

Non-Epistemic Deniability

Sam Berstler
Massachusetts Institute of Technology

Forthcoming in *MIND*

Abstract. This paper develops an analysis of non-epistemic deniability. On my analysis, a speaker has non-epistemic deniability for *G-ing* when non-acknowledgment social norms make it impermissible for others to retaliate against the speaker for *G-ing*. I identify two kinds of non-acknowledgment norms that generate non-epistemic deniability: two-tracking norms, which function to contain conflict within a group, and open secrecy norms, which function to inhibit the group from acting on shared knowledge. Narrowly, this paper builds on Alexander Dinges and Julia Zakkou's recent landmark analysis of deniability. Dinges and Zakkou argue that non-epistemic deniability does not exist. I disagree. But I also use their account of epistemic deniability in order to motivate my own analysis of non-epistemic deniability. Broadly, my paper provides a case study in how speakers strategically leverage non-acknowledgment norms in order to protect their own interests at the expense of others'.

Key words: deniability; implausible deniability; strategic speech; insinuation; off-the-record

1. Introduction

If you want to communicate something nasty or awkward or legally risky, it's often best to communicate it in an indirect way. Consider this scene from a faculty meeting:

Feuding colleagues 1

While glaring at her long-time rival, Michael, Lucy says:

Lucy: 1. I know that nobody would ever dream of violating the confidentiality of these proceedings.

What she really means is:

2. You, Michael, have violated our confidentiality policies before.

Why didn't Lucy just say¹ what she meant? Here's one familiar reason: Lucy wanted *deniability* for insulting Michael.

[Acknowledgements placeholder]

¹ According to Grice (1975), one says that *p* only if one means that *p*. Lucy only made as if to say (1). For the sake of simplicity, I won't make this distinction.

What is it to have deniability? Philosophers, linguists, and social scientists generally assume that deniability is an epistemic sort of thing (Walton 1996; Pinker 2007; Pinker et al 2008; Lee and Pinker 2010; Fricker 2012; Peet 2015; Davies 2019; Mazzarella 2021; Dinges and Zakkou 2023). Lucy has deniability insofar as the relevant people, in the relevant context, believe that she didn't insult Michael or believe that it's at least plausible that she didn't or even just don't know that she did.

But I am going to argue that while deniability is *sometimes* an epistemic sort of thing, it isn't always. There are non-epistemic² kinds of deniability, and speakers regularly have these kinds of deniability in everyday contexts.

On my analysis, a speaker has non-epistemic deniability for some act G when certain kinds of social norms, which I call *non-acknowledgment norms*, make it impermissible for others to retaliate against her for G -ing. I focus on two types of such norms: two-tracking norms, which function to contain conflict within a group, and open secrecy norms, which function to inhibit groups from acting on shared knowledge. Crucially, whether these norms prohibit others from retaliating against the speaker depends not just upon *what* message the speaker communicates but upon *how* the speaker communicates it.

This paper both criticizes and builds on Alexander Dinges and Julia Zakkou's recent landmark analysis of deniability. On both my and Dinges and Zakkou's (D&Z) proposal, a speaker has deniability insofar as her addressee cannot permissibly respond to her in certain ways. On D&Z's proposal, the norm that prohibits the addressee from acting in these ways is an epistemic norm. But I argue that certain social norms can *also* play this role. When this happens, the speaker has a form of non-epistemic deniability. In doing so, I respond to D&Z's argument that non-epistemic deniability does not exist.³ But I also extend, and therefore demonstrate the fecundity, of D&Z's central insight.

Here's the plan. In the next section, I introduce D&Z's analysis of epistemic deniability. I then generalize their analysis and show how this generalization can guide our inquiry into non-epistemic deniability. In the third section, I develop my own analysis of non-epistemic deniability. In the fourth section, I respond to D&Z's argument against non-epistemic deniability. I also compare my analysis to Elisabeth Camp (2018)'s earlier, competing account of non-epistemic deniability.

2. What epistemic deniability is

In idiomatic speech, we often treat "deniability" as interchangeable with "plausible deniability." This suggests an analysis on which a speaker has deniability for G -ing when she can believably or convincingly or plausibly deny G -ing (Peet 2015: 31). For example:

Epistemic deniability as believability

S has epistemic deniability for G -ing, relative to an agent A , iff: if S denies to A that she G -ed,⁴ A will believe that her denial is plausible.

The problem is that this analysis doesn't capture the full breadth of cases. Sometimes, we judge that a denial is implausible. Nonetheless, we find ourselves thinking: "If only the speaker had said out loud what she just insinuated! *Then* I really could have nailed the bastard!" Consider this case:

² Berstler (2019) and Dinges and Zakkou (2023) sometimes call this implausible deniability. I'm avoiding this term because it can be misleading. As I'll discuss later, it's possible to have deniability that is both epistemic and intuitively implausible (Dinges and Zakkou 2023: 380).

³ Here D&Z argue against Camp (2018) and Berstler (2019).

⁴ D&Z (2023: 385 and *passim*) here have " S denies to A that she intended to G ." For the cases at issue in this paper, my simplification doesn't matter.

Racist realtor

Barbara is white realtor showing a Black couple, Dan and Jasmine, homes in a tony suburb. A neighboring town, Ashwood, is historically Black.

- Barbara:** 3. Perhaps you'd be more comfortable in Ashwood, which is, ahhh, a more *transitional* neighborhood.
- Dan:** 4. What, because we're Black?
- Barbara:** 5. Oh goodness no, that has nothing to do with it.

(adapted from Camp 2018: 48)

Let us suppose that Dan finds Barbara's denial outrageously implausible. And yet, as Camp put it, Barbara has successfully "rhetorically frustrat[ed]" him (Camp 2018: 47). In what way?

To answer this question, D&Z propose that we shift how we're thinking about deniability. In the initial presentation of their analysis, D&Z propose that a speaker has deniability because of what the addressee doesn't know:⁵

Epistemic deniability as falling short of knowledge

S has epistemic deniability for *G-ing*, relative to an agent *A*, if and only if: if *S* denies that she *G-ed*, then *A* does not know whether *S G-ed*.

(adapted from D&Z 2023: 385)

We can find *p* implausible but still fail to know that *p* is false. So on this analysis, not all epistemic deniability is plausible deniability.

Why would a speaker like Barbara desire this minimal form of deniability (D&Z 2023: 378)? D&Z propose that Barbara has *normatively* frustrated Dan. Initially, D&Z endorse the knowledge norm on actionability (Hawthorne and Stanley 2008):

Knowledge norm on actionability

It is epistemically permissible⁶ to treat *p* as a reason for action if and only if one knows that *p*.

(adapted from D&Z 2023: 386)

Thus, since Dan doesn't know what Barbara insinuated, he cannot permissibly accuse her of lying, lodge an anti-discrimination complaint, or even privately resent her (D&Z 2023: 387).

What's essential to D&Z's analysis isn't the knowledge norm on actionability. Rather, what's essential is their appeal to epistemic permissibility facts. In their final analysis, D&Z drop the assumption that knowledge is the norm on actionability and offer a more neutral account:

⁵ Earlier, less developed versions of this idea appear in Pinker (2007: 453, 456-457), Pinker et al (2008: 836-837), Lee and Pinker (2010: 796-797), and Berstler (2019: 27).

⁶ Following Stanley and Hawthorne (2008), D&Z use the word "proper" here (D&Z 2023: 386). Stanley and Hawthorne (2008: 578) gloss propriety as permissibility. I revert to the more familiar "permissible" in order to emphasize the continuity between this norm and the norms I later discuss.

Epistemic deniability as impermissibility

S has epistemic deniability for *G-ing*, relative to an agent *A* and some action *H*, if and only if: if *S* denies that she *G-ed*, then it is not epistemically permissible for *A* to treat the proposition that *S G-ed* as a reason for *H-ing*.

(adapted from D&Z 2023: 391)

Thus, rather than thinking about deniability in terms of what the *speaker can* do, D&Z ask us to think about deniability in terms of what the *audience cannot* do. On their story, the speaker cares about the addressee's epistemic state because she cares about the normative facts that obtain in virtue of it. The actionability norm is a kind of extra-linguistic tool, through which the speaker can indirectly exert control over her interlocutor's set of permissible options.

D&Z's account of deniability more or less ends here. But there's a natural way to extend it. When we act, we are robustly sensitive to all *sorts* of permissibility facts—not just the epistemic ones. What we do is robustly sensitive to what we can morally, socially, and legally do (and so on).⁷ Do speakers ever take advantage of *these* sorts of norms, in a way that's relevantly similar to what Lucy and Barbara do when they take advantage of the actionability norm?

Given our aims, we want to find norms that we can slot into the below schema and thereby produce a theoretically viable analysis of deniability:

Deniability recipe

S has **x-deniability** for *G-ing*, relative to an agent *A* and a context *c*,⁸ if and only if: in *c* it is impermissible [**according to norm x**] for *A* to [**some acts H**].

What counts as a theoretically viable analysis of deniability? Following D&Z, I will assume that it must conform to the following two constraints.

First, whatever *H* is, there must be a core set of cases in which *S* doesn't want *A* to *H* (D&Z 2023: 378). This is because we want to explain why speakers want deniability in the first place. Second, there must be a core set of cases in which speaking indirectly, rather than directly, is a good way to achieve x-deniability (D&Z 2023: 373). This is because deniability is, in this context, a "semi-technical term," which we want to put to work in our theory of conversational behavior (D&Z 2023: 373). And one of the things we want to explain is why strategic speakers sometimes choose to speak indirectly, rather than directly.⁹

To see how the deniability recipe works, let's consider an example from the legal realm. Thanks to *Bronston v. United States* (1973), it is more or less legal in the US court system to knowingly and intentionally make false implicatures under oath:

⁷ We don't always act in ways that are morally, socially, or legally permissible. But we don't always act in epistemically permissible ways either.

⁸ I've relativized D&Z's definition to a context and deleted the clause, "if *S* denies that she *G-ed*" on the righthand side. This is for two reasons. First, in the cases that I'll go onto consider, the permissibility facts do not depend upon whether *S* denies that she *G-ed*. So including the conditional is misleading. Second, my adjustment eliminates the problem of denials that actually *undermine* a speaker's deniability. As D&Z point out, a bad liar can provide her addressee with evidence that she *did* in fact *G* and thereby lose deniability (D&Z 2023: 386 fn18). In some of the cases I consider, inartful denials can also eliminate a speaker's non-epistemic deniability. Note that a speaker will have deniability relative to many contexts that both she and the theorist will find uninteresting. See D&Z 2023: 376-377 for further motivation of the context parameter.

⁹ Like D&Z, I don't assume that when a speaker speaks indirectly, she always gains some form of deniability. Nor do I assume that if a speaker speaks directly, she never gains deniability (D&Z 2023: 375, 381).

Perjury principle

It is legally permissible to bring a perjury case against someone only if she knowingly and intentionally asserts something false under oath.

This norm, when slotted into the deniability recipe, produces a theoretically viable notion of deniability:

Legal deniability¹⁰

S has legal deniability for *G-ing*, relative to an agent *A* and a context *c*, if and only if: in *c*, it is legally impermissible, according to perjury law, for *A* to bring perjury charges against her for *G-ing*.

To see why, let's check our two constraints. First, except in rare cases, no one wants the prosecutor's office to bring perjury charges against her. So, courtroom witnesses and defendants who want to mislead the judge or jury will generally have robust reason to want to secure legal deniability for themselves. Legal deniability is desirable.

Second, the perjury principle ensures that speaking indirectly, rather than directly, is almost always a good way to secure legal deniability. For example, consider this exchange, adapted from the one at issue in *Bronston v. United States*:

Clever defendant

Prosecutor: 6. Have you ever had a Swiss bank account?

Bronston: 7. My company had one there for about six months.

What Bronston really meant was:

8. I have never had a Swiss bank account.

It turned out that Bronston did have a Swiss bank account, which he used to commit financial improprieties. So in saying (7), Bronston intentionally and knowingly communicated something false under oath. But because he didn't say (8), he had legal deniability for communicating this.

This form of legal deniability only comes into play in special institutional contexts. If our goal is to analyze the micro-strategies of informal conversation, we will rarely appeal to it. But I will now argue that there are other forms of non-epistemic deniability that do regularly come into play in everyday contexts.

3. Why non-acknowledgment norms generate non-epistemic deniability

In everyday life, we often share knowledge that we nonetheless do not acknowledge, discuss, or rely upon for joint action.¹¹ Sometimes, this is because we have no need to do so. But other times, we are keeping this shared knowledge at a remove because we are conforming to norms of politeness, discretion, and other standards of social propriety. I don't acknowledge the obvious holes in a colleague's sweater, even though he can see that I can see them, and I can see that he can see that I can see them, and so on... In such a situation, I am conforming to what I call a *non-acknowledgment norm*.

¹⁰ D&Z briefly endorse a version of legal deniability, albeit one formulated in terms of admissibility rules and not perjury law (D&Z 2023: 390fn25).

¹¹ In many of these cases, our shared knowledge will rise to the level of joint knowledge: we will know *p* and know that we know *p* and know that we know that we know *p* and so on (Lewis 1969).

Non-acknowledgment norms impose constraints on when and how groups deal with knowledge that they share.¹² Such norms tend to emerge through tacit agreements and implicit coordination. We enforce them through subtle patterns of esteem and reward, on the one hand, and disapproval and retaliation, on the other. But like many social norms, we may not be able to articulate their precise content. Sometimes all we can say is: “That sort of thing *just can’t be said* around here.”

Not every non-acknowledgment norm generates the possibility of non-epistemic deniability. Suppose that every Friday morning, a group of faculty meditate together. According to an informal norm, no one speaks during it. One Friday morning, Jamie goes to the meditation space and vents all of his workplace frustrations. Jamie is taking advantage of a non-acknowledgment norm: his colleagues, if they want to conform to the meditation norm, will remain silent. But notice that *how* Jamie communicates his complaints—whether he says them directly or indirectly—doesn’t affect whether his colleagues can permissibly respond to him. Thus, Jamie doesn’t have deniability in the technical sense in which we are interested.

One kind of non-acknowledgment norm that does generate the right explanatory pattern is what I call *two-tracking norms*. When a group institutes a two-tracking norm, it tacitly agrees to divide its communication, in certain contexts, between an “official,” explicit, open channel and an “unofficial,” inexplicit, or hidden channel. The rule is usually something like this. You can be as obnoxious, resentful, aggressive, or uncooperative as you like, *so long as* you confine your obnoxious, resentful, aggressive, or uncooperative communications to the unofficial track. In the official track, you should present yourself as courteous, pro-social, and otherwise morally upstanding. As a rough heuristic, our direct speech acts default to the official track. Our indirect speech acts and forms of communication that don’t rise to the level of meaning, like hints, default to the unofficial track.

The “double-speak” of white-collar professionals exemplifies two-tracking. Consider another exchange between our two feuding colleagues:

Feuding colleagues 2

Lucy: 8. Of course, *you* wouldn’t like this job candidate.
9. *Implicature: You’re too myopic to understand the appeal of this candidate.*

Michael: 10. I just don’t share your deep appreciation for contemporary metaphysics.
11. *Implicature: Well, you only want this candidate because you happen to work in the same area as her!*

Two-tracking norms generally function to help groups regulate conflict. Most obviously, they limit how hostile an exchange can become. Lucy and Michael can permissibly insult each other, but they can’t shout or curse. Second, two-tracking increases the cognitive burden required to communicate an uncooperative message and so disincentivizes doing so. Even for the most communicatively adept, indirect sneering requires more creativity, timing, and skill than just shouting, “You’re a jerk!” Finally, two-tracking makes it more difficult for outsiders and newcomers to learn the details of intra-group conflict. An eavesdropping student or dean may detect the biting edge to Lucy and Michael’s exchange but not understand more than this.

According to one school of thought, two-tracking may be a remarkably widespread strategy. The mid-century sociologist Erving Goffman (1959, 1967) and the sociolinguists that he inspired (Brown and Levinson [1978] 1987) proposed that we each have a ritual social self or *face* and that interlocutors generally work together to preserve their own and each other’s. Our face is fragile. I can damage yours through such banal acts as declining an invitation to your party, asking you to turn down your music, or too enthusiastically praising your

¹² I will remain neutral on whether we ever have moral reason to conform to these norms. (See Nagel (1998) for an argument that we sometimes do.) My argument works so long as we generally have prudential reason to conform.

outfit.¹³ If that's right, then everyday communication, even among cooperative speakers, becomes overloaded with social risk. (Unfortunately, to explain why face is fragile in these ways requires more discussion than I can provide here.)

In order to decrease the riskiness of basic interaction, Goffman and Brown and Levinson imagine a kind of society-wide two-tracking. What is in the unofficial track doesn't "count" for the purposes of face: what we "unofficial[ly]" communicate "need not be faced up to" (Goffman 1967: 30). If I tell you, "I don't want to hang out with you because I'm just not interested in being your friend," I catastrophically attack your face. But if I demur, "I'm just really busy right now, but I'll let you know if I free up!" I reject you, but I don't do the same social damage, even if it's mutually obvious from my tone and the context that I am communicating that I don't want to be your friend. On this view, face is "largely a matter of surface appearances" (Brown and Levinson 1978: 212; see also 69, 73, 211-213).

How does two-tracking generate non-epistemic deniability? First, a two-tracking practice entails an *anti-contamination norm*:

Anti-contamination norm

If a speaker *S* performs some communicative act *H* within the unofficial track, then:

- (i) if the speaker would have violated the official track norms, had she *H-ed* within it,
- (ii) it is not socially permissible to presuppose,¹⁴ for the purposes of communicating within the official track, that the speaker *H-ed*.

Anti-contamination norms ensure that the bad behavior within the unofficial track remains confined to the unofficial track. Thus, the anti-contamination norm prohibits Michael from saying (12) or (13):

- Michael:**
- 12. I'm hurt that you'd suggest that I'm narrow-minded
 - 13. Could you explain why you think I'm so myopic? I just want to understand where you're coming from.

If Lucy had openly insulted Michael, she would have officially insulted him. Within the official track, Michael would need to respond to Lucy with professional decorum. But since the anti-contamination norm wouldn't apply, he could at least say (12) or (13). Thus, Lucy and Michael have two reasons, respectively, to insult each other in the unofficial track. In doing so, Lucy and Michael conform to the two-tracking norm. And they further restrict how they can each permissibly respond to each other.

Thus, in the above exchange, Lucy and Michael both have *two-track deniability*:

Two-track deniability

S has two-track deniability for *G-ing*, relative to an agent *A* and a context *c*, if and only if: in *c*, it is not socially permissible, according to a two-track anti-contamination norm, for *A* to presuppose, for the purposes of communicating within the official track, that *S G-ed*.

A second kind of non-acknowledgment norm that generates non-epistemic deniability is what I call *open secrecy norms*. When a group institutes an open secrecy norm, it cordons off some of its shared knowledge so that members cannot rely on it for the purposes of conversation, collective deliberation, or other joint activity. When open secrecy norms succeed, group members may seem to suffer from willful ignorance. But their problem, insofar as they have one, is not that they are refusing to form a belief. It is that they are refusing to acknowledge what they know to be true.

¹³ What counts as a face-threatening act is culturally mediated (Brown and Levinson [1978] 1987: 13-15).

¹⁴ I am using Stalnaker's notion of presupposition (1999, 2002, 2014). For our purposes, it's safe to gloss presupposing *p* as taking *p* for granted or acting as if *p* is shared information.

Here are two examples. Suppose that in the Jones family, it's an open secret that Uncle Henry is an alcoholic. When Uncle Henry shows up drunk for Thanksgiving, the Jones family ignores his slurring speech and plays along with his claim that he lost his job due to lay-offs. Or suppose that in a biomedical company, it's an open secret that the lab is illegally dumping chemicals into the river. Everyone jointly knows about it. But company executives won't even concede that there are false rumors about the dumping. Employees who insist on discussing the policy with their managers are quietly fired for "poor performance."

Open secrecy norms impose heavy-duty constraints on group members. They require members not to acknowledge the content of the open secret *or* the existence of the open secrecy norm *or* even that the group jointly believes the open secret (Zerubavel 2006, Berstler forthcoming):

Open secrecy

If p is an open secret in a group g , then

- (i) the majority of g jointly believes¹⁵ that p ,
- (ii) it's not socially permissible in g for group members to presuppose, for the sake of interacting with g members, that p , and
- (iii) (i) and (ii) are also open secrets in g .

Open secrecy norms are self-camouflaging in a way that run-of-the-mill non-acknowledgment norms are not. Consider this ordinary non-acknowledgment norm. In their hiring deliberations, faculty agree not to take into account where an applicant received her PhD. If a colleague waxes lyrical about a candidate's PhD from Harvard, the chair can remind her, "We agreed to ignore this sort of thing." In contrast, if Clara insists at Thanksgiving, "Of course Henry is an alcoholic!" her family members cannot say, "We don't talk about that around here!" without violating the very norm that they are trying to enforce. Because of this, open secrecy norms always offer some minimal form of protection *whenever* someone violates them.¹⁶

Therefore, if the communicator is clever enough, she can leverage the open secrecy norm in order to both violate the open secrecy norm and prevent others from openly criticizing her for doing so. Consider this story:

Affair

In a workplace, it's an open secret that two employees, Mark and Coraline, are having an extra-marital affair with each other. At a meeting, Mark asks how the office can build better rapport among its employees. Alejandro mutters:

Alejandro: 14. I think this office needs fewer extracurricular activities, not more.

What he really means is:

15. It would improve our rapport if Coraline and Mark ended their extra-marital affair.

A few people snicker. Mark doesn't reply.

Why doesn't Mark say any of the following?

¹⁵ This clause is meant to capture the sense in which an open secret is not just widely believed information but shared information within the group.

¹⁶ It's very minimal. For example, Clara's family can permissibly insist that she's "crazy."

- Mark:**
16. Your accusation is ridiculous.
 17. That kind of joke is unprofessional.
 18. I'm going to pretend that I didn't hear that.

To see why, we need to think about how Mark does and could, in principle, recognize Alejandro's implicature. Let p pick out the proposition that Coraline and Mark are having an affair. In order to conversationally implicate (15), Alejandro has presupposed p . Because he presupposes that other people will recognize that he is presupposing p and so recognize what he means, he is presupposing that everyone jointly knows p . In fact, people do jointly know p . This is why they have no trouble identifying Alejandro's implicature.

The problem is that if Mark says (16-18), Mark is obviously presupposing that Alejandro meant (15). In saying (15-17), Mark is not acknowledging that p is true. So he's not falling afoul of clause (ii). But Mark *is* presupposing that the majority of g jointly believes p .¹⁷ After all, if Mark weren't presupposing this, Mark could not, in context, identify what Alejandro really meant in saying (14). But when p is an open secret, it's *also* an open secret that the majority of the group jointly believes p . In other words, it's not permissible to presuppose that the open secret is a rumor, even a malicious or false one.

Thus, the very same open secrecy norm that normally protects Mark now requires him to pretend that Alejandro's assertion in (14) is innocent. Mark must act as clueless as an office newcomer, who knows none of the sordid backroom gossip. If Alejandro had just said (15), Mark wouldn't be in this position, because he wouldn't need to presuppose p in order to interpret Alejandro. Thus, Alejandro, despite violating the open secrecy norm, also has non-epistemic deniability in virtue of it:

Open secrecy deniability

S has open secrecy deniability for G -ing, relative to an agent A and a context c , if and only if: in c , it is socially impermissible, according to an open secrecy norm in a group g , for A to presuppose, for the sake of interacting with g members, that S G -ed.

In general, whether speakers can gain non-epistemic deniability depends upon the particular architecture of the social norms in their environment. So when we analyze controversial real world cases or re-evaluate some of the cases from the secondary literature, we will need to hypothesize about the structure of the local social norms. To give you a flavor for how we might do this, consider two such cases.

First, consider **Racist Realtor** and cases like it. In these cases, a speaker (usually, implicitly American) makes a bigoted remark indirectly, but not directly, and thereby gains some kind of protection against censure. Could polite, liberal, multi-racial American society two-track racist and other bigoted communication? (If so, this would obviously be an unjust practice.) For Barbara to gain non-epistemic deniability, we need only assume that speakers cannot permissibly contaminate the official track with bigotry expressed within the unofficial ones.¹⁸ This hypothesis would, in part, explain why those who "call out" obviously problematic insinuations are nonetheless sometimes condemned as crass, aggressive, or uncooperative.

¹⁷ Mark could try to rebuke Alejandro without presupposing that the group jointly believes p . He could say something like, "No one understands what you're insinuating, Alejandro." This is still an odd sort of pretense, however. If Mark actually believed that only he and Alejandro know what Mark meant, Mark probably wouldn't rebuke Alejandro in public. It needlessly calls attention to his unprofessional conduct, and other faculty members might think that Mark is imagining things. To my ear, Mark's rebuke sounds only marginally permissible. That is exactly what these considerations predict.

¹⁸ Mendelburg (2001), Stanley (2015), and Khoo (2017) propose that in these same contexts, there is a norm against racist speech *simpliciter*. Speakers like Barbara speak indirectly in order to gain epistemic deniability for violating this norm. My hypothesis is meant to rival theirs.

Or consider a case that D&Z and others have extensively discussed (Lee and Pinker 2010: 794; D&Z 2023: 374). A patron, Saira, asks a waiter, Van:

- Saira:** 19. Is there a way to shorten my wait for a table?
20. *Implicature:* Will you take a bribe?

If it's an open secret in the local community that waiters take bribes, then we can tell a story on which Saira has non-epistemic deniability. It works much like the one in **Affair**. Saira's insinuation is subtle enough that if Van didn't know the open secret, Van wouldn't recognize what Saira is asking. Van can't openly accept Saira's bribe; he must find a subtler way to indicate his acceptance. But he also can't openly decline it either. To say, "How dare you suggest that this establishment is corrupt!" or even "Sorry, we don't take bribes here" is to violate the open secrecy norm. In such a case, Saira would have open secrecy deniability for bribing Van.

4. Comparison

Why are D&Z skeptical of the existence of non-epistemic deniability? To answer this question, we'll need to zoom out and consider the structure of the entire debate. In the first papers that discuss non-epistemic deniability, Camp (2018) and Berstler (2019: 27-28) appeal to our intuitions about specific cases. Their arguments go something like this. Here is a case in which a speaker denies *G-ing* and seems to have deniability for doing so. Nonetheless, her interlocutor knows that the speaker *G-ed*. So the speaker must have some non-epistemic form of deniability.

The problem is that intuitions about subtle cases can be murky. D&Z's strategy is to simply explain them away (D&Z 2023: 393-394). They propose that in some of the cases, like **Racist Realtor**, the speaker has implausible but epistemic deniability. In still other cases, the speaker has what they call untouchability:

Untouchability

S has untouchability for *G-ing*, relative to an agent *A* and some action *H*, if and only if: if *S* denies that she *G-ed*, then it is practically irrational for *A* to *H*.

(adapted from Dinges and Zakkou 2023: 393)

Consider **Feuding Colleague 1**. Suppose that everyone in the room jointly knows that Lucy has just insinuated that Michael has violated the confidentiality policy. But Lucy also controls a grant fund to which Michael desperately needs access. Michael knows that if he responds to Lucy's insinuation, Lucy will reject his application for funding. Let us suppose that it's therefore irrational for Michael to call out Lucy. Lucy has untouchability, with respect to Michael, for her accusation.

Once we are armed with D&Z's analysis of epistemic deniability and their analysis of untouchability, do we have reason to reject the analysis of non-epistemic deniability that I've been sketching? I don't think so. This is because my argument hasn't proceeded in the same way that Camp's and Berstler's original arguments did. I don't argue, "Here is some data, and the only way to explain the data is to posit the existence of non-epistemic deniability." Instead, I started with independently motivated hypotheses about the structure of certain social norms. If these social norms have the content that I have hypothesized, it becomes rational for strategic speakers to leverage the norms in the way that I have described. And when speakers leverage the norms in those ways, I have argued, they have something deserving of the name deniability.

Independently of its dialectical role, D&Z's analysis of untouchability raises a second question: is untouchability yet another form of non-epistemic deniability? D&Z briefly float the hypothesis that it is. But

they observe that this leads to “odd” results (D&Z 2023: 394). I agree with their ultimate verdict: untouchability is not a form of non-epistemic deniability. But I think that they misdiagnosis the problem.

In defense of their verdict, D&Z ask us to consider a case like the following. A politician, Lorna, drunkenly reveals to her constituent, Ben, “I’m only supporting this bill because I’ve taken money from Big Oil.” Lorna’s staffers tell Ben, “If you keep that remark to yourself, we’ll give you \$1,000,000.” Let us suppose that it becomes rational for Ben to decline to speak to the press. So Lorna has a form of untouchability. Then the press tells Ben, “If you tell us whatever Lorna told you, we’ll give you \$5,000,000.” Let us suppose that it now becomes rational for Ben to speak to the press. D&Z comment that it’s “odd to say that speakers [could] gain or lose deniability in this way” (D&Z 2023: 394)

I don’t think this sort of case is as probative as D&Z do. The problem is not that Lorna can and lose gain untouchability without knowing that she has. Nor is the problem that Lorna can *in some case or other* gain and lose untouchability for communicating *p* for reasons that have nothing to do with how she communicated *p*. I could tell a parallel story about epistemic deniability. Imagine that when Ben goes home after the rally, he slips and suffers a concussion. He loses his memory of the rally, and so Lorna gains epistemic deniability. (We could even add that the wicked staffers intentionally push Ben, hoping that he’ll lose his memory.) But then after treatment, he regains his memory, and Lorna loses epistemic deniability again.

Rather, the problem is that there is no *systematic* explanatory relationship, across a robust class of cases, between speaking indirectly and gaining untouchability. Insofar as there is a systematic explanatory relationship between whether I speak indirectly and whether it’s practically rational for others to retaliate against me, it is only because it is often practically irrational to violate the actionability, anti-contamination, and open secrecy norms. Those are the norms that are doing the heavy explanatory lifting. Thus, a norm like “don’t act in practically irrational ways” isn’t the kind of norm that we can slot into the deniability recipe.

Finally, I should mention one last point of agreement between D&Z and myself. In the first and only other full-length treatment of non-epistemic deniability, Camp (2018) argues that the phenomenon motivates a revisionary analysis of speech acts and discourse. In order to explain non-epistemic deniability, Camp meshes together psychological and commitment-based theories of communication (Camp 2018: 52-62).¹⁹ The idea is that a speaker gains non-epistemic deniability when she performs a communicative act, aimed at changing her interlocutor’s psychological beliefs, but fail to perform the normative act that would commit her to what she has done.

When D&Z undercut Camp’s motivation for positing the existence of non-epistemic deniability, they also undercut her motivation for positing this revisionary theory. Although I believe that non-epistemic deniability does exist, my analysis, like D&Z’s, undercuts Camp’s argument. The mere existence of non-epistemic deniability doesn’t *force* us into adopting any particular foundational view about speech acts and discourse. My account can explain non-epistemic deniability while remaining neutral on such big picture questions.

More generally, Camp and I locate the source of non-epistemic deniability in different places. On Camp’s analysis, what generates the possibility of non-epistemic deniability is something about the nature of speech acts and discourse themselves. On my analysis, what generates the possibility of non-epistemic deniability are extra-linguistic social norms, which happen to be sensitive to how speakers communicate their message.

¹⁹ Representative psychological theories include Bach and Harnish (1979), Grice (1989) and Stalnaker (1999, 2002, 2014). Representative commitment theories include in Peirce (1934) and Brandom (1983, 1994). For further discussion of how these schools of thought have developed, see Fogal, Harris, and Moss (2018).

5. Conclusion

What is it to have deniability? According to Dinges and Zakkou, a speaker has deniability insofar as an epistemic norm shields her from certain forms of retaliation. To this, I have added an addendum. Sometimes, what shields the speaker is not an epistemic norm but a social norm. The trick was to clarify exactly what sorts of social norms can play this role and then to theorize about whether any such norms exist. I have argued that at least two such norms do: two-tracking norms and open secrecy norms. Whether there are any more such norms, I leave as an open question.

Works Cited

- Bach, Kent and Robert M. Harnish. *Linguistic Communication and Speech acts*. Cambridge, MA: MIT Press, 1979.
- Berstler, Sam. “The Structure of Open Secrets.” *The Philosophical Review*. Forthcoming.
- , “What’s the Good of Language? On the Moral Distinction between Lying and Misleading.” *Ethics* 130 (2019): 5-31.
- Brandom, Robert. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press, 1994.
- , “Asserting.” *Nous* 17 (1983): 637-650.
- Brown, Penelope and Stephen C. Levinson. *Politeness: Some Universals in Language Usage*. Cambridge, UK: Cambridge University Press, 1978.
- Camp, Elisabeth. “Insinuation, Common Ground, and the Conversational Record.” In *New Work on Speech Acts*. Edited by Daniel Fogal, Daniel W. Harris, and Matt Moss. 44-46. New York: Oxford University Press, 2018.
- Davis, Alex. “Testimony, Recovery and Plausible Deniability: a Response to Peet.” *Episteme* 16 (2019): 18-38.
- Dinges, Alexander and Julia Zakkou. “On Deniability.” *Mind* 132 (2023): 372-401.
- Fogal, Daniel, Daniel W. Harris, and Matt Moss. “Speech Acts: the Contemporary Theoretical Landscape.” In *New Work on Speech Acts*. Edited by Daniel Fogal, Daniel W. Harris, and Matt Moss. 1-39. New York: Oxford University Press, 2018.
- Fricke, Elizabeth. “Stating and Insinuating.” *Aristotelian Society Supplementary Volume* 86 (2012): 61-94.
- Goffman, Erving. *Interaction Ritual: Essays on Face-to-Face Behavior*. New York: Random House, 1967.
- , *The Presentation of the Self in Everyday Life*. New York: Doubleday, 1959.
- Grice, H.P. *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press, 1989.
- , “Logic and Conversation.” In *Syntax and Semantics 3: Pragmatics*. Edited by Peter Cole and Jerry L. Morgan. 41-58. New York: Academic Press, 1975.

- Hawthorne, John and Jason Stanley. "Knowledge and Action." *Journal of Philosophy* 105 (2008): 571-590.
- Khoo, Justin. "Code Words in Political Discourse." *Philosophical Topics* 45 (2017): 33-64.
- Lee, James J. and Steven Pinker. "Rationales for Indirect Speech: the Theory of the Strategic Speaker." *Psychological Review* 117 (2010) 785-807.
- Lewis, David. *Convention: A Philosophical Study*. Cambridge, MA: Blackwell, 1969.
- Mazzarella, Diana. "'I Didn't Mean to Suggest Anything Like That!': Deniability and Context Reconstruction." *Mind and Language* 38 (2021): 218-236.
- Mendelberg, Tali. *The Race Card: Campaign Strategy, Implicit Message, and the Norm of Equality*. Princeton, NJ: Princeton University Press, 2001.
- Nagel, Thomas. "Concealment and Exposure." *Philosophy and Public Affairs* 27 (1998): 3-30.
- Peet, Andrew. "Testimony, Pragmatics, and Plausible Deniability." *Episteme* 12 (2015): 29-51.
- Peirce, Charles S. *Belief and Judgment*. Cambridge, MA: Harvard University Press, 1934.
- Pinker, Steven. "The Evolutionary Social Psychology of Off-Record Indirect Speech Acts." *Intercultural Pragmatics* 4 (2007): 437-461.
- , Martin A. Nowak, and James J. Lee. "The Logic of Indirect Speech." *Proceedings of the National Academy of Sciences* 105 (2008): 833-838.
- Stalnaker, Robert. *Context*. New York: Oxford University Press, 2014.
- , "Common Ground." *Linguistics and Philosophy* 25 (2002): 701-721.
- , *Context and Content: Essays on Intentionality in Speech and Thought*. New York: Oxford University Press, 1999.
- Stanley, Jason. *How Propaganda Works*. Princeton, NJ: Princeton University Press, 2015.
- Walton, Douglas. "Plausible Deniability and Evasion of Burden of Proof." *Argumentation* 10 (1996): 47-58.
- Zerubavel, Eviatar. *The Elephant in the Room: Silence and Denial in Everyday Life*. New York: Oxford University Press, 2006.