WN-A It is epistemically proper to treat p as a reason for action iff one is warranted in believing p to a degree that is adequate relative to one’s present deliberative context.

Here, ‘warrant’ is used as ‘a broad label for epistemic rationality’ (Gerken, 2017, p. 10), while the following is a non-exhaustive list of factors that determine the ‘deliberative context’: ‘alternative *courses of action*, availability of *further evidence*, considerations of *urgency*, the *stakes* associated with the action, the *social roles* and *conventions* associated with the action’ (133). Depending on these factors, the warrant required for actionability may go beyond knowledge according to Gerken (141–143).

Importantly, knowledge and actionability may come apart specifically in putative cases of implausible deniability. Suppose a strategic speaker insinuates racist messages. The stakes for calling her out might be high due to the high social costs associated with being a racist. Mistaken accusation can have dramatic consequences.²⁷ This might raise the epistemic standards beyond knowledge. At least, the *perceived* epistemic standard may be very high. Similarly, police officers and judges, say, occupy specific social roles with associated conventions. These conventions might place very specific epistemic burdens on them that might go beyond knowledge. Lee and Pinker (2010, p. 790) relatedly appeal to the ‘high standard of proof beyond a reasonable doubt’ in court to explain deniability.

Notably, Gerken’s account nicely accommodates Camp’s (2018, pp. 51–52) suggestion that deniability in the context of sexual harassment is constrained by a ‘reasonable person’ standard in some legal and administrative domains. Why should such legal and administrative standards have any bearing on the scope of a speaker’s deniability? In the present framework, we can explain this as follows. By establishing such standards, we change the conventions

²⁷See e.g. Mendelberg, 2001 on the ‘Norm of Racial Equality’.

associated with adversarial reactions towards sexual harassment. Such conventions are part of the deliberative context. Thus, on Gerken's view, they set the epistemic standard for adversarial reactions and, on our definition, the epistemic standard for deniability.²⁸

In sum, we do not think that cases of implausible deniability exist. But even if they do, it does not follow that our definition fails. One still has to establish that knowledge suffices for actionability in the relevant cases. As indicated, this is not obvious, and some proponents of implausible deniability might be particularly open to rejecting this assumption.

9 Conclusion

Having offered an account of deniability, we want to end by drawing out some conclusions on how to counter strategic speech. Our account helps to distinguish some general strategies to avoid crediting our interlocutors with more deniability than necessary, thereby offering a framework for implementing more concrete proposals. First, one should heed the distinction between deniability and untouchability. Limiting the former is different from limiting the latter. Untouchability often results from power asymmetries, fear of conflict, etc. These factors are usually irrelevant for deniability. Second, focusing on deniability, one should heed the distinction between evidential deniability and psychological deniability. Limiting evidential deniability is a matter of epistemic vigilance. If you pay close attention to contextual cues, your evidential basis may be strong enough to nail down the speaker on a communicated message even if she denies it. Meanwhile, we can limit psychological deniability primarily by believing what our evidence tells us to believe, not letting ourselves be fooled into e.g. taking irrelevant error-possibilities seriously.²⁹ Finally, if we assume that conventions and social roles partially determine the epistemic standard for actionability, we can limit deniability by changing these conventions and social roles; for instance, we can introduce reasonable person standards where this is possible. These strategies remain to be fleshed out, but we hope the presented theoretical framework facilitates this task.

References

- Benton, Matthew A. (n.d.). "Knowledge Norms". In: *The Internet Encyclopedia of Philosophy*. URL: <http://www.iep.utm.edu/kn-norms/>.
- Berstler, Sam (2019). "What's the good of language? On the moral distinction between lying and misleading". In: *Ethics* 130.1, pp. 5–31.
- Blome-Tillmann, Michael (2008). "Conversational implicature and the cancellability test". In: *Analysis* 68.2, pp. 156–160.
- (2014). *Knowledge and Presuppositions*. Oxford: Oxford University Press.
- Boogaart, Ronny, Henrike Jansen, and Maarten van Leeuwen (2020). "'Those are your words, not mine!' Defence strategies for denying speaker commitment". In: *Argumentation* Latest Articles, pp. 1–27.
- Brown, Jessica (2008). "Subject-sensitive invariantism and the knowledge norm for practical reasoning". In: *Noûs* 42.2, pp. 167–189.
- Buchak, Lara (2014). "Belief, credence, and norms". In: *Philosophical Studies* 169.2, pp. 285–311.

²⁸On certain forms of anti-intellectualism (e.g. Wright, 2011 and McKenna, 2020), knowledge itself depends on social roles and associated conventions. Such revisionary accounts of knowledge would allow one to accommodate Camp's suggestion while assuming the previously stated knowledge norms of action and emotion.

²⁹If we assume epistemic contextualism or anti-intellectualism, we can further limit non-evidential deniability by resisting inappropriate attempts to raise the epistemic standard; see footnote 23.

- Camp, Elisabeth (2018). "Insinuation, common ground, and the conversational record". In: *New Work on Speech Acts*. Ed. by Daniel Fogal, Daniel W. Harris, and Matt Moss. Oxford: Oxford University Press, pp. 44–66.
- Davies, Alex (2019). "Testimony, recovery and plausible deniability: A response to Peet". In: *Episteme* 16.1, pp. 18–38.
- DeRose, Keith (2009). *The Case for Contextualism*. Vol. 1. Knowledge, Skepticism, and Context. Oxford: Oxford University Press.
- Dinges, Alexander (2018a). "Anti-intellectualism, egocentrism and bank case intuitions". In: *Philosophical Studies* 175.11, pp. 2841–2857.
- (2018b). "Knowledge and availability". In: *Philosophical Psychology* 31.4, pp. 554–573.
- Dretske, Fred I. (1970). "Epistemic operators". In: *Journal of Philosophy* 67.24, pp. 1007–1023.
- Fricker, Elizabeth (2012). "Stating and insinuating". In: *Aristotelian Society Supplementary Volume* 86.1, pp. 61–94.
- García-Carpintero, Manuel (2018). "Sneaky assertions". In: *Philosophical Perspectives* 32.1, pp. 188–218.
- Gerken, Mikkel (2017). *On Folk Epistemology: How We Think and Talk about Knowledge*. Oxford: Oxford University Press.
- Goldberg, Sanford C. (2020). *Conversational Pressure: Normativity in Speech Exchanges*. Oxford: Oxford University Press.
- Grice, Herbert Paul (1989). *Studies in the Way of Words*. Cambridge and MA: Harvard University Press.
- Hawthorne, John (2004). *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Hawthorne, John and Jason Stanley (2008). "Knowledge and action". In: *Journal of Philosophy* 105.10, pp. 571–590.
- Jamieson, Kathleen Hall (1992). *Dirty Politics: Deception, Distraction, and Democracy*. Oxford: Oxford University Press.
- Khoo, Justin (2017). "Code words in political discourse". In: *Philosophical Topics* 45.2, pp. 33–64.
- Lackey, Jennifer (2007). "Norms of assertion". In: *Noûs* 41.4, pp. 594–626.
- Lasonen-Aarnio, Maria (2010). "Unreasonable knowledge". In: *Philosophical Perspectives* 24.1, pp. 1–21.
- Lee, James J. and Steven Pinker (2010). "Rationales for indirect speech: the theory of the strategic speaker". In: *Psychological Review* 117.3, pp. 785–807.
- Lewis, David (1996). "Elusive Knowledge". In: *Australasian Journal of Philosophy* 74.4, pp. 549–567.
- Maguire, Barry (2018). "There are no reasons for affective attitudes". In: *Mind* 127.507, pp. 779–805.
- Mazzarella, Diana et al. (2018). "Saying, presupposing and implicating: How pragmatics modulates commitment". In: *Journal of Pragmatics* 133.133, pp. 15–27.
- McKenna, Robin (2020). "Feminist epistemology and pragmatic encroachment". In: *Social Epistemology and Relativism*. Ed. by Natalie Ashton et al. London: Routledge.
- Mendelberg, Tali (2001). *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton, NJ: Princeton University Press.
- Michaelson, Eliot (2021). "Speaker's reference, semantic reference, sneaky reference". In: *Mind & Language* Early View, pp. 1–20.
- Myers-Schulz, Blake and Eric Schwitzgebel (2013). "Knowing that P without believing that P". In: *Noûs* 47.2, pp. 371–384.
- Nagel, Jennifer (2010). "Knowledge ascriptions and the psychological consequences of thinking about error". In: *The Philosophical Quarterly* 60.239, pp. 286–306.

- Peet, Andrew (2015). "Testimony, pragmatics, and plausible deniability". In: *Episteme* 12.1, pp. 29–51.
- (2018). "Testimonial knowledge without knowledge of what is said". In: *Pacific Philosophical Quarterly* 99.1, pp. 65–81.
- Pepp, Jessica (2018). "Assertion, lying and untruthfully implicating". In: *The Oxford Handbook of Assertion*. Ed. by Sanford Goldberg. Oxford University Press, pp. 829–850.
- Pinker, Steven, Martin A. Nowak, and James J. Lee (2008). "The logic of indirect speech". In: *Proceedings of the National Academy of Sciences of the United States of America* 105.3, pp. 833–838.
- Reed, Baron (2010). "A defense of stable invariantism". In: *Noûs* 44.2, pp. 224–244.
- Reins, Louisa M. and Alex Wiegmann (2021). "Is lying bound to commitment? Empirically investigating deceptive presuppositions, implicatures, and actions". In: *Cognitive Science: A Multidisciplinary Journal* 45.2, pp. 1–35.
- Rose, David and Jonathan Schaffer (2013). "Knowledge entails dispositional belief". In: *Philosophical Studies* 166.S1, pp. 19–50.
- Saul, Jennifer M. (2017). "Racial figleaves, the shifting boundaries of the permissible, and the rise of Donald Trump". In: *Philosophical Topics* 45.2, pp. 97–116.
- (2018). "Dogwhistles, political manipulation, and philosophy of language". In: *New Work on Speech Acts*. Ed. by Daniel Fogal, Daniel W. Harris, and Matt Moss. Oxford: Oxford University Press, pp. 360–383.
- Sharot, Tali (2012). *The Optimism Bias: Why We're Wired to Look on The Bright Side*. London: Constable & Robinson Ltd.
- Stanley, Jason (2015). *How Propaganda Works*. Princeton, NJ: Princeton Univ. Press.
- van Elswyk, Peter (2020). "Deceiving without answering". In: *Philosophical Studies* 177.5, pp. 1157–1173.
- Viebahn, Emanuel (2020). "Lying with presuppositions". In: *Noûs* 54.3, pp. 731–751.
- Walton, Douglas (1996). "Plausible deniability and evasion of burden of proof". In: *Argumentation* 10.1, pp. 47–58.
- Weiner, Matthew (2005). "Must we know what we say?" In: *The Philosophical Review* 114.2, pp. 227–251.
- (2006). "Are all Conversational Implicatures Cancellable?" In: *Analysis* 66.290, pp. 127–130.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford: Oxford University Press. ISBN: 0-19-825043-6.
- (2005). "Contextualism, subject-sensitive invariantism and knowledge of knowledge". In: *The Philosophical Quarterly* 55.219, pp. 213–235.
- (forthcoming). "Justifications, excuses, and sceptical scenarios". In: *The New Evil Demon*. Ed. by Julien Dutant and Fabian Dorsch. Oxford: Oxford University Press.
- Wright, Sarah (2011). "Knowledge and social roles: A virtue approach". In: *Episteme* 8.1, pp. 99–111.
- Zakkou, Julia (2018). "The cancelability test for conversational implicatures". In: *Philosophy Compass* 13.12, pp. 1–14.