

The Puzzle of Plausible Deniability

Andrew Peet

April 5, 2024

Abstract

How is it that a speaker S can at once make it obvious to an audience A that she intends to communicate some proposition p , and yet at the same time retain plausible deniability with respect to this intention? The answer is that S can bring it about that A has a high justified credence that ' S intended p ' without putting A in a position to know that ' S intended p '. In order to achieve this S has to exploit a sense in which communication can be lottery-like. After defending this view of deniability I argue that it compares favorably to a rival account recently developed by [Dinges and Zakkou \(2023\)](#).

The authors have no relevant financial or non-financial interests to disclose.

1 The Puzzle

Bassa is caught going 50mph in a 30 zone. He is pulled over and an officer approaches his window. Bassa wants to keep his driving record clean, so he decides to offer a bribe. The dialogue proceeds as follows:

Bassa: Is there a problem officer?

Officer: Do you realize you were going 50 in a 30 zone?

Bassa: I didn't realize, I'm very sorry. Look, I have plenty of cash on me. Is there any way that we could settle this right now?

Officer: You know it is illegal to bribe a police officer?

Bassa: Of course! I wasn't trying to bribe you! I was just wondering if I could pay the fine here and now! C.f. [Lee & Pinker \(2010\)](#), p 790

Bassa intends to offer a bribe. In order to be successful he must make his intention clear. However, he does this in a roundabout way. This allows him to maintain plausible deniability: when challenged he is able to insist that he wasn't offering a bribe. There is an intuitive sense in which he avoids *committing* to offering a bribe. Our officer is not in an epistemic position to hold Bassa

responsible in the same way she would if he had simply said 'why don't I just pay you to let me off the hook?'. This situation is typical of plausible deniability. When an agent ϕ s, but retains plausible deniability, they prevent those who would have reason to hold them responsible from being in an epistemic position to do so.

In this sense, plausible deniability stands in contrast to what we might call 'implausible deniability'. I use the term 'implausible deniability' to denote the form of deniability that exists in any case in which there is no epistemic barrier to the audience's holding the speaker responsible, but there is a non-epistemic barrier to their doing so. That is, implausible deniability exists in any case in which the speaker lacks plausible deniability, but is nonetheless able to avoid being held responsible for their act via the presentation of a denial. I suspect that there are many distinct mechanisms of implausible deniability. But the most obvious examples of implausible deniability concern power relations. Imagine a powerful and well-connected male business executive talking to their young female intern. They stand uncomfortably close to the intern, make a statement with a clear lewd implication, and accompany their statement with a wink. The intern understands the remark and knows exactly what the executive means. There is nothing the executive could say to defeat this knowledge. *Epistemically*, the intern is in a position to hold the executive responsible. A senior female executive with just as much power as the male would (let us imagine) have no problem holding the male executive accountable in the same circumstances. However, the intern is cannot hold the executive responsible. The executive could simply deny that they meant anything lewd by their remarks - perhaps offering an unrealistic or far-fetched reconstruction of their intended meaning. In these circumstances the intern is powerless. By pushing the matter further she would face social sanction. The executive is thus able to avoid being held responsible. Our focus in this paper is exclusively on plausible deniability - not implausible deniability.^{1,2}

Plausible deniability is puzzling. Consider the following:

1. Bassa makes it clear to the police officer that he is offering a bribe.
2. The officer is not in an epistemic position to hold him responsible for offering a bribe.
3. If it is clear to a subject A that a subject B has ϕ ed then, *ceteris paribus*, A can hold B responsible for ϕ ing.

¹It is controversial whether such situations really constitute cases of deniability. Dinges and Zakkou (2023) deny the possibility of true implausible deniability, holding that all apparent cases of implausible deniability are really examples of what they call 'untouchability'. I suspect that this issue is merely terminological. Moreover, for our purposes, it is of little interest whether cases of 'implausible deniability' really exist since our focus is only on plausible deniability. See Camp (2018) for discussion of implausible deniability.

²For recent discussions of deniability see Lee & Pinker (2010), Fricker (2012), Peet (2015), Stanley (2015), Khoo (2017), Camp (2018), Saul (2018), Mazarella (Forthcoming), and Dinges and Zakkou (2023).

(1)-(3) are jointly inconsistent. Yet, taken alone, each of (1)-(3) seems true. I'll say a few words on each:

(1): It is important to the non-accidental fulfillment of Bassa's intention that he makes it clear to the officer that he's offering a bribe. If he does not make it clear that he's offering a bribe then, even if the officer is willing to accept a bribe, his attempt to keep his record clean will likely fail. Of course, there is always the possibility that, by dropping an obscure hint, Bassa will make the officer form the belief that he's offering a bribe. And this would suit Bassa's purposes just as well. But if he drops an obscure hint then he's leaving a lot to chance. He is best served by framing his utterance in such a way that its clear what he's doing. And, fortunately for Bassa, he *does* make it clear that he's offering a bribe: Bassa has provided the officer with compelling evidence that he's offering a bribe. So, the officer is in a very strong epistemic position with respect to the proposition 'Bassa offered a bribe'. Indeed, [Lee and Pinker \(2010\)](#) found that, when presented with a similar dialogue, most subjects estimated that there was a 99% probability that the speaker was offering a bribe.

(2): In order to properly hold a subject responsible for an act of ϕ ing our judgment that they ϕ ed must have sufficient epistemic standing. We are criticizable if we hold a subject responsible without sufficient epistemic standing. We'll turn to the question of what conditions are sufficient for such standing in section 2. For now, it suffices to note that if our police officer was in an epistemic position to hold Bassa responsible for offering a bribe then she would have done so. After all, there is no social or emotional cost associated with holding him responsible. Indeed, we might imagine that there would be benefits: our officer would become known for her integrity. This might be something that she values highly. There are practical barriers to her holding him responsible, but these all seem to derive from the fact that she is not epistemically well-positioned to do so. For example, she knows that the charge would never stick, and this is precisely because Bassa's careful phrasing leaves her with insufficient resources to pin him down as having offered a bribe.

(3): In general, if (A) a subject has ϕ ed, (B) they are responsible for having ϕ ed, and (C) we have compelling undefeated evidence for (A) and (B), then we can hold them responsible for ϕ ing.³ Certainty is not required. If it was, then we would hardly ever be in a position to hold others responsible for wrongdoing. The standard required for criminal prosecution is that the admissible evidence put the defendant's guilt 'beyond reasonable doubt'. This falls short of certainty. And, plausibly, much less is required for us to merely blame or hold others responsible on a non-institutional level. So, if a subject ϕ ed, and it genuinely is clear to us

³A subject can ϕ without being responsible for having ϕ ed. In such circumstances, it will be inappropriate to hold the subject responsible for ϕ ing. Furthermore, there are different forms of responsibility with distinct preconditions. Plausibly, to hold a subject responsible in a given manner, one must be epistemically well positioned not only with respect to the proposition that they ϕ ed, but also the proposition that the further preconditions for the relevant form of responsibility are satisfied. For influential discussions of the different forms of responsibility see [Watson \(1996\)](#) and [Shoemaker \(2011\)](#).

that they ϕ ed, we ought to be able to hold them responsible for ϕ ing.

So, each of (1)-(3) has some prima facie plausibility. Yet, they can't all be true. Hence the puzzle. The puzzle is not difficult to solve. Indeed, astute readers may have already identified the solution: The term 'clear', as it appears in (1), and (3), is ambiguous. There is a sense in which it is clear to the officer that Bassa is offering a bribe. But the fact that Bassa's intentions were 'clear' in this sense does not put the officer in a position to hold him responsible. Our first task, then, is to identify a sense of 'clear' according to which it can be clear to a subject A that a subject B ϕ ed without A being in a position to hold B responsible for ϕ ing (this, in turn, will require saying a little more about the conditions under which we are in an epistemic position to hold a subject responsible for ϕ ing). Our second task will be to provide a picture of the epistemology of communication that explains how it is possible for Bassa to render his intention clear in this (and only this) sense.

We'll proceed as follows: in section 2 I will outline an analogous puzzle that has already received significant attention. A popular solution to this puzzle distinguishes between evidence that puts us in a position to know that a subject has ϕ ed, and evidence that merely rationalizes an arbitrarily high credence that the subject has ϕ ed. Only evidence of the former type puts us in a position to hold subjects responsible. In section 3 I'll present a picture of the epistemology of communication that allows us to extend this solution to the puzzle of deniability. I'll argue that, in cases of plausible deniability, audiences are provided with evidence that rationalizes a high credence that the speaker has ϕ ed without putting them in a position to know that the speaker ϕ ed. So, there is a sense in which it is clear to our police officer that Bassa was offering a bribe: The officer's evidence supports a very high credence that he was offering a bribe. But there is another sense in which it was not clear that Bassa was offering a bribe: the officer is not in a position to know that he was offering a Bribe.

The account of deniability I present has much in common with several recent approaches. It is a close relative of the view of deniability recently advocated by [Dinges and Zakkou \(2023\)](#). It also has much in common with the views of [Fricker \(2012, 2023\)](#), [Peet \(2015\)](#), and [Camp \(2018\)](#). Indeed, it shares with these views the features that lead Dinges and Zakkou (2023) to reject them - features that, according to Dinges and Zakkou (2023), lead to an untenable skepticism about testimonial knowledge. Thus the following questions arise: 1. What are the advantages of the view advocated here over Dinges and Zakkou's (2023) account? 2. Can the skeptical result be avoided? I address these questions in sections 4 and 5. In section 4 I argue that Dinges and Zakkou's (2023) approach does not provide a satisfactory resolution to the puzzle of deniability. In section 5 I will argue that the skeptical result can be resisted.

2 A Parallel Problem

Suppose a rodeo is attended by 1000 people. However, very few attendees have purchased tickets (10 to be exact). The rest of the attendees gatecrashed. Furthermore, suppose we know that Mickaël was at the rodeo, and that he is no more or less prone to gatecrashing than anyone else. Given our evidence, there is a 99% probability that he jumped the fence without paying. Yet, most feel that it would be illegitimate to convict him on this basis (C.f. [Cohen \(1977\)](#)). The evidence establishes a high probability of guilt. But this does not seem sufficient for conviction.⁴

Why is this? The answer seems to be that we are in a lottery-like situation with respect to Mickaël's guilt. Suppose that you possess a ticket in a very large lottery. You know that the chances of winning are extremely low. So, it is rational for you to assign a high credence to the proposition that your ticket is a loser. However, you also know that one ticket will win. So you don't *know* that your ticket is a loser. Your belief fails to constitute knowledge because it could easily be false, and you are unable to rule out the (relevant) possibility that your ticket is a winner.⁵ Arguably, you are not even justified in believing that it is a loser ([Nelkin \(2000\)](#), [Smith \(2010, 2016\)](#), [Littlejohn \(2012\)](#)). After all, if you know that you don't know that p , then surely you are not justified in judging that p . If you were, you would be justified in judging the Moore paradoxical ' p and I don't know that p ' ([Littlejohn \(2012\)](#)). Moreover, your judgement lacks what [Smith \(2010, 2016\)](#) calls 'normic support': Given your evidence, it could just so happen that your ticket is a winner. This would not call out for explanation. So, you are not justified in believing that it is a loser.

Of course, you are justified in assigning a high credence to the proposition that your ticket is a loser, and likewise judging that a loss is highly probable.

Our situation with respect to Mickaël is similar. We know that 99% of the rodeo attendees didn't pay. If we pick any arbitrary attendee there will be a 99% probability conditional on our evidence that the attendee gatecrashed. So, it looks like we have good reason to believe of any arbitrary attendee that they gate crashed. However, if we applied this reasoning across the board we would be sure to form a false belief. After all, we know that ten attendees did pay. So, we do not know, and may not be justified in outright believing, that Mickaël gatecrashed the rodeo. Despite the high probability of guilt, Mickaël could easily be innocent. We are unable to rule out the (relevant) possibility that he was one of the 10 legitimate ticket holders. This would not cry out for explanation or constitute a deviation from normality.

Of course, in the lottery case it will be natural to act in many ways as if your ticket is a loser. For example, you won't book any expensive holidays. And in the gatecrasher case it will be natural to act in many ways as if Mickaël gatecrashed.

⁴This problem is widely discussed. See e.g. [Buchak \(2014\)](#), [Blome-Tillman \(2017\)](#), [Pritchard \(2018\)](#), [Smith \(2018\)](#), [Gardiner \(2019\)](#), and [Littlejohn \(2020\)](#) for characteristic discussions.

⁵See [Dretske \(1970\)](#), and [Gardiner \(2019\)](#) on relevant alternatives.

For example, you'll be sure to check that Mickaël has payed if you see him at your own rodeo. However, it won't be permissible to act in *all* respects as if he gatecrashed. Most relevantly, it will not be fitting to hold him responsible for gatecrashing. There are several possible explanations for this. The two most obvious suggestions are as follows: First, perhaps it is unfitting to hold a subject responsible for ϕ ing if we don't *know* that they ϕ ed.⁶ Secondly, it may be that it is unfitting to hold somebody responsible for ϕ ing if we are not justified in outright believing that they ϕ ed. If we accept that outright belief is not justified in lottery cases then we will, again, have a straightforward explanation of the failure of our evidence to support blaming attitudes.

In what follows we'll assume that lack of knowledge, rather than lack of justification, is playing the crucial role here. That is, we'll assume that it is epistemically inappropriate to hold an agent responsible for ϕ ing if we do not know that they ϕ ed. Little hangs on this. An analogous account of deniability could easily be developed on the assumption that it is lack of justification that blocks standing to blame. Ultimately, the crucial lesson is that when we have a justified high credence that '*S* ϕ ed' but cannot rule out all the relevant or nearby '*S* didn't ϕ ' possibilities we can't fittingly hold *S* responsible for ϕ ing. The account of deniability I develop in what follows should be seen as conditional on this claim.

3 The Communicative Lottery

The problem of plausible deniability is similar to the gatecrasher problem. In both cases our evidence supports a high credence that a subject performed a particular action ϕ . However, in each case there is an epistemic obstacle standing in the way of our holding the subject responsible.

The gatecrasher case seemed lottery-like. In both lottery and gatecrasher cases there are a number of situations which, for all we know, we could easily be in. These situations differ with respect to which ticket wins, and who gatecrashed the rodeo. In the vast majority of these situations our judgement that 'this ticket is a loser' or 'Mickaël gatecrashed' is true. So, each proposition has a high probability. However, these propositions are each false in a small subset of the situations that are both live possibilities, and consistent with our evidence.

There is a sense in which communication can be similarly lottery-like. Communicative exchanges typically take place in circumstances of epistemic uncertainty. When a speaker crafts an utterance and an audience interprets it they will each rely on their representations of the context. That is, they will rely on their judgements regarding what is presupposed by their conversational partner, what is mutually salient between them, the goal of the conversation, and innumerable other factors. The speaker, when crafting their utterance, will have to estimate how the audience

⁶Why would fitting responsibility judgements require knowledge? Well firstly, it could be that knowledge is the norm of practical reason (Hawthorne & Stanley (2008)). Secondly, it may be that one must know that *p* in order to act for the reason that *p* (Hyman (1999)).

is representing the context. And the audience, when interpreting an utterance, will have to estimate how the speaker is representing the context.

Typically, this uncertainty will not lead to communicative difficulties. Usually when there is uncertainty about the context, we will craft our utterances in such a way that roughly the same thing is expressed whatever context we could reasonably be inhabiting (c.f. [Stalnaker \(1978\)](#)). At the very least, we will usually craft our utterances in such a way that we make a relevant and cooperative contribution in every context we take to be a live possibility. However, we don't always succeed in doing so. Sometimes we misjudge which contexts are live possibilities. Such misalignment in the speaker and hearer's representations of the context is a common source of misinterpretation. Suppose that in a context C_1 an utterance of S would express p . And suppose that in another context C_2 an utterance of S would instead express q . If a speaker states S thinking they are in context C_1 but the audience thinks they are in C_2 then they will miscommunicate. The speaker will expect the audience to recover p , but the audience will actually recover q . Situations like this are by no means uncommon. As good as we are at communicating, we do sometimes get it wrong.⁷

This is what allows for the possibility of plausible deniability. When interpreting an utterance there will be a number of possible situations the audience could, for all they know, easily be in. These situations correspond to the ways in which the speaker could realistically be representing the context. When a speaker maintains plausible deniability they craft their utterance in such a way that it expresses one claim (say, p) relative to almost all the contexts that are live possibilities for the audience. However, they ensure that their utterance expresses something else, say q , relative to a small subset of the contexts that are live possibilities for the audience.⁸ This puts the audience in a difficult situation. They are justified in having a high credence that the speaker intended to communicate p . After all, p is expressed relative to the vast majority of the situations that, for all they know, they could easily be in. However, there will still be a small set of contexts that are plausible enough to be live possibilities in which the utterance expresses some alternative proposition q . So, just like in the lottery and gatecrasher cases, the audience does not *know* that the speaker intended to communicate p . Depending on our view of justification, they may not even be justified in outright believing that the speaker intended p .⁹

⁷There is an extensive psychological literature on the sources of miscommunication, and the biases that lead us to underestimate just how frequently miscommunication occurs. See [Mustajoki \(2012\)](#) for a useful overview.

⁸It is important that these contexts genuinely are live possibilities for the audience. That is, it is important that the audience could, for all they know, easily be occupying a context in which q is expressed. Otherwise the speaker's denial will not be plausible, and their ability to avoid responsibility will not be due to any *epistemic* barrier faced by the audience.

⁹[Mazarella \(Forthcoming\)](#) provides a detailed account of one way in which uncertainty about context can be manipulated (or induced) to produce successful denials. Mazarella's aims are slightly different to my aims here. My aim is to identify the general epistemic structure of cases in which speakers *have* plausible deniability. Mazarella provides an account of one cognitive mechanism (among potentially many) by which speakers are able to produce successful denials.

To illustrate this, let us return to our opening dialogue. Bassa says to the officer 'Look, I have plenty of cash on me. Is there any way that we could settle this right now?'. This provides the officer with strong evidence that Bassa is offering a bribe. It is rational for our officer to have a high credence that he is offering a bribe. However, our officer cannot rule out the possibility that he is trying to pay a fine. It is unlikely that Bassa is this naive. But the population is large, and certainly contains at least a few people who are this naive. Moreover, our police officer knows nothing about Bassa. To our officer he is little more than a random sample from this large population. Thus, the officer cannot rule out that Bassa is among the sub-population naive enough to believe they can pay speeding fines at the roadside.¹⁰

So, our officer is justified in having a high credence that Bassa was offering a bribe. But she does not know that he was offering a bribe, and may not be justified in outright believing that he was offering a bribe. As mentioned earlier, a high credence in p typically renders it permissible to act in most ways as if p is true. To use an example from Ross (2021, p 16), suppose you have strong statistical evidence that your neighbour is a serial killer. Even if you don't know, and are not justified in outright believing that your neighbour is a serial killer, your high rational credence will influence your behaviour in many ways. You won't let your neighbour babysit your kids, you won't share your spare key with them, and you won't accept their offer to go on a late night hike across the moors. Similarly, even if our police officer does not know, and is not justified in outright believing that Bassa was offering a bribe, she will be justified in acting in many ways just as if Bassa has offered a bribe. For example, if she is willing to accept a bribe she might refrain from confronting him, and instead respond to his initial statement along the lines of 'yes, I think we can work something out'. Similarly, if she is not willing to accept a bribe, she can warn him of the legal consequences of bribing a police officer. However, she won't be justified in acting in all ways as if Bassa offered

However, the two accounts dovetail nicely. Mazarella suggests that an audience's willingness to accept a denial will, at least in part, be a function of the extent to which the denier's proposed re-interpretation achieves relevance (i.e. strikes an optimal balance between cognitive effects and cognitive effort). If the proposed re-interpretation achieves a high level of relevance then this constitutes evidence that it was the originally intended interpretation. Hence, it serves to either A) render epistemically relevant the possibility that the proposed re-interpretation was the originally intended interpretation (hence, the denial generates deniability), or else B) raise to salience the fact that this interpretation always was epistemically relevant, and thus that the speaker had deniability all along. Conceivably, a well crafted denial along the lines set out by Mazarella may also raise to *psychological* salience possibilities that ought to be ignored (i.e. possibilities that do not truly achieve the status of epistemically relevant possibilities). Thus, it is possible that such denials sometimes generate implausible deniability (at least when the audience is sufficiently credulous - such cases will be briefly discussed later). For recent empirical results that provide some support for a picture along the lines of Mazarella's see Bonalumi, Bumin, Scott-Phillips, & Heintz (2023).

¹⁰What if our officer did know Bassa? Well, if our officer knew that Bassa knew the norms for paying traffic fines then Bassa would lack plausible deniability. There would be nothing stopping our officer from holding Bassa responsible.

a bribe. In particular, she will not be justified in blaming Bassa, or holding him responsible for offering a bribe. Blaming or holding responsible requires knowledge.

So far the model of deniability I have presented is restricted in scope. I have given a story about how speakers are able to utilize indirect speech, context sensitivity, and the general messiness of communication to maintain deniability with respect to their intention. However, the basic idea behind this picture is extremely simple, and generalizes beyond this limited range of cases. An agent maintains plausible deniability with respect to their act of ϕ ing insofar as others are unable to know that they ϕ ed. For example, there are cases of deniability in speech that do not depend on the manipulation of context sensitivity or indirectness. As [Saul \(2017\)](#) notes, speakers will sometimes utilize what she calls 'fig-leaves': qualifications of a statement that serve to obfuscate the original claim, and place the speaker at some distance from the natural interpretation. Insofar as fig-leaves generate *plausible* deniability, they do so by preventing the audience from knowing what was meant. Similarly, deniability can be enhanced by mumbling, or speaking in a noisy environment. After all, mumbling or background noise increases the scope for interpretative error. Indeed, the core idea generalizes beyond deniability in speech: For example, when a military officer has plausible deniability with respect to a clandestine operation those with respect to whom they have deniability are unable to *know* that the officer knew of or authorized the operation.

4 A Rival Account

The view of deniability developed in the previous section holds that an agent maintains deniability with respect to their act of ϕ ing only when others are unable to know that they ϕ ed. [Dinges and Zakkou \(2023\)](#) argue that, at least on certain readings, the views of deniability advocated by [Fricker \(2012, 2023\)](#), [Peet \(2015\)](#), and [Camp \(2018\)](#) also have this consequence. They worry that any such view will entail an untenable testimonial skepticism. Their argument is not set out explicitly, but can be reconstructed as follows (C.f. [Dinges and Zakkou \(2023\)](#) p379 & p384):^{11,12}

1. A subject *S* maintains deniability with respect to their action ϕ only if

¹¹Thanks to Alexander Dinges & Julia Zakkou (personal communication) for clarifying the structure of their argument.

¹²This is not the only criticism [Dinges and Zakkou](#) raise against these views. They also maintain that speaker trustworthiness increases deniability, and that the aforementioned accounts fail to account for this fact. It is not clear to me that trustworthiness does increase *plausible* deniability. It may be that a trusted speaker is better able to convince their audience after the fact that they intended something different. Indeed, it may be that a denial offered by a trustworthy speaker more easily *generates* plausible deniability than a denial offered by an untrustworthy speaker, since such a denial renders more epistemically relevant the possibility that they really did mean something different (i.e. interpretative possibilities become epistemically relevant for the genuinely trustworthy speaker that remain irrelevant for the untrustworthy speaker). But this does not entail that they have any more deniability at the time of utterance (i.e. before a denial is offered) than the untrustworthy speaker.

those with respect to whom they maintain deniability are not in a position to know that $S \phi$ ed (assumption for reductio).

2. Testimonial knowledge requires knowledge of what is said.
3. Deniability is widespread - speakers very often have plausible deniability with respect to the contents of their testimony.
4. Therefore, audiences very often fail to know what the speaker has said (from (1) & 3).
5. Therefore, audiences very often fail to gain testimonial knowledge (from (2) & 4).

If sound, this argument suggests that audiences very often fail to gain testimonial knowledge even when the speaker is trustworthy, rationally trusted, and speaking from knowledge. Dinges and Zakkou deem this result to be unacceptable. So, they suggest, one of the premises must be false. They take issue with premise (1), and provide an account of deniability consistent with its rejection. That is, they hold that an agent S can maintain deniability with respect to their act of ϕ ing even when those with respect to whom they maintain deniability are in a position to know ' $S \phi$ ed'. In this section I'll discuss Dinges & Zakkou's account, arguing that despite its similarity to the account I have argued for here, it is unable to provide an adequate solution to the puzzle of plausible deniability. I'll return to the skeptical argument in section 5.

Dinges and Zakkou provide two versions of their account. The first, which we will call the 'conditional knowledge account' is as follows:

Conditional Knowledge: 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 . [Dinges & Zakkou, 2023](#), p 385.

The second, what we will call the 'conditional rationality account' is as follows:

Conditional Rationality: 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 actions or emotions. [Dinges & Zakkou, 2023](#), p 391.

On both accounts, deniability with respect to ϕ is consistent with the audience knowing (until the denial is actually produced) that the speaker ϕ ed. The conditional knowledge account merely requires that the audience would no longer know that the speaker ϕ ed were the speaker to offer a denial. The conditional rationality approach is even less demanding. It merely requires that the audience would no longer be in an epistemic position to treat ' $S \phi$ ed' as a reason for action or emotion in the event of a denial. If the conditions for rationally treating p as

a reason for action or emotion are even stronger than knowledge then the conditional rationality account will render deniability consistent with the audience's retaining knowledge that the speaker ϕ ed even after a denial has been offered.

In this section we'll primarily focus on the conditional knowledge account. This is partly for reasons of space, but also because the conditional knowledge account is closer to the non-conditional account defended in section 3. Many of the criticisms provided generalize straightforwardly to the conditional rationality account.

The conditional account appears to provide an alternative solution to the puzzle of deniability: Firstly, we can hold that the officer in our core example initially knew that Bassa was offering a bribe. This captures the sense in which it was clear to the officer that Bassa was offering a bribe. Secondly, we can capture the sense in which there was an epistemic barrier preventing the officer from holding Bassa responsible by holding that once Bassa denies offering a bribe the officer is no longer be in a position to know that he was offering a bribe (or, at least, the officer is no longer in an epistemic position to act on this proposition).

This response to the puzzle of plausible deniability faces a number of problems. Firstly, consider the claim that our officer loses her knowledge after Bassa denies offering a bribe. There are two ways this could happen. 1. Bassa's denial could rationally defeat the officer's belief. 2. His denial could cause the officer to lose her belief without rationally defeating it. Lets consider the second case first: the officer loses her belief even though Bassa's denial does not rationally defeat it. In this case her change in belief will be irrational. If so, the puzzle of plausible deniability is left unresolved. After all, this diagnosis rests on the assumption that our officer is epistemically irrational. But the officer need not be irrational for Bassa to maintain deniability. A central feature of the puzzle was the existence of an *epistemic* barrier preventing our officer from holding Bassa responsible. On this diagnosis there is no epistemic barrier. Our officer still has sufficient evidence to hold Bassa responsible.

To put the point another way, this approach renders our opening case an instance of implausible deniability rather than one of plausible deniability. And this highlights a more general problem with the conditional approach: it classifies many implausible denials as plausible denials. Consider the following:

Irrational Denial: Adam and Stefano are trying to decide where to go for lunch. Adam says 'the Italian restaurant round the corner is good'. Stefano replies 'no, its terrible'. Adam then responds 'I know, by 'good' I meant nearby'. Because of his irrational and unyielding trust in Adam Stefano accepts Adam's denial.

The conditional account predicts that Adam has plausible deniability. Were Adam to deny meaning that the restaurant was good (which he did) Stefano would no longer know (because he no longer believes) that Adam meant that the restaurant was good. However, this is clearly not a plausible denial. The success of Adam's denial rests not on its plausibility, but upon Stefano's irrational trust. It

is, thus, a case of implausible deniability.¹³ In light of this, it is natural to modify the account as follows:¹⁴

Conditional Knowledge +: S has deniability relative to the proposition that she ϕ ed iff: if S denies that she ϕ ed, then S's audience *is not in a position to know* that she ϕ ed.

The idea here would be that a subject has plausible deniability with respect to her having ϕ ed when either the audience is not in a position to know that she has ϕ ed, or else her denial would rationally defeat the audience's knowledge that she ϕ ed. And this takes us back to the first possibility mentioned above regarding our core case: that Bassa's denial rationally defeats the officer's knowledge. So, could Bassa's denial have rationally defeated our officer's knowledge?

A number of epistemologists have denied the very possibility of knowledge defeat (Lasonen-Aarnio (2010), Baker-Hytch & Benton (2015)). So, for these epistemologists the answer will be 'no'. However, if this is the case then the conditional theorist can simply fall back on the conditional rationality account, holding that denials can epistemically defeat rational standing to blame without defeating knowledge.

So, let's assume that knowledge can be defeated. Is it plausible that audiences in paradigmatic deniability cases start off knowing what the speaker was doing, only to have that knowledge defeated by the speaker's denial? No. In paradigmatic deniability cases - cases like that with which we started - audiences know that if they challenge the speaker they will be met with a denial. What they do not know is whether or not this denial will be sincere. Consider our core case again: If Bassa was offering a bribe he would deny this when challenged, and if he was merely attempting to pay a roadside fine he would also deny attempting to offer a bribe. Given this, it is unclear how a denial could possibly defeat our officer's knowledge. If you already know that you are about to receive evidence E , and you already know the content of E , then it is not clear how one's actually receiving E could defeat one's knowledge (or any other positive epistemic status). Indeed, it is not clear in what sense E would really constitute new evidence at all.

If this is right, then the conditional account does not resolve the puzzle of plausible deniability. In order to capture the sense in which it is obvious to our officer that Bassa was offering a bribe the conditional theorist has to maintain that the officer starts out knowing that he was offering a bribe, but loses this knowledge in the face of Bassa's denial. But, unless we assume that the officer is irrational (which we needn't) the only way for the officer to lose her knowledge is for Bassa's denial to defeat her knowledge (or rational standing to blame). And his denial cannot have this effect.

¹³This objection extends to the conditional rationality account as long as we assume that one must believe p in order for it to be epistemically proper for one to treat p as a reason for action or emotion.

¹⁴Dinges & Zakkou 2023 consider a similar modification to their account on p 385, fn 17.

To be clear, I am not suggesting that acts of denial *never* have any impact on audiences. The claim is far weaker: All that is required is that there are cases of deniability (such as our core example - which I take to be paradigmatic of plausible deniability) in which the speaker possesses deniability with respect to their having ϕ ed even though the actual act of offering a denial would do nothing to change the audience's epistemic position with respect to the proposition that the speaker ϕ ed. There are many other ways for a denial to have an impact on an audience. For example, it is possible for a speaker to possess deniability without the audience's realizing they have deniability. That is, the audience may mistakenly believe that they know that the speaker ϕ ed despite their inability to rule out an epistemically relevant possibility in which the speaker did not ϕ . In this case the act of denial may cause the audience to realize that the speaker possessed deniability all along. Secondly, an act of denial may generate plausible deniability when the initial utterance was not presented in a manner that generated immediate deniability. That is, a skillful denial may render epistemically relevant an error possibility that was, initially, too far-fetched to be taken seriously. Thirdly, as already noted, an act of denial can cause the audience to believe that the speaker has plausible deniability even when they do not. That is, the act of denial may cause a sufficiently credulous audience to take seriously error possibilities that they ought to ignore (I take **Irrational Denial** to be an example of this).¹⁵

I believe the above discussion suffices to demonstrate that the non-conditional account of plausible deniability is preferable to the conditional account. However, for the sake of completeness it is worth mentioning a number of further worries. Firstly, the conditional theorist assumes that before Bassa offers his denial the officer knows that he is offering a bribe. If this is the case then, at least initially, the officer is in a position to hold Bassa responsible. But this seems wrong: the officer is never in a position to hold Bassa responsible because of the availability of his denial. Otherwise, the officer would be wise to arrest Bassa on the spot, before he has a chance to deny his intention.¹⁶

Furthermore, plausible deniability undermines commitment - when a speaker has deniability regarding their intention to communicate p they are not committed to p . But the conditional knowledge account tells us that in some cases of deniability the audience will start out knowing precisely what the speaker intended, only to have that knowledge potentially defeated by a later denial. Since such audiences know what the speaker intended they should be able to hold the speaker responsible and blame them until the denial is offered. But this entails that the

¹⁵Indeed, it is consistent with what I have said here that there are cases in which A) the original utterance generates deniability, and B) the audience knows that a denial will be issued if the speaker is challenged, but C) the actual act of denial serves to generate even more deniability. This will occur when a denial is offered that is 1. distinct from the form of denial expected by the audience, and 2. such as to render epistemically relevant new error possibilities in addition to those that were rendered relevant by the original framing of the utterance.

¹⁶This assumes that knowledge that a subject ϕ ed is sufficient for properly holding the subject responsible for ϕ ing. The conditional rationality account is not committed to this claim, and thus avoids this additional worry.

audience can fittingly hold the speaker to p and blame them if p is false despite the speaker's not being committed to p .

These problems show that the conditional theorist's response to the puzzle of deniability is inadequate.¹⁷ But, before moving on to consider the skeptical argument that motivated the conditional account, it is worth considering a final more general problem for the conditional account: it is subject to counterexample. Consider the following:

Bad Liar: Yoshiyuki and Futaba are discussing their friend Miho. Futaba asks whether Miho is seeing anyone. Yoshiyuki replies 'well, she's been making lots of trips to town'. The implication is that Miho is seeing somebody in town. But this is deniable. Somebody in Yoshiyuki's position could reply 'I didn't mean that, I meant that she could be hoping to meet someone in town'. However, suppose that Yoshiyuki is extremely nervous and is, as a result, a terrible liar. If he was actually challenged he would respond as follows: 'Well I um, well no thats not, thats....I think thats not what I...umm..no, I didn't mean that I meant that... well I guess...erm, maybe I meant she's trying to meet somebody in town?'

Yoshiyuki's incompetent denial will do little to weaken Futaba's epistemic position. On the contrary, it will strengthen it. So, the conditional accounts tell us that Yoshiyuki lacks deniability. But this is wrong. He has deniability, he's just disposed to mess up the denial. Indeed, his act of denial destroys his deniability.

Of course, such counterexamples can be dealt with by complicating the conditional account. We might, for example, hold that a subject has deniability only if a *competent* denial would defeat the audience's knowledge.¹⁸ But, as is often the case with conditional analyses in philosophy, such modifications simply give rise to more baroque counter-examples (for example, we may posit that counter-evidence will systematically materialize whenever a competent denial is offered).¹⁹ This, in

¹⁷It might be thought that the conditional theorist can, in response, simply piggyback on the non-conditional account given in section 3. That is, they could accept that the officer never knew that Bassa was offering a bribe, and that the sense in which Bassa's intentions were clear to the officer can be captured in terms of rational credence. But if they do so then it is hard to see why we should accept the conditional account over the non-conditional account. After all, the puzzle of deniability is a general one. It is a general feature of plausible deniability that we are not in an epistemic position to hold agents responsible for the actions with respect to which they have deniability. The diagnosis offered in section 3 is similarly general: when an agent has deniability with respect to having ϕ ed those with respect to whom they have deniability are not in a position to know that the agent ϕ ed. If this leads to skeptical consequences for the non-conditional theorist, then it is hard to see how an approach that simply piggybacks on the non-conditional account's diagnosis can avoid skepticism. Moreover, I'll argue in what follows that the skeptical result is not worrying. Thus, even if the conditional theorist can avoid it, it obtains little advantage from doing so.

¹⁸Indeed, Dinges & Zakkou consider a similar worry and propose a similar response (p 386, fn18.).

¹⁹Here is a variant of **Bad Liar** along these lines: Yoshiyuki is a perfectly competent liar. However, unbeknownst to both Yoshiyuki and Futaba there is a lie detecting robot being tested

turn, leads us down a rabbit hole of more and more convoluted variants on the original analysis.

Furthermore, there is a more general problem underwriting these counterexamples: Intuitively, an utterance's level of deniability depends on how it is framed and presented. A cleverly framed utterance has a high level of deniability. A ham-fisted attempt does not. A speaker's ability to actually capitalize on the deniability of their utterance is a different matter. The conditional account does not capture this. On the conditional account, deniability ceases to be a function of how the original utterance is framed. Two identical utterances produced in identical circumstances by identical seeming speakers can differ radically with respect to how much plausible deniability the speaker has at the time of utterance.

In light of this, the prospects for the conditional account look dim. The conditional theorist's diagnosis of the puzzle of deniability does not work, and the view, at least as currently formulated, is subject to counterexample. In light of this, it is worth returning to the skeptical argument that was used to motivate the conditional analysis over non-conditional accounts like the one developed here (as well as those presented by Fricker (2012, 2023), Peet (2015), and Camp (2018)).

5 Testimonial Skepticism?

Dinges and Zakkou's (2023) skeptical argument runs as follows (C.f. p4 & p10):

1. A subject S maintains deniability with respect to their action ϕ only if those with respect to whom they maintain deniability are not in a position to know that S ϕ ed (assumption for reductio).
2. Testimonial knowledge requires knowledge of what is said.
3. Deniability is widespread - speakers very often have plausible deniability with respect to the contents of their testimony.
4. Therefore, audiences very often fail to know what the speaker has said (from (1) & 3).
5. Therefore, audiences very often fail to gain testimonial knowledge (from (2) & 4).

I have suggested that we should retain premise (1). So, what about premises (2) and (3)? Premise (2) is highly controversial. Longworth (2018), Peet (2018, 2019) and Pollock (Forthcoming) have all argued at length that testimonial knowledge does not require knowledge of what is said. We don't have space to go

in the next room. If Yoshiyuki offered a denial the robot would overhear, burst into the room, and call him out for lying. Here Yoshiyuki has deniability (Futaba can't rule out the possibility that he meant that Miho was going to town in the hopes of meeting somebody - she does not *know* that this is what he meant). However, the conditional fails. If Yoshiyuki offered a denial he would no longer have deniability.

through these arguments here. But it is notable that some cases of deniability constitute plausible examples of testimonial knowledge without knowledge of what is said: A speaker can maintain deniability by ensuring that their utterance admits of multiple reasonable interpretations. And, in doing so, they can ensure that their utterance expresses something true on every reasonable interpretation. If they do so then there will be no way for a reasonable and competent audience member to form a false belief by accepting what they take the speaker to be asserting. The audience will, thus, be safe from error in their testimonial belief (although they will not be safe from error in their belief that 'S intended p '). In such cases, my inclination would be to attribute testimonial knowledge to the audience.

Still, this conclusion will likely be controversial. The argument above was very quick and, as Dinges and Zakkou point out, it is still commonly accepted that testimonial knowledge requires knowledge of what is said. So, let us accept for the sake of argument that testimonial knowledge does require knowledge of what is said.

This leaves us with premise (3), which tells us that plausible deniability is very widespread. Is this correct? Well, linguistic communication is rife with context sensitivity. And, context sensitivity very often gives rise to *deniability*. As I have argued in previous work (Peet (2015)), a sufficiently inventive speaker will often be able to give a story according to which the audience's interpretation of a context sensitive sentence was mistaken. And, in doing so, they will often be able to induce sufficient doubt about their intention that the audience will not attempt to hold them responsible. In light of this, it may appear that deniability is extremely widespread.

But our interest here isn't in deniability in general. Our interest is in *plausible* deniability. A speaker can lack plausible deniability but still be able to induce sufficient doubt in the audience that they get away with their denial. **Irrational Denial** gave a stark illustration of this. But there are more subtle ways by which speakers can manipulate context sensitivity to retain deniability. For example, suppose a speaker ϕ ed, and their audience knows that they ϕ ed. The speaker may be able to induce a revision of this belief by raising to salience a far fetched error possibility according to which they did not ϕ . After all, merely mentioning such an error possibility can alter the audience's perception of its relevance or likelihood (Hawthorne (2004), Williamson (2005), Nagel (2010), Gerken (2017), & Dinges (2018)). Such implausible deniability is consistent with the audience knowing, at least initially, that the speaker ϕ ed. So, the widespread nature of such deniability does not give rise to skeptical worries.

Speakers typically maintain *plausible* deniability by exploiting the fact that there's usually a number of slightly different contexts the audience could, for all they can reasonably tell, be occupying. To communicate p in a deniable way the speaker frames their utterance in such a way that it expresses p relative to the majority of these possible contexts, and some other proposition q relative to a small subset of these contexts. So, for there to be widespread plausible deniability regarding the contents of assertions it will have to be the case that assertions very

often admit of more than one reasonable interpretation. That is, it will have to be the case that, very often, there are epistemically nearby possibilities – the sort of relevant possibilities we have to be able to rule out in order to possess knowledge, in which assertions express something other than the proposition intended by the speaker.

It is impossible to settle a priori how often this situation arises. Thankfully, we don't need to settle this issue in order to defend the non-conditional account of plausible deniability. There are two reasons for this. Firstly, the non-conditional account is not committed to the claim that plausible deniability is widespread. So, the non-conditional theorist is free to deny premise (3). Secondly, plausible deniability will only be widespread if a fairly extreme form of communicative unreliability is also widespread. And, if such unreliability *is* widespread then skeptical worries arise regardless of our views about deniability (at least if we assume that testimonial knowledge requires knowledge of what is said). That is, premise (3) is only plausible insofar as the conclusion (5) is already antecedently plausible. So, if (3) is accepted then the conclusion (5) should be accepted as well.²⁰

This defense of the non-conditional view might be met with resistance. An opponent could respond as follows: The non-conditional account assumes that the existence of nearby error possibilities consistent with the audience's evidence precludes knowledge of what is said (call this the 'error assumption'). If we discover that nearby error possibilities are present in most normal communicative exchanges then the error assumption would entail an unpalatable testimonial skepticism. Thus, if we discover that such error possibilities are widespread then we should reject the error assumption, and with it the non-conditional account of deniability. This rejoinder only gets off the ground if we have prior to reason to accept that nearby error possibilities are present in most normal communicative exchanges. But this is a possibility we should take seriously. After all, there is empirical research that suggests we are significantly less reliable as communicators than we often assume.²¹

Should the non-conditional theorist be worried? No. Firstly, the error assumption only generates skeptical worries in conjunction with the assumption that testimonial knowledge requires knowledge of what is said. As noted above, this assumption has come under sustained critique in the recent literature, and it is certainly less fundamental to our epistemic theorizing than the hypothesis that nearby error possibilities undermine knowledge. So, a more promising response to the skeptical threat would be to reject the knowledge of what is said assumption.

But even this might not be necessary. The force of the skeptical challenge rests on the assumption that testimonial skepticism is untenable. But the skepticism that arises from communicative unreliability is far less worrying than it initially appears. Granted, testimonial skepticism sounds pretty bad. The vast majority of our beliefs are acquired via (or are in some way dependent on beliefs acquired

²⁰Thanks to an anonymous reviewer for helping with the framing of these points.

²¹See for example Ferreira & Patson (2007), Karimi & Ferreira (2016), and Christianson (2016).

via) communication with others. So, if the majority of communicative exchanges are rendered epistemic failures then much of what we take ourselves to know about the world will be called into question. However, the skepticism that arises from communicative unreliability does not render the majority of communicative exchanges epistemic failures.

Firstly, even if the audience can't *know* what is said, their evidence will still typically support a high rational credence that the speaker intended a particular proposition p . And if they rationally trust the speaker, this will be sufficient to justify a high rational credence that p . This will often be enough for practical purposes.

Secondly, even if the audience cannot immediately come to know the fine-grained proposition that they take the speaker to be asserting, there will usually be some coarse-grained proposition in the neighborhood that they can still come to know. For example, they can still come to know the diagonal proposition, the disjunction of the possible propositions the speaker could reasonably intend, or the proposition that is part (in the sense of [Fine \(2017\)](#)) of every proposition the speaker could realistically be attempting to express (c.f. [Davies \(Forthcoming\)](#)). This proposition will be true in every context the audience could, for all they can reasonably tell, be occupying. The speaker will lack deniability with respect to this proposition.

To illustrate the epistemic position of audiences in contexts where error possibilities give rise to plausible deniability, consider the following cases:

Honest Hiking: Lynn and Ron are hiking. Ron asks Lynn how far they are from the campsite. Lynn knows that they are still several miles from the campsite, but she is not sure of the exact distance. So, she tells Ron 'we're still some distance from the campsite'.

Dishonest Hiking: Lynn and Ron are hiking. Ron asks Lynn how far they are from the campsite. Lynn knows that the campsite isn't around the next bend. But beyond that she does not know how far they are from the campsite. Nonetheless, as the more experienced hiker she wants to give the impression of knowing roughly where they are. So, she says 'we're still some distance from the campsite'.

Ron will form the same beliefs in either case. He will very likely form the belief that they are a long way (several miles) from the campsite. After all, the phrase 'some distance' would usually be used in such contexts to communicate that they are a significant distance from the campsite. However, Lynn is able to maintain plausible deniability regarding the proposition that the campsite is several miles away. If the campsite turns out to be just a mile away she can claim that she meant they were not practically on top of it. Although not typical, it is certainly a realistic possibility that a normal speaker could mean such a thing by 'some distance away'. So, in neither case can Ron come to know that they are several miles from the campsite.

Nonetheless, at least in **Honest Hiking**, Ron can come to know something: he comes to know that they are not right on top of the campsite. To put it another way, he does come to know that they are *some* distance from the campsite. He just doesn't know how far *some* distance is. Moreover, in both cases, Ron can form a high rational credence that they are several miles from the campsite. This rationalizes him in, for example, conserving some of his water, or sitting down to have a rest. Indeed, some of these behaviors will allow him to narrow down what Lynn could have meant by 'some distance'. For example, if Lynn agrees with him that they should sit down to rest or conserve their water then he will be in a position to rationally believe and hold Lynn to the claim that they are several miles from the campsite. Normal conversations will often provide similar opportunities for audiences to narrow down what a speaker meant by their earlier utterances.²² Thus, even if the audience does not immediately acquire knowledge, the subsequent dialogue will often be sufficient to bump what was previously mere belief or high credence up to full-blown knowledge.

In light of this, it would not be especially worrying if cases like **Honest Hiking** turned out to be the norm. The form of skepticism we would be forced to accept (if we insisted on retaining the assumption that testimonial knowledge requires knowledge of what is said) would be extremely limited. Indeed, it is exactly the sort of skeptical result we should expect and accept in the event of the discovery that we are unreliable communicators.²³ Overall, then, skeptical worries put very little pressure on the non-conditional account of deniability.

6 Conclusion

The puzzle of plausible deniability is resolved. How is it that a speaker is able to make their intention clear to their audience whilst at the same time maintaining an epistemic barrier that prevents the audience from holding them responsible? They must exploit the potentially lottery like nature of linguistic communication. They must construct their utterance in such a way that it expresses one thing relative to the majority of contexts that are live possibilities, and something else relative to a minority of these contexts. This allows the audience to have a justified high credence regarding the speaker's intention. But it does not allow them to know the speaker's intention. It may not even allow them to justifiably believe that the speaker has a particular intention. And it certainly doesn't allow them to hold the speaker responsible for their intention.

²²For discussion of the opportunities dialogue affords us to narrow down and even negotiate intended meaning see Clark (1996), Allott (2016), Elder (2019), and Drożdżowicz (Forthcoming).

²³In fact, the resultant skepticism would be no more worrying than skeptical positions that have already been explicitly defended by Peet (2016), and Pollock (2023).

7 Acknowledgements

Thanks to Alexander Dinges, Julia Zakkou, and several anonymous reviewers for helpful comments and discussion. Thanks also to (remote) audiences the University of St Andrews, and at the New Models of Knowledge Communication conference at the University of Oslo where parts of this material were presented. The research leading to these results received funding from the European Research Council (Horizon 2020) under Grant Agreement No 818633.

References

- [1] Allott, N. 2016. Misunderstandings in verbal communication. In A. Rocci, & de L. Saussure (Eds.), *Verbal Communication*. Berlin: Walter De Gruyter. 485–507. doi: [10.1515/9783110255478-025](https://doi.org/10.1515/9783110255478-025)
- [2] Baker-Hyatt, M. & Benton, M. 2015. Defeatism Defeated. *Philosophical Perspectives* 29 (1). 40-66. doi: [10.1111/phpe.12056](https://doi.org/10.1111/phpe.12056)
- [3] Blome-Tillman, M. 2017. 'More Likely than Not' - Knowledge First and the Role of Statistical Evidence in Courts of Law. In Carter, A., Gordon, E., and Jarvis, B. (Eds) *Knowledge First Approaches in Epistemology and Mind* Oxford: Oxford University Press. 278-292. doi: [10.1093/oso/9780198716310.003.0014](https://doi.org/10.1093/oso/9780198716310.003.0014)
- [4] Buchak, L. 2014. Belief, Credence, and Norms. *Philosophical Studies* 169 (2). 285-311. doi: [10.1007/s11098-013-0182-y](https://doi.org/10.1007/s11098-013-0182-y)
- [5] Camp, E. 2018. Insinuation, Common Ground, and Conversational Record. In Fogal, D., Harris, D., & Moss, M. (Eds) *New Work on Speech Acts*. Oxford: Oxford University Press. 44-66. doi: [10.1093/oso/9780198738831.001.0001](https://doi.org/10.1093/oso/9780198738831.001.0001)
- [6] Christianson, K. 2016. When language comprehension goes wrong for the right reasons: Good-enough, underspecified, or shallow language processing. *Quarterly journal of experimental psychology* 69 (5). 817–828. doi: [10.1080/17470218.2015.1134603](https://doi.org/10.1080/17470218.2015.1134603)
- [7] Clark, H. H. 1996. *Using language*. Cambridge: Cambridge university press. doi [10.1017/CBO9780511620539](https://doi.org/10.1017/CBO9780511620539)
- [8] Cohen, L. J. 1977. *The Probable and the Provable* Oxford: Clarendon Press. doi: [10.1093/acprof:oso/9780198244127.001.0001](https://doi.org/10.1093/acprof:oso/9780198244127.001.0001)
- [9] Davies, A. 2021. Communicating in Contextual Ignorance. *Synthese* 199. 12385–12405. doi: [10.1007/s11229-021-03337-8](https://doi.org/10.1007/s11229-021-03337-8)
- [10] Dinges, A. 2018. Knowledge and Availability. *Philosophical Psychology* 31 (4). 554–573. doi: [10.1080/09515089.2018.1438594](https://doi.org/10.1080/09515089.2018.1438594)

- [11] Dinges, A. & Zakkou, J. 2023. On Deniability. *Mind*. doi:[10.1093/mind/fzac056](https://doi.org/10.1093/mind/fzac056)
- [12] Dretske, F. 1970. Epistemic Operators. *Journal of Philosophy* 67 (24). 1007–1023. doi: [10.2307/2024710](https://doi.org/10.2307/2024710)
- [13] Drożdżowicz, A. Forthcoming. Making it precise—Imprecision and underdetermination in linguistic communication. *Synthese*. 1-27. doi [10.1007/s11229-022-03544-x](https://doi.org/10.1007/s11229-022-03544-x)
- [14] Elder, C. H. F. 2019. Negotiating what is said in the face of miscommunication. In Piotr Stalmaszczyk (ed.) *Philosophical Insights into Pragmatics*. Berlin: Walter De Gruyter. 107-127. doi: [10.1515/9783110628937-006](https://doi.org/10.1515/9783110628937-006)
- [15] Ferreira, F., & Patson, N. D. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1(1–2). 71–83. doi: [10.1111/j.1749-818X.2007.00007.x](https://doi.org/10.1111/j.1749-818X.2007.00007.x)
- [16] Fine, K. 2017. Truthmaker Semantics. In B. Hale, C. Wright, & A. Miller (Eds.), *A Companion to the Philosophy of Language*. Wiley Blackwell. 556–577. doi: [10.1002/9781118972090.ch22](https://doi.org/10.1002/9781118972090.ch22)
- [17] Fricker, E. 2012. Stating and Insinuating. *Proceedings of the Aristotelian Society* 86 (1). 61-94. doi: [10.1111/j.1467-8349.2012.00208.x](https://doi.org/10.1111/j.1467-8349.2012.00208.x)
- [18] Fricker, E. 2023. Testimony, Luck, and Conversational Implicature. In A. Fairweather & C. ;pntemayor (Eds.) *Linguistic Luck: Safeguards and Threats to Linguistic Communication*. New York: Oxford University Press. 181-222. doi: [10.1093/oso/9780192845450.003.0008](https://doi.org/10.1093/oso/9780192845450.003.0008)
- [19] Gardiner, G. 2019. The Reasonable and the Relevant: Legal Standards of Proof. *Philosophy & Public Affairs* 47 (3). 288-318. doi: [10.1111/papa.12149](https://doi.org/10.1111/papa.12149)
- [20] Gerken, M. 2017. *On Folk Epistemology: How We Think and Talk about Knowledge*. Oxford: Oxford University Press. doi: [10.1093/oso/9780198803454.001.0001](https://doi.org/10.1093/oso/9780198803454.001.0001)
- [21] Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press. doi: [10.1093/0199269556.001.0001](https://doi.org/10.1093/0199269556.001.0001)
- [22] Hawthorne, J., & Stanley J. 2008. Knowledge and Action. *Journal of Philosophy* 105 (10). 571-590. doi: [10.5840/jphil20081051022](https://doi.org/10.5840/jphil20081051022)
- [23] Hyman, J. 1999. How Knowledge Works. *Philosophical Quarterly* 49 (197). 433-451. doi: [10.1111/1467-9213.00152](https://doi.org/10.1111/1467-9213.00152)
- [24] Karimi, H., & Ferreira, F. 2016. Good-enough linguistic representations and online cognitive equilibrium in language processing. *The Quarterly Journal of Experimental Psychology* 69 (5). 1013–1040. doi: [10.1080/17470218.2015.1053951](https://doi.org/10.1080/17470218.2015.1053951)

- [25] Khoo, J. 2017. Code Words in Political Discourse. *Philosophical Topics* 45 (21). 33–64. doi: [10.5840/philtopics201745213](https://doi.org/10.5840/philtopics201745213)
- [26] Lasonen-Aarnio, M. 2010. *Unreasonable knowledge*. *Philosophical Perspectives*. 24 (1). 1–21. doi: [10.1111/j.1520-8583.2010.00183.x](https://doi.org/10.1111/j.1520-8583.2010.00183.x)
- [27] Lee, J.J., & Pinker, S. 2010. Rationales for Indirect Speech: The Theory of the Strategic Speaker. *Psychological Review* 117 (3). 785-807. doi: [10.1037/a0019688](https://doi.org/10.1037/a0019688)
- [28] Littlejohn, C. 2012. Lotteries, probabilities, and permissions. *Logos and Episteme*, 3, 509–514. doi: [10.5840/logos-episteme20123330](https://doi.org/10.5840/logos-episteme20123330)
- [29] Littlejohn, C. 2020. Truth, Knowledge, and the Standard of Proof in Criminal Law. *Synthese* 197. 5253-5286. doi: [10.1007/s11229-017-1608-4](https://doi.org/10.1007/s11229-017-1608-4)
- [30] Longworth, G. 2018. Understanding What is Said. *Synthese* 195. 815-834. doi: [10.1007/s11229-016-1243-5](https://doi.org/10.1007/s11229-016-1243-5)
- [31] Mazzarella, D. I Didn't Mean to Suggest Anything Like That!: Deniability and Context Reconstruction. *Mind and Language*. doi: [10.1111/mila.12377](https://doi.org/10.1111/mila.12377)
- [32] Mustajoki, A. 2012. A Speaker-Oriented Multidimensional Approach to Risks and Causes of Miscommunication. *Language and Dialogue* 2 (2). 216 - 243. doi: [10.1075/ld.2.2.03mus](https://doi.org/10.1075/ld.2.2.03mus)
- [33] Nagel, J. 2010. Knowledge Ascriptions and the Psychological Consequences of Thinking About Error. *The Philosophical Quarterly* 60 (239). 286–306. doi: [10.1111/j.1467-9213.2009.624.x](https://doi.org/10.1111/j.1467-9213.2009.624.x)
- [34] Nelkin, D. 2000. The lottery paradox, knowledge and rationality. *Philosophical Review* 109. 373– 409. doi: [10.2307/2693695](https://doi.org/10.2307/2693695)
- [35] Peet, A. 2015. Testimony, Pragmatics, and Plausible Deniability. *Episteme* 12 (1). 29-51. doi: [10.1017/epi.2014.31](https://doi.org/10.1017/epi.2014.31)
- [36] Peet, A. 2016. Testimony and the Epistemic Uncertainty of Interpretation. *Philosophical Studies* 173 (2). 395-416. doi: [10.1007/s11098-015-0498-x](https://doi.org/10.1007/s11098-015-0498-x)
- [37] Peet, A. 2018. Testimonial Knowledge without Knowledge of What is Said. *Pacific Philosophical Quarterly* 99 (1). 65-81. doi: [10.1111/papq.12137](https://doi.org/10.1111/papq.12137)
- [38] Peet, A. 2019. Knowledge-Yielding Communication. *Philosophical Studies* 176. 3303-3327. doi: [10.1007/s11098-018-1175-7](https://doi.org/10.1007/s11098-018-1175-7)
- [39] Pollock, J. 2023. Linguistic Understanding and Testimonial Warrant. *Erkenntnis* 88. 457–477. doi: [10.1007/s10670-020-00362-w](https://doi.org/10.1007/s10670-020-00362-w)
- [40] Pollock, J. Forthcoming. Content Internalism and Testimonial Knowledge. *Inquiry*. doi: <https://doi.org/10.1080/0020174X.2021.1977691>

- [41] Pritchard, D. 2018. Legal Risk, Legal Evidence, and the Arithmetic of Criminal Justice. *Jurisprudence* 9 (1). 108-119. doi: [10.1080/20403313.2017.1352323](https://doi.org/10.1080/20403313.2017.1352323)
- [42] Ross, L. 2021. Rehabilitating Statistical Evidence. *Philosophy and Phenomenological Research* 102 (1). 3-23doi: [10.1111/phpr.12622](https://doi.org/10.1111/phpr.12622)
- [43] Saul, J. 2017. Racial Figleaves, the Shifting Boundaries of the Permissible, and the Rise of Donald Trump. *Philosophical Topics* 45 (2). 97-116. doi: [10.5840/philtopics201745215](https://doi.org/10.5840/philtopics201745215)
- [44] Saul, J. 2018. Dogwhistles, Political Manipulation, and Philosophy of Language. In D. Fogal, D. Harris, & M. Moss (Eds.) *New Work on Speech Acts*. Oxford: Oxford University Press. 360-383. doi: [10.1093/oso/9780198738831.003.0013](https://doi.org/10.1093/oso/9780198738831.003.0013)
- [45] Shoemaker, D. 2011. Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics* 121 (3). 602-632. doi: [10.1086/659003](https://doi.org/10.1086/659003)
- [46] Smith, M. 2010. What Else Could Justification Be? *Noûs* 44 (1). 10-31. doi: [10.1111/j.1468-0068.2009.00729.x](https://doi.org/10.1111/j.1468-0068.2009.00729.x)
- [47] Smith, M. 2018. When does Evidence Suffice for Conviction? *Mind* 127 (508). 1193-1218. doi: [10.1093/mind/fzx026](https://doi.org/10.1093/mind/fzx026)
- [48] Stalnaker, R. 1978. Assertion. In Cole, P. (Ed.) *Syntax and Semantics, Vol 9: Pragmatics* New York Academic Press. Reprinted in Stalnaker, R. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press. 78-95. doi: [10.1093/0198237073.003.0005](https://doi.org/10.1093/0198237073.003.0005)
- [49] Stanley, J. 2015. *How Propaganda Works*. Princeton, NJ: Princeton University Press.
- [50] Watson, G. 1996. Two Faces of Responsibility. *Philosophical Topics* 24 (2). 227-248 doi: [10.5840/philtopics199624222](https://doi.org/10.5840/philtopics199624222)
- [51] Williamson, T. 2005. Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge. *The Philosophical Quarterly* 55 (219). 213-235. doi: [10.1111/j.0031-8094.2005.00396.x](https://doi.org/10.1111/j.0031-8094.2005.00396.x)