

Gaslighting, First and Second Order

Paul-Mikhail Catapang Podosky
UNIVERSITY OF MELBOURNE

In what sense does one doubt their understanding of reality when subject to gaslighting? I suggest that an answer to this question depends on the linguistic order at which a gaslighting exchange takes place. This marks a distinction between *first order* and *second order* gaslighting. The former occurs when there is disagreement over whether a shared concept applies to some aspect of the world, and where the use of words by a speaker is apt to cause a hearer to doubt her interpretive abilities *without* doubting the accuracy of her concept. The latter occurs when there is disagreement over which concept should be used in a context, and where the use of words by a speaker is apt to cause a hearer to doubt her interpretive abilities *in virtue of* doubting the accuracy of her concept. Many cases of second order gaslighting are *unintentional*: its occurrence often depends on contingent environmental facts. I end the paper by focusing on the distinctive epistemic injustices of second order gaslighting: (1) metalinguistic deprivation, (2) conceptual obscuration, and (3) perspectival subversion. I show how each reliably have sequelae in terms of psychological and practical control.

Type ‘gaslighting’ into any search engine and it’ll return a definition of the following kind: Gaslighting is a form of psychological manipulation, the effect of which induces doubt in a target’s understanding of reality. The intention behind this paper is to get clear on what this means. Of course, we all have a basic grip on the notion: Gaslighting is a pernicious medium of psychological and practical control. Psychological control in the sense that an offender of gaslighting can cause a target to form certain attitudes concerning her own reliability; practical control in the sense that an offender of gaslighting can motivate a target to perform actions that she otherwise wouldn’t. And this control can occur at different ontological levels. Political figures, such as Donald Trump, have been accused of gaslighting entire nations; but more commonly gaslighting is used to refer to abusive behavior that occurs in domestic settings. This distinguishes between *individual* and *collective* gaslighting.

All of this I will take for granted. What interests me is the following: *In what sense does gaslighting cause someone to doubt their understanding of reality?* Let me narrow the scope of this question. It is not my aim to offer an empirical thesis that explains what happens in someone’s head when they are gaslit. Nor is it my aim to explicate different theories of doubt and relate them to gaslighting. In this paper, I am concerned with the relationship between the concepts used to understand reality, and the doubt that gaslighting serves to induce. Using insights from the philosophy of language, my ultimate aim is to explain the epistemic effects on the subject of gaslighting, and how this leads to psychological and practical control. Specifically, I suggest that gaslighting can cause someone to doubt their understanding of reality in different ways depending on the linguistic order at which a conversational exchange takes place. To see this, consider a pair of conversations that I will return to throughout the paper:

(1) Woman: John brushed up against my butt, that’s sexual harassment.

Man: Sexual harassment? I’m sure it was an accident.

(2) Woman: John brushed up against my butt, that’s sexual harassment.

Man: That’s not sexual harassment. It’s so trivial.

Both (1) and (2) are cases of gaslighting. Yet, I argue that they are importantly different.

Conversation (1) is what I call *first order gaslighting*. It occurs when there is disagreement over whether a shared concept applies to some aspect of the world, and where the use of words by a speaker is apt to cause a hearer to doubt her interpretive abilities *without* doubting the accuracy of her concept. In (1), the man's use of words is apt to cause the woman to doubt that sexual harassment occurred, without her doubting that she has the right concept to pick it out.

Conversation (2) is what I call *second order gaslighting*. It occurs when there is disagreement over which concept should be used in a context, and where the use of words by a speaker is apt to cause a hearer to doubt her interpretive abilities *in virtue of* doubting the accuracy of her concept. In (2), the man's use of words is apt to cause the woman to doubt that sexual harassment occurred *because* she doubts that her concept can reliably pick it out.

Much attention has been paid to first order gaslighting, and rightly so. However, second order gaslighting has been almost unnoticed as a distinctive phenomenon.¹ This is a significant drawback in our pursuits in the project of social justice. First and second order gaslighting involve different wrongs, and require different solutions. In the final section of this paper, I bring to light three unique wrongs of second order gaslighting. I suggest that each wrong constitutes a distinctive form of *epistemic injustice*: (i) metalinguistic deprivation, (ii) conceptual obscuration, and (iii) perspectival subversion. I argue that each of these wrongs reliably have sequelae in terms of psychological and practical control.

One thing to note before I continue. Ordinary understanding of gaslighting is ambiguous across multiple readings. It can be understood as either having a success condition or not. To avoid confusion, I will not assume that gaslighting must be successful. Instead, I will understand gaslighting as behavior that is *apt* to cause another to doubt their interpretive abilities. This suggests that gaslighting is a probabilistic connection between behavior and doubt, thus accommodating the possibility that gaslighting behavior might not bring about doubt in a hearer.

1. What is Gaslighting?

Gaslighting is a contested notion.² To get clear on how I'll make sense of it, I want to start with ordinary understanding. A quick online search of gaslighting returns a common theme. This theme is best captured in Stephanie Sarkis's (2017) *Psychology Today* article where she writes:

Gaslighting is a tactic in which a person or entity, in order to gain more power, makes a victim question their reality (2017, 1).

Roughly, this definition suggests that gaslighting is particular behavior by one party that is causally responsible for making another come to doubt their understanding. This is vague, but it serves as a useful starting point.

The first thing to notice is that this definition makes a strong causal claim. According to Sarkis, gaslighting tactics *make* a victim question their reality. However, as suggested at the end of the last section, I want to accommodate cases where such tactics are unsuccessful. Thus, I take gaslighting to involve behavior that is *apt* to make a victim question their reality.

To fill out more of the details, we should ask: What kind of behavior is gaslighting? I will treat gaslighting as a socio-linguistic phenomenon. I accept that it could encompass much more than exchanging words, but it is not my aim to shed light on such possibilities. My interest is in forms of gaslighting that involve language use. Specifically, the use of language *apt for inducing certain attitudes in a hearer*. Which attitudes? Those that concern one's *interpretive abilities*. This includes: the ability to get facts right, and the ability to properly evaluate situations (Abramson 2014, 8). Thus, to say that gaslighting involves linguistic behavior that is apt to cause someone to 'question their understanding of reality' is to say that such behavior is apt to cause someone to doubt their interpretive abilities.

What this *doesn't* say is that gaslighting merely involves getting someone to believe that their understanding is mistaken. Gaslighting affects *higher-order* attitudes. For someone to be gaslit, they must form

negative attitudes about their attitude-forming mechanisms. In other words, gaslighting affects one's stance towards one's own reliability. Consider a familiar case of gaslighting. Imagine a wife who sees her husband kiss another woman. When she confronts the husband about this, he says that he was simply greeting the other woman and that she is 'crazy' for thinking otherwise. Here, the husband is not simply telling his wife that she is mistaken. He is saying something about the *means* by which she is mistaken. He is asserting that her ability to form appropriate attitudes about the situation is unreliable. For gaslighting to be successful, the wife must come to doubt her interpretive abilities, not merely her understanding of the event.

We can put this point another way: Gaslighting targets intellectual self-trust.³ And intellectual self-trust is comprised of *cognitive* and *affective* elements (Jones 2012). A person who is subject to gaslighting might come to form a *belief*, or make a *judgment*, about the unreliability of their interpretive abilities (e.g., forming the belief that they are 'crazy').⁴ However, cognitive elements do not account for all failures of self-trust. Self-trust can also be undermined by *affective attitudes*. After all, the doubt in one's own interpretive abilities may not be a response to reasons. Instead, it is due to certain emotions: anxiety, depression, hopelessness, lack of confidence, the need to come to a conclusion, etc. (2012, 6). To see this, consider the example above. When the wife confronts the husband after seeing him kiss another woman, she has sufficient reason to believe that her husband is cheating on her. However, let's say that owing to the power operating in the marriage, and the costliness of leaving, imagine that the wife still doubts what she saw. This needn't be because she believes that her perceptual and cognitive faculties have failed her. The niggling feeling of doubt is not a response to reasons. Instead, it is an affective stance toward her own interpretive abilities: The wife has her attention on the *bare* possibility that she might be mistaken.⁵

With this in mind, a first stab at defining gaslighting might run as follows:

Naïve Gaslighting: Gaslighting occurs when (i) a speaker uses words (ii) that are apt to cause a hearer to doubt her own interpretive abilities.

It is easy to produce counterexamples to *Naïve Gaslighting*. After all, it makes it seem that gaslighting is pervasive. I think that it is, but not as pervasive as this definition lets on. It allows for situations in which someone quite innocently uses words that are causally responsible for another coming to doubt their interpretive abilities. Imagine a couple that disagrees on what the total cost of dinner will be, and the one who is wrong comes to form the belief that they are cognitively defective because of this. It is far too much of a stretch to say that this falls within the proper bounds of gaslighting.⁶ Thus, *Naïve Gaslighting* doesn't specify the relevant conditions to be a definition of gaslighting.⁷ So, there is a residual question: In order for speech to count as gaslighting, what is the relevant relationship between the use of words by a speaker and the inducement of certain higher-order attitudes in a hearer?

A simple answer to this is that an offender of gaslighting has to *intend* to induce doubt in a hearer. We can call this *intentional gaslighting*:

Intentional Gaslighting: Gaslighting occurs when (i) a speaker uses words with (ii) the intention that a hearer comes to form (iii) negative attitudes toward her own interpretive abilities.⁸

All that's important to notice about *Intentional Gaslighting* is that it requires that one intend to gaslight. Take the case above where the wife sees her husband kiss another woman. According to this definition, for the husband to count as gaslighting his wife, he must intend for her to doubt her interpretive abilities. This is perhaps what we have in mind when we think of paradigm cases of gaslighting.

One might take exception to *Intentional Gaslighting* on the grounds that not all intuitive cases of gaslighting involve an intention of this sort. Think of a situation in which a man brushes up against a woman's butt in the office, and the woman reports this to a colleague who responds, 'I'm sure it was innocent, John isn't the kind of guy to act inappropriately at work'. As a result, the woman comes to doubt her ability to recognize sexual harassment. In this case, the colleague may not intend for the woman to doubt her interpretive abilities, even though he expressed doubt in her testimony. Nevertheless, this looks like a cut and dry case of

gaslighting.⁹ The lesson seems to be that not all cases of gaslighting are intentional, though perhaps a great many of them are.

If we can't appeal to intentions to account for all cases of gaslighting, it seems that we have to look beyond psychological facts. A useful starting point is to recognize that the world is unjust. Certain groups of people are privileged, others are subordinated. One sense in which the world is unjust is that certain people are excluded from, or lack full participation in, the economy of epistemic relations; the social network in which we exercise our capacity as knowers. Exclusion from (full) participation in such relations is called *epistemic injustice* (Fricker 2007). This is a wrong that occurs when someone is undermined in their capacity as an epistemic subject. There are many ways that one might suffer from epistemic injustice. One might receive an unfair credibility assessment, and thus suffer from *testimonial injustice* (Fricker 2007, chap. 1); one might be unable to render intelligible experience, and thus suffer from *hermeneutical injustice* (2007, chap. 7); and/or one might be subject to ill-fitting interpretations, and thus suffer from *contributory injustice* (Dotson 2012) or *wilful hermeneutical ignorance* (Pohlhaus Jr. 2011), etc. Whatever the case, the important point to focus on is that identity-prejudice plays a substantive role in determining who is afflicted most by epistemic exclusion. Such injustice exists in our structures, and dictates the character of our epistemic lives.

If gaslighting affects one's attitudes toward their interpretive abilities, then it is easy to see why epistemic injustice is worth bringing up. Given that certain groups are excluded from, or lack full participation in, the economy of epistemic relations, then this serves as partial explanation for why members of such groups might be inclined to downgrade their self-trust in the face of doubt. After all, they have never really been treated as full epistemic subjects. Because of this, we might think that an unintentional story of gaslighting should be sensitive to such facts. Thus, we can define unintentional gaslighting as follows:

Unintentional Gaslighting: Gaslighting occurs when, (i) a speaker uses words *without* the intention that a hearer come to form negative attitudes toward her own interpretive abilities, but (ii) the use of such words is apt to cause the hearer to form such attitudes (iii) owing to the hearer being subject to systematic epistemic injustice that has disposed her to do so.

Unintentional Gaslighting allows us to distinguish between cases in which it seems that gaslighting has occurred and cases in which it seems that it hasn't, but where the only relevant difference is that one person is subject to systematic epistemic injustice, and the other is not.

For example, consider the case above where the woman testifies to being deliberately brushed up against by a man in the office, and her colleague responds, 'John isn't the kind of guy to act inappropriately at work'. In a society that systematically fails to treat the testimony of women seriously when it comes to allegations of sexual harassment (and sexual assault), gaslighting is more of a possibility. Why? The tendency of members of society to disbelieve women when such allegations are made disposes women to doubt their own experience involving sexual harassment. And this doubt is not just about *particular* cases of sexual harassment, but the ability to identify it *generally*. Women are pushed to doubt their interpretive abilities (in a domain, or across domains). They are told: 'nice guys don't sexually harass people'; 'you asked for it by wearing provocative clothing'; 'just accept that's what men are like', etc.

Compare this case with an otherwise identical one, but where it is a man who gets brushed up against and is told 'John isn't the kind of guy to act inappropriately at work'. On the unintentional reading, this doesn't count as gaslighting. I think that this gets the right result. This is not to say that sexual harassment didn't occur. It is only to say that when the man's allegation of sexual harassment is dismissed, it isn't a case of gaslighting. Why? Men enjoy a great deal of credibility. They are not disposed to doubt themselves in the face of conflicting understanding. And this is due to the fact that men don't suffer from pernicious stereotypes concerning their ability to get facts right or properly evaluate situations (e.g., men are not associated with being too emotional or sensitive).¹⁰ The important difference between the two cases is this: Women are subject to systematic credibility deflations and hermeneutical deficiencies, whereas men are not. Unintentional gaslighting is sensitive to such facts.

Given the foregoing considerations, a definition of gaslighting should accommodate cases that are intentional and unintentional. That is, intentional and unintentional gaslighting are complements – both are relevant to gaslighting. Thus, I suggest that following *disjunctive* definition:

Gaslighting: Gaslighting occurs when (i) a speaker uses words and either (ii) the speaker intends for the use of such words to cause a hearer to form (iii) negative attitudes toward her own interpretive abilities, or (iv) a speaker uses words without such an intention, but (v) the use of words is apt to cause the hearer to doubt her interpretive abilities (vi) owing to the hearer being subject to systematic epistemic injustice that has disposed her to do so.¹¹

With this definition in mind, I want to discuss one further aspect of gaslighting that is paramount. Bringing to light the nature of unintentional gaslighting exposes its *temporal* aspect. We might tend to think of gaslighting as one-off events, but it is perhaps better thought of as something that occurs within a system of oppressive patterned behavior *over time*. We can explain why it is that people are more disposed to doubt themselves in the face of certain utterances by looking at the social environment in which they are situated, and the information that is available in such environments that enables coordination between members.

Social environments consist of *background patterns*. Such patterns are the regular and coordinated behavior of individuals that comprise a network of social relations. This coordinated behavior is a product of the culturally available information that agents draw on that frames expectation, enabling (quasi) rule-governed social interaction. This includes drawing on tropes, narratives, social meanings, schemas, roles, and the like.¹² Cultural tools of this sort allow social agents to render intelligible experience, and facilitates intentional engagement with their surroundings. In unjust social environments, the extant patterns of interaction unfairly privilege some and subordinate others.

To coordinate at a given time, social agents draw on certain cultural tools that narrow the range of expected behavior. The salience of such tools can be made explicit by the interacting parties, or the context in which the parties are situated. For example: In an oppressive patriarchal society, pernicious patterns of interaction are activated when the schema ‘woman’ is salient in certain domains. A pertinent case is the systematic bias in assessment of intellectual contributions by women in philosophy. Here, it is the context of

philosophy that makes gender a salient cultural tool to use when evaluating intellectual credibility. Other contexts include male dominated workplaces and domestic settings.

Much of gaslighting, intentional and unintentional, depends on the salience of cultural tools. This is because such tools determine the range of behavior fitting for the interaction. Consider the example above where the woman testifies to being sexually harassed, and the colleague responds by saying 'John isn't the kind of guy to act inappropriately at work'. Despite the word 'woman' not being explicitly mentioned in the interaction, an associated cultural tool is made salient given the nature of the conversation, and the context in which the conversation takes place. After all, a *woman* has made an accusation against a *man* in a male dominated workplace situated in a patriarchal social environment. Due to women's association with 'hysteria', 'over-sensitivity', and 'paranoia' in such an environment, what is fitting for the woman to do in the face of doubt is to downgrade her self-trust. In other words, women are disposed to downgrade their self-trust given persistent challenges to their epistemic reliability. Thus, the success of gaslighting owes much to the salience of certain cultural tools (e.g., identity-prejudicial stereotypes) that govern social interaction.

In sum, gaslighting is not something that usually happens in one-off, isolated instances. Instead, it is often a function of the social environment in which one is situated that has an extended history of unjust treatment of (some of) its members. The success of gaslighting is dependent on the background patterns and culturally available information that gaslighters are able to tap into.

2. Gaslighting, First and Second Order

Up to this point, I have only discussed how I'll make sense of gaslighting: It is a socio-linguistic phenomenon that involves the use of words apt for causing a hearer to doubt her interpretive abilities. This can be intentional or unintentional, and often requires making salient pernicious cultural tools (e.g., stereotypes). I am now in a position to offer an answer to the primary research question: *In what sense does gaslighting cause someone to doubt their understanding of reality?* Given the foregoing, we can rephrase the question: *In what sense does gaslighting cause someone to doubt their interpretive abilities?* What I will suggest is that there are two broad

ways of answering this question. Each answer depends on the linguistic order at which a gaslighting exchange takes place. *First order* gaslighting occurs when a gaslighter and the subject of gaslighting¹³ aim to coordinate on a presupposed shared concept, but take themselves to disagree as to whether the concept applies to some aspect of the world. *Second order* gaslighting occurs when a gaslighter and the subject of gaslighting *do not* aim to coordinate on a presupposed shared concept and, instead, disagree as to which concept should be used in a context.¹⁴ Second order gaslighting has been overlooked as a distinctive form of gaslighting in the philosophical literature. This may be due to the fact that very many instances of second order gaslighting seem unintentional – it is not the sort of thing that comes to mind when we think about clear cases of gaslighting. Nevertheless, it is a form of injustice that should be made explicit.

We can see gaslighting as occurring at different linguistic orders by comparing the pair of conversations given at the beginning of the paper:

- (1) Woman: John brushed up against my butt, that's sexual harassment.
 Man: Sexual harassment? I'm sure it was an accident.
- (2) Woman: John brushed up against my butt, that's sexual harassment.
 Man: That's not sexual harassment. It's so trivial.

Both conversations represent a fairly typical gaslighting exchange: Someone offers testimony, and this testimony is rejected out hand.¹⁵ So, what we can see is *disagreement*. By 'disagreement', I don't mean a rational dispute between epistemic peers – there is no reasons-giving dialogue. Instead, I mean only to say that there is a difference in what the woman and man take to be the nature of the event that took place. In both (1) and (2), the woman thinks that John brushing up against her is sexual harassment, and the man does not.

What we can also see is that the *grounds* on which the man rejects the woman's understanding of the event as sexual harassment differs between (1) and (2). In (1), the man objects to the woman on the grounds that it must have been an accident. Perhaps this is due to his affection for John, or his general skepticism of

workplace harassment. In (2), the man objects to the woman on the grounds that she doesn't understand what sexual harassment is. The man has a different idea of the range of things that count as sexual harassment. In short: (2) involves a presupposed *conceptual difference*; (1) does not.

A short interlude: What do I mean by 'concept'? Providing a story of what concepts are is notoriously difficult, so I won't attempt it here.¹⁶ All that I need to say for my purposes is that concepts have extensions; the range of things that fall under a concept. Now, this can be interpreted in, at least, a couple of ways. We can say that 'concept' is just a stand in for word meanings, and word meanings are simply the extensions (and intensions) of *lexical items* (Cappelen 2018). Or we can say that concepts are more than the extensions of lexical items, but instead *cognitive structures* that have extensions in addition to a rich body of associated (non-reference fixing) information that guides social interaction (e.g., Elisabeth Camp's (2015) notion of *characterizations*). I am inclined to accept the latter, however I don't wish to get involved in this debate.¹⁷ Whatever the right story, concepts either are or have extensions. I will remain agnostic about how such extensions are determined.

Now that we have some clue as to what concepts are, I want to discuss different kinds of disagreement involving concepts. This will help us get clearer on the difference between (1) and (2).

2.1. Orders of Usage and Disagreement

Speakers use words to convey information about the world. When I say 'Fil is tall', I am describing how things are: There is someone named Fil, and he has the property of being tall. However, sometimes words are used by speakers to convey information about the appropriate or correct usage of those words in a context. In such cases, when I say 'Fil is tall' I am pragmatically advocating for how to use 'tall' in the context. Call the former way of using words *first order usage*, the latter *second order usage*.¹⁸

Disagreements can occur at both orders of usage. I can say 'Fil is tall', and someone can disagree with me by saying 'Fil is not tall'. What ensues is a disagreement that centers around whether or not a threshold of tallness has been met. We can call this *first order disagreement*. First order disagreements concern what the

world is like, where speakers are committed to coordinating on a presupposed shared concept – e.g., speakers agree on a threshold that must be met for being tall, but disagree as to whether someone has met this threshold. However, disagreements can also occur over how ‘tall’ *should* be used in a context. Speakers can disagree as to which threshold must be met in order for someone to be tall. David Plunkett and Tim Sundell (2013) call this *metalinguistic negotiation*. We can call it *second order disagreement*.

Second order disagreement is fundamentally about conceptual choice: Out of a competing range of concepts, which is the one that should be expressed by a word in a context? (2013, 3). Settling this question is a matter of discerning which concept we should employ for the purposes at hand, not (directly) about the truth of the literal content of an utterance. For example, consider an argument between friends:

A: This martini is delicious!

B: That’s not a martini, it has vodka in it.

We can think of this exchange as first order disagreement. It concerns whether or not the drink is, in fact, a martini. However, we can also see it as a disagreement that takes place at the second order. It is a disagreement over which concept of martini, out of a competing range, should be expressed by ‘martini’. Each friend advocates for their usage, taking into consideration things such as authenticity and common usage.

One might ask: How can tell when a disagreement is second order? Amie Thomasson (2017) lists three diagnostic markers: (i) disputants agree on the facts, and the introduction of more facts will not resolve the disagreement; (ii) the disagreement cannot be settled by appealing to prior or actual usage of the term; (iii) the disagreement persists even when the disputants realize they are using terms differently. In the case above, we can see how all three criteria might be met. On (i), we can imagine that A and B agree on all of the bare facts – they accept that the drink in question contains gin and vodka (as opposed to gin and vermouth). On (ii), we can imagine that A makes an appeal to common usage and B makes an appeal to authenticity, but such

appeals do not settle the disagreement. On (iii), we can imagine that A and B come to realize that they are using the term ‘martini’ differently, but nevertheless the disagreement persists.

2.2. First order Gaslighting

With this distinction in mind, let’s consider (1) once more:

- (1) Woman: John brushed up against my butt, that’s sexual harassment.
Man: Sexual harassment? I’m sure it was an accident.

Here, it appears that the woman and man are aiming to coordinate on a shared concept of *sexual harassment*. This is evidenced by the fact that the man contests whether the act was *deliberate*, and does not contest whether the deliberate act *counts* as sexual harassment. That is, the woman and the man both agree that *if* the act of brushing up against the woman’s butt were deliberate, it would have been sexual harassment. There is no presupposed conceptual difference. Thus, (1) is first order disagreement.¹⁹

How does first order disagreement turn into first order gaslighting? It depends on whether the man intends to gaslight. On the intentional reading, the man recognizes that the woman has a competing interpretation of the event, and with this recognition, he intends to cause the woman to doubt her ability to make reliable judgments about whether an event falls within the range of their shared concept of *sexual harassment*.

On the unintentional reading, the man recognizes that the woman has a competing interpretation of the event, however this recognition does not undergird an explicit intention of getting the woman to doubt her interpretive abilities. Instead, the man simply takes the woman to be wrong about the bare facts of the situation. He dismisses her testimony. However, as suggested before, the context of this dismissal matters when it comes to gaslighting: Given the content of the conversation, the salience of certain cultural tools causes the woman downgrade her self-trust in the face of doubt. That is, the woman is part of a society in

which allegations of sexual harassment are regularly dismissed, and this disposes her to doubt her interpretive abilities when her testimony is met with disbelief. Despite the man not intending to gaslight, he counts as gaslighting owing to contingent environmental facts.

With this in mind, we can define first order gaslighting as:

First order gaslighting occurs only if (i) a gaslighter and the subject of gaslighting aim to coordinate on a presupposed shared concept, but where (ii) the gaslighter's use of words is apt to cause the subject of gaslighting to doubt her interpretive abilities, and (iii) this is *not* due to the subject of gaslighting doubting the accuracy of her concept.²⁰

So, in what sense does someone doubt their interpretive abilities when subject to first order gaslighting? According to this definition: the subject to gaslighting doubts their ability to get facts right, or properly evaluate situations, but this doubt is not a result of mistrusting the accuracy of the concept used to pick things out in the world.

2.3. *Second order Gaslighting*

Consider (2) once more:

(2) Woman: John brushed up against my butt, that's sexual harassment.²¹

Man: That's not sexual harassment. It's so trivial.

Here, it appears that the woman and man are not aiming to coordinate on a presupposed shared concept of *sexual harassment*. They appear to have different concepts. This isn't to say that the woman and man are not aiming to coordinate on *some* concept. In a conversation where speaker's recognize that there is conceptual difference, it stands to reason that they will do as much as possible to ensure that there is conceptual

similarity.²² After all, without establishing conceptual similarity, conversation is difficult to advance – either speakers will be talking past each other, or at least one speaker will not understand what is being said. What I am claiming is simply that the man and woman did not enter the conversation with the same concept. There is presupposed conceptual difference. Nevertheless, there is also an assumed background aim of coordinating on a concept that they can both agree on.

Not only are the man and woman failing to coordinate on a presupposed shared concept, they are also disputing which concept should be expressed by ‘sexual harassment’. Why believe this? We can appeal to Thomasson’s diagnostic markers. On (i), the man and woman both accept the fact that John brushed up against the woman. Unlike (1), the man does not contest whether the act of brushing up against the woman’s butt was deliberate. He accepts that it is deliberate, but still fails to count as sexual harassment. Thus, the disagreement is over how this deliberate act ought to be conceptualized. On (ii), we can imagine the exchange continuing where appeals are made to previous or actual linguistic usage, but where such appeals do not settle the disagreement. The man might say to the woman, ‘sexual harassment is about actually groping women, what you experienced was just harmless flirting’, and the woman might respond, ‘that’s not how people use the term anymore’. On (iii), we can imagine that the man and woman realize that they are using the term ‘sexual harassment’ in different ways, but nevertheless they continue to disagree. The woman thinks the term should be used to include cases in which one is deliberately brushed up against; the man thinks that such cases are far too trivial to count. From this, we can see is that (2) meets the diagnostic criteria for second order disagreement.

Now, how does second order disagreement turn into second order gaslighting? We can tell roughly the same story as we did about first order gaslighting. The primary difference is *why* the subject of gaslighting comes to doubt her interpretive abilities. In first order gaslighting, the gaslighter’s use of words is apt to cause the subject of gaslighting to doubt her ability to recognize sexually harassment *without* her questioning the accuracy of her concept of *sexual harassment*. That is, the subject of gaslighting comes to doubt her ability to recognize sexual harassment, but not because she believes that she has the wrong concept to identify it. Contrast this to second order gaslighting. The gaslighter’s use of words is apt to cause the subject of

gaslighting to form negative attitudes about her ability to recognize sexual harassment *in virtue of doubting the accuracy of her concept*. In other words, the subject of gaslighting comes to doubt her ability to judge whether sexual harassment has occurred *because she doubts that she has the right concept to pick it out*.²³

With this in mind, we can define second order gaslighting as:

Second order gaslighting occurs only if, (i) a gaslighter and the subject of gaslighting *do not* aim to coordinate on a presupposed shared concept, and where (ii) the gaslighter's use of words is apt to cause the subject of gaslighting to doubt her interpretive abilities, (iii) *owing* to the subject of gaslighting doubting the accuracy of her concept.

We can ask: In what sense does someone doubt their interpretive abilities when subject to second order gaslighting? The answer: The subject of gaslighting doubts her ability to get facts right, and properly evaluate situations, in virtue of doubting the accuracy of the concept used to interpret some aspect of the world. It strikes me that much of second order gaslighting is unintentional. Of course, this claim needs empirical backing. However, it seems that the second order gaslighter does not explicitly intend to cause the subject of gaslighting to doubt her interpretive abilities through challenging concepts. To accept otherwise would make gaslighting far too cognitively demanding, meaning that we would miss out on explaining many of its occurrences. Despite this, we might say that when a gaslighter challenges a concept, the challenge *implicitly* carries the assumption that the concept in question is wrong, and stands in need of improvement. However, what's important to note is that this does not involve an explicit intention to cause another to doubt their interpretive abilities.²⁴

I want to finish this section by raising a concern: Why would second order disagreement cause someone epistemic distress given that the dispute could be resolved by simply introducing a new term to express one of the concepts (e.g., 'sexual pestering' as opposed to 'sexual harassment')? In some cases, this might be the best strategy. It may be too difficult to convince someone to take up a new concept to be expressed by a word that has an established word-concept pairing. However, for a good deal of cases, we have an investment in words

owing to their normative valence and function in our social practices – what Herman Cappelen (2018) calls ‘lexical effects’. In particular, words like ‘sexual harassment’ do serious normative and explanatory work. People already have the term ‘sexual harassment’, and it carries cognitive and affective associations (e.g., a serious moral wrong) that enables inferences to be made, informing our decisions-making practices (e.g., staying away from a perpetrator) and normative judgements (e.g., the perpetrator is the appropriate target of blame, resentment, etc.). And this investment in words can also partially explain why someone might be motivated to (intentionally) gaslight at the second order. A gaslighter might recognize the normative significance of the word under dispute, and, owing to this, will want to either keep the word or avoid it depending on his purposes.

3. Wrongs of Second order Gaslighting

Why bother distinguishing between first and second order gaslighting? First off, it is important to make as much sense of injustice as we can. When we think of gaslighting, it strikes me that we tend to see it as occurring only at the first order; so revealing the nature of second order gaslighting is useful insofar as we can identify its occurrence. But perhaps more importantly, first and second order gaslighting involve different wrongs, and require different solutions. The wrongs of first order gaslighting have been discussed in much detail in the philosophical literature (Abramson 2014; McKinnon 2017, forthcoming; Stark 2019). Thus, I will spend time explicating distinctive wrongs of second order gaslighting. I suggest that there are (at least) three: (1) *metalinguistic deprivation*, (2) *conceptual obscuration*, (3) *perspectival subversion*. Each of which constitute a form of *epistemic injustice*. I show how each of the injustices reliably have sequelae in terms of psychological and practical control.

3.1. *Metalinguistic Deprivation*

In cases of second order gaslighting, the gaslighter challenges the accuracy of the concept that the subject of gaslighting has used in her interpretation of an event. This challenge is apt to cause the subject of gaslighting to doubt the accuracy of her concept. Owing to this doubt (confusion, loss of self-trust, etc.), the subject of gaslighting becomes unable to fully participate in discussion about which concept should be expressed by shared words (i.e., resolving metalinguistic negotiation). The subject of gaslighting does not feel that their concept is best for the context – the accuracy of the concept is doubted, after all. This, I maintain, constitutes an *epistemic injustice* (Fricker 2007). Why?

Epistemic injustice was originally characterized as ‘a wrong done to someone in their capacity as a *knower*’ (2007, 1, my emphasis). Yet, we needn’t articulate epistemic injustice in terms of knowledge. Instead, we can think of it as a wrong done to someone in their capacity as an *epistemic subject* (Fricker 2013; Gerken 2019). There are different ways that we can think of what it means to be an epistemic subject. However, the primary reason behind using the language of ‘subject’ over ‘knower’ is to express that there are aspects of our epistemic lives that involve representing and interpreting the world that falls short of knowledge.

We can think of an epistemic subject as someone who *interprets* the world on the basis of the concepts to which they have access, whether or not this interpretation constitutes knowledge. This is evident in Fricker’s discussion of *hermeneutical injustice* (Fricker 2007, 148). Here, it seems that there is something essential that epistemic subjects do: They interpret the world in line with the concepts at their disposal. Being denied access to (certain) concepts, under certain conditions, that would enable one to render intelligible experience constitutes a distinctively epistemic wrong.

Additionally, an epistemic subject is someone who not merely passively interprets the world, but actively contributes to the pool of epistemic resources shared by members of a community of thinkers and speakers. This thought is present in the work on pernicious ignorance (Pohlhaus Jr., 2011; Dotson 2012). The idea is that dominantly situated epistemic subjects have failed to recognize the knowledge that marginalized folk have to offer, and in virtue of this fail to understand marginalized experience. This constitutes an epistemic injustice. The lesson is that the decision as to which epistemic resources a diverse community of thinkers and speakers ought to possess isn’t something that should fall in the hands of a privileged few.

Such epistemic resources aren't just shared within global communities. They are also present in hyper-local contexts such as conversation. As we have seen, speakers negotiate the concepts that will be expressed by shared words. In such contexts, *individuals* can be prevented from contributing understanding in much the same way that *groups* are prevented in global contexts. From this, we can say that, like groups, individuals are undermined in their capacity as an epistemic subject when this occurs. Thus, an epistemic subject is, *inter alia*, someone *who is entitled to have their conceptual understanding given consideration*²⁵ *in the decision as to which concept will be expressed by a shared term in a context*. When one is prevented from full participation in concept-determining conversation, this constitutes a denial of someone's capacity as an epistemic subject. Call it *metalinguistic deprivation*:

Metalinguistic deprivation is an epistemic injustice that occurs when one is prevented, or restricted, from contributing to the processes involved in determining the concept that will be expressed by a word, or set of words, in a context (especially a context in which this determination matters).

3.1.1. Discriminatory Metalinguistic Deprivation

Metalinguistic deprivation comes close to Fricker's notion of *hermeneutical marginalization* (2007, 152). This occurs when members of a marginalized group are prevented from *hermeneutical participation* with respect to some significant area of social experience. Being prevented from hermeneutical participation is the idea that one is denied access to equal production of epistemic resources, such as concepts, to be expressed by words across social space (Fricker 2013, 1319).

We can see a similarity. Both hermeneutical marginalization and metalinguistic deprivation involve preventing someone from participating in the processes that determine the available epistemic resources. However, there are differences. First, hermeneutical participation occurs at the level of groups, whereas metalinguistic deprivation operates at the individual level. Second, hermeneutical marginalization is *structural* (Fricker 2007, 155). It is something that marginalized groups suffer from, and from which privileged groups

benefit. Metalinguistic deprivation does not require this. But, this is not to say that metalinguistic deprivation cannot be structural in some sense, while agential in another. Metalinguistic deprivation becomes a *discriminatory epistemic injustice*²⁶ when one is prevented, or restricted, from the processes that give rise to the concepts to be used in conversation *owing to prejudicial stereotypes that are made salient in the context*. Call this *discriminatory metalinguistic deprivation*. In cases of gaslighting, this can happen in two broad ways.

On the intentional story, a gaslighter might intend to induce doubt in the subject of gaslighting by saying things like, ‘You women are hysterical’, ‘You’re probably on your rags’, ‘You’re so emotional’, and the like. In such cases, the gaslighter is attempting to make salient prejudicial stereotypes (i.e., cultural tools that narrow the range of expected behavior) associated with a particular social category with the aim that the subject of gaslighting *come to believe that such stereotypes accurately represent who she is*. If the subject of gaslighting comes to form such a belief, this sets up the conditions for successful metalinguistic deprivation. The subject of gaslighting will feel that she cannot properly contribute to concept-determining conversation in virtue of doubting her interpretive abilities, which is a consequence of thinking that a pernicious stereotype aptly applies to herself.

On the unintentional story, the gaslighter does not intend to induce doubt in the subject of gaslighting, but nonetheless brings to salience a prejudicial stereotype that causes the subject of gaslighting to doubt herself. This can be as simple as dismissing the testimony of women in contexts where they tend to be disbelieved (e.g., ‘that’s not sexual harassment’ activates certain cultural tools). Again, the subject of gaslighting becomes restricted in her ability to contribute to concept-determining conversation in virtue of thinking that a pernicious stereotype aptly applies to herself.

3.1.2. *Metalinguistic Deprivation and Pernicious Ignorance*

Another form of epistemic injustice close to (discriminatory) metalinguistic deprivation is *wilful hermeneutical ignorance* (Pohlhaus Jr. 2011). This occurs when dominant knowers use ‘epistemic resources that do not allow for the intelligibility of what [marginalized people have] to say’ (2011, 725), while at the same time

‘refusing to learn the epistemic resources developed from marginalized situatedness’ (2011, 722). This is closely related to Kristie Dotson’s (2012) notion of *contributory injustice*.

Wilful hermeneutical ignorance, and contributory injustice, are present in a great deal of second order gaslighting. After all, both involve a dominant knower imposing an ill-fitting epistemic resource onto marginalized experience. However, the wrongs of wilful hermeneutical ignorance and contributory injustice are distinct from the wrong of metalinguistic deprivation. The wrong of metalinguistic deprivation is to be found in someone being *prevented from, or restricted in, concept-determining conversation*. The wrong of wilful hermeneutical ignorance appears to occur *after* this process. The latter occurs after it has been determined that, say, ‘sexual harassment’ will express the dominant knower’s preferred concept. Thus, we might say, at least in some cases, metalinguistic deprivation *gives rise* to wilful hermeneutical ignorance. This is because a dominant knower can have his preferred concept accepted as being expressed by a shared word, and used to interpret experience, by depriving his interlocutor from contributing conceptual understanding.

3.1.3. *Psychological and Practical Control*

Apart from being an epistemic injustice, (discriminatory) metalinguistic deprivation can be a source of psychological and practical control. Broadly, I take ‘control’ to refer to the ability to influence someone contrary to their interests.²⁷ This can be innocent. I am interested in cases in which control, in this sense, is *unfair*. And I take it that when one is influenced contrary to their interests *in virtue of* suffering from epistemic injustice, that this constitutes unfair control.

We can say that *S* has *psychological control* over *T* when *S* is able to influence *T*’s psychological states, such as how *T* feels and how *T* sees the world; and *S* has *practical control* over *T* when *S* is able to influence the way that *T* engages with their social environment. In cases of (second order) gaslighting, psychological and practical control is unfair. Let’s discuss the former in more detail first.

Metalinguistic deprivation can be a source of psychological control insofar as the gaslighter can influence the beliefs that the subject of gaslighting has, and motivate certain affective attitudes. In the example that

we've been exploring, the woman acquiesces to the concept of sexual harassment endorsed by her gaslighter. Because of this, she must come to accept certain things. She must accept John did not sexually harass her. Further, she must take on the attitude that she was wrong to accuse John of sexual harassment.²⁸ Accepting such things is apt for inducing certain affective states. The woman might feel anxious or insecure around her John, second-guess herself when brushed up against, feel unwelcome in certain spaces, and take herself to be 'bad' for ever thinking she was harassed, etc. The subject of gaslighting is under the psychological control of her gaslighter in virtue of being excluded from concept-determining conversation that causes her to take on attitudes, both cognitive and affective, that are contrary to her interest in being safe.

Metalinguistic deprivation is also a source of practical control. This is owes to the practical consequences of accepting a concept as apt or appropriate for a context. Consider the case earlier in the paper where a wife catches her husband kissing another woman. Imagine that the husband denies that kissing counts as 'cheating'. Let's say that the husband suggests this because he recognizes that if he accepts his wife's preferred concept, he will be in the wrong. Now, suppose that the wife, owing to metalinguistic deprivation, accepts the husband's concept of cheating. A number of things practically follow from this. One consequence is that the wife must no longer interact with her husband as a cheater - she no longer has grounds for accusing him of wrongdoing, leaving him, asking for an apology, being emotionally distant, demanding that he change his ways, etc. And, because she may feel bad for accusing of her husband of cheating, this might mean that she changes the way she relates to him, such as being overly apologetic, accommodating, cautious not to anger or upset him. Moreover, the husband may continue to kiss other women, even in front of his wife, yet the wife feels that she isn't licensed to call him out. Other consequences abound. Through second order gaslighting, the gaslighter can push the subject of gaslighting to accept a concept, which determines the fittingness of certain actions for the context, thereby influencing the subject of gaslighting to relate to her social environment contrary to her interest in being respected by her partner. Gaslighting can motivate a target to perform actions that she otherwise wouldn't.

3.2. *Conceptual Obscuration*

A further wrong of second order gaslighting occurs when the subject of gaslighting forms the attitude that she has an inaccurate concept and, owing to this, *adopts the preferred concept of her gaslighter*. For example, the woman who is brushed up against may take on board her gaslighter's words that such an act is far too trivial to count as sexual harassment, and thus comes to accept that it does not constitute sexual harassment. Clearly, this has contextual implications. 'Sexual harassment' will express the gaslighter's preferred concept of *sexual harassment*. But it also has cross-contextual implications. The woman will become disposed to token her gaslighter's preferred concept in any context that she deems relevantly similar (i.e., any context in which she, or others, experience particular unwanted sexual advances). And given that the gaslighter's preferred concept of *sexual harassment* is inaccurate, this means that the woman will not recognize sexual harassment when it occurs. Because of this, the woman *loses important knowledge about the world*. Knowledge that is in her best interest to have. Call this *conceptual obscuration*.

Conceptual obscuration undermines someone in their capacity as an epistemic subject. It constitutes an epistemic injustice. How? The first thing to notice is that it involves someone *diminishing the knowledge of another by obscuring their epistemic resource*. However, this does not yet constitute an epistemic injustice – obscuring one's epistemic resource might be quite innocent. For example, in a friendly conversation one might be told that couches are chairs, therefore obscuring their concept of *chair*. It is far from intuitive to say that this constitutes an injustice. To locate the injustice, we must look at *the means by which an epistemic resource is obscured*. And the means by which obscuring an epistemic resource constitutes an epistemic injustice is through *metalinguistic deprivation*. It is one thing to be wronged by being denied full participation in concept-determining conversation, it is another thing to be wronged by being denied this *and* lose knowledge in the process (in virtue of adopting an inaccurate concept endorsed by an interlocutor). The wrong is located not simply in denying someone full participation in concept-determining conversation, but in the fact that this denial is causally responsible for someone losing knowledge about the world. Thus, we can define conceptual obscuration as:

Conceptual obscuration is an epistemic injustice that occurs when *S* possesses an accurate concept *c*, but in virtue of suffering from metalinguistic deprivation, *S* comes to replace *c* with an nearby inaccurate concept *c**.²⁹

This is distinct from other forms of epistemic injustice. It is not hermeneutical injustice insofar as there is no lacuna in the collective hermeneutical resource. It is not *just* wilful hermeneutical ignorance or contributory injustice insofar as it isn't simply about imposing an ill-fitting concept onto marginalized experience. Instead, it is about one losing knowledge they once had, and the means by which this occurs.³⁰

Losing knowledge entails that there are things to which the subject of gaslighting is no longer licensed. Like metalinguistic deprivation, this sets up the conditions for psychological and practical control. Consider again the example of the cheating husband. In this case, the wife accepts her husband's concept of cheating. When this occurs, she loses knowledge of what cheating is; she can no longer reliably pick it out. This means she is no longer licensed to believe her husband is cheating, or accuse him of doing so – her newly acquired concept doesn't take kissing to count. This constitutes psychological control insofar as the gaslighter is able to influence what the subject of gaslighting is able to *know* or *rationaly believe*. Moreover, this translates into practical control. The wife's practical reasoning is affected insofar as she cannot, say, rationally choose to leave her husband, or confront him for his actions, given that he hasn't 'cheated' under the relevant definition.³¹

3.3. *Perspectival Subversion*

The final wrong of second order gaslighting that I will discuss is *perspectival subversion*. Again, this constitutes a form of epistemic injustice. It is not about being denied full participation in concept-determining conversation, or coming to adopt inaccurate concepts. Instead, it concerns the impairment of one's ability to *independently* interpret situations. In other words, it is a wrong that occurs when a subject of gaslighting is exposed to persistent challenges to her concepts over time, and because of this reaches a point where she

forms the attitude (i.e., belief, anxiety) that her interpretive abilities are so unreliable that she must defer to her gaslighter's conceptual judgment (i.e., the right or appropriate concept to token). This comes in degrees. At the extreme end, it involves a total loss of self-trust.³² Such cases are perhaps rare. More commonly, a subject of gaslighting will lose self-trust in particular domains – such as whether an event counts as sexual harassment. In any case, this captures a unique and pernicious wrong of gaslighting:

Perspectival subversion is an epistemic injustice that occurs when a subject of gaslighting is the target of persistent conceptual challenges over time such that they come to doubt their ability to make conceptual judgments (in a domain, or across domains), and so they defer to the conceptual judgment of their gaslighter.³³

Why 'perspectival'? It concerns how gaslighting affects one's *disposition* to construct interpretations of events (Camp 2018). Contrast this with the wrongs discussed above. Metalinguistic deprivation and conceptual obscuration involve the disposition to token a concept. In the former, it is the deployment of a concept that is challenged; in the latter, it is the deployment of a concept that is challenged, and a new concept is adopted to be used. Perspectival subversion, on the other hand, involves a disposition *not* to deploy concepts. It involves a subject of gaslighting refraining from making conceptual judgments because they take themselves to be unreliable (in a domain, or across domains) after being exposed to persistent challenges to their conceptual understanding. For example, consider the case above where the wife sees her husband kiss another woman. Let's say that in this marriage, the wife regularly sees her husband kiss other women, and when she confronts him about it each time, he consistently and repeatedly tells her that she doesn't really understand what 'cheating' is – he tells her that kissing doesn't count. He aggressively accuses her of being crazy when she brings it up, and says that she needs to stop being so sensitive. After a while of this treatment, the wife doubts the accuracy of her concept of cheating. However, she doesn't replace it with the one endorsed by her husband. Instead, she is so confused about what things count as cheating, and what things

don't, that she decides her interpretive abilities aren't good enough to reliably pick it out, or to discern the correct concept.³⁴ The wife becomes disposed not to make conceptual judgments about cheating.

Making a conceptual judgment is often necessary for practical reasons. For example, interpreting a situation as cheating allows one to act in response to the situation *as* cheating. Because of this, the subject of gaslighting must defer to the conceptual judgment of her gaslighter if she is to have some idea as to how to act 'appropriately' in a context. This is a form of psychological and practical control. It forces the subject of gaslighting to 'outsource' her interpretive abilities. Her judgment about an event is dependent on how her gaslighter sees things. Because the subject of gaslighting does not have sufficient control over her interpretation, owing to persistent challenges to her conceptual understanding, she is undermined in her capacity as an epistemic subject. She suffers from epistemic injustice.

4. A Remaining Question

The foregoing aimed to answer the following question: *In what sense does one doubt their understanding of reality when gaslit?* I suggested that gaslighting can cause someone to doubt their understanding of reality in different ways depending on the linguistic order at which a gaslighting exchange takes place. This marks a distinction between *first order* and *second order* gaslighting. Given its omission in the philosophical literature, I focused on second order gaslighting: A form of gaslighting that involves inducing doubt in someone by targeting their conceptual understanding. I argued that it involves three wrongs: (1) *metalinguistic deprivation*, (2) *conceptual obscuration*, and (3) *perspectival subversion*. Each are types of epistemic injustice. And each reliably have sequelae in terms of psychological and practical control. I'll end by leaving open a question for future research: What can we do in the face of such wrongs?

References

- Abramson, Kate. 2014. Turning up the lights on gaslighting. *Philosophical Perspective* 28 (1), 1 – 30.
- Barker, Chris. 2002. The dynamics of vagueness. *Linguistics and Philosophy* 25, 1 -36.
- Berenstain, Nora. Forthcoming. White feminist gaslighting. *Hypatia*.
- Bicchieri, Cristina. 2017. *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford: Oxford University Press.
- Camp, Elisabeth. 2015. Logical concepts and associative characterizations. In *The conceptual mind*, ed. Eric Margolis and Stephen Laurence, 591 – 621. MA: MIT Press.
- Camp, Elisabeth. 2018. Perspectives in imaginative engagement with fiction. *Philosophical Perspectives* 31 (1), 73 – 102.
- Cappelen, Herman. 2018. *Fixing language: An essay on conceptual engineering*. Oxford: Oxford University Press.
- Dotson, Kristie. 2012. A cautionary tale: On limiting epistemic oppression. *Frontiers: A Journal of Women Studies* 33 (1), 24 – 47.
- Fricke, Miranda. 2007. *Epistemic injustice: Power and the ethics of knowing*. Oxford: Oxford University Press.
- Fricke, Miranda. 2013. Epistemic injustice as a condition of political freedom? *Synthese* 190, 1317 – 1332.

Gerken, Mikkel. 2019. Pragmatic encroachment and the challenge from epistemic injustice. *Philosophers' Imprint* 19.

Haslanger, Sally. 2012. *Resisting reality: Social construction and social critique*. Oxford: Oxford University Press.

Haslanger, Sally. 2020. Going on, not in the same way'. In *Conceptual ethics and conceptual engineering*, ed. Alexis Burgess, Hermen Cappelen, and David Plunkett. Oxford: Oxford University Press.

Hawthorne, John and Stanley, Jason. 2008. Knowledge and action. *Journal of Philosophy* 105: 571 – 590.

McKinnon, Rachel. 2017. 'Allies' behaving badly: Gaslighting as epistemic injustice'. In *The Routledge handbook of epistemic injustice*, ed. Ian James Kidd, Gaile Pohlhaus Jr., and Jose Medina. London: Routledge.

Plunkett, David and Tim Sundell. 2013. Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint* 13.

Pohlhaus Jr., Gaile. 2011. Relational knowing and epistemic injustice: Toward a theory of willful hermeneutical ignorance. *Hypatia* 27 (4): 715 – 735.

Pohlhaus Jr., Gaile. 2017. Varieties of epistemic injustice. In *The Routledge handbook of epistemic injustice*, ed. Ian James Kidd, Gaile Pohlhaus Jr., and Jose Medina. London: Routledge.

Ruíz, Elena. Forthcoming. Cultural gaslighting. *Hypatia*.

Sarkis, Stephanie. 2017. 11 Warning Signs of Gaslighting. *Psychology Today*.
<https://www.psychologytoday.com/au/blog/here-there-and-everywhere/201701/11-warning-signs-gaslighting>.

Spear, Andrew D. 2018. Gaslighting, confabulation, and epistemic innocence. *Topoi* 39 (1): 229 – 241.

Spear, Andrew D. 2019. Epistemic dimensions of gaslighting: Peer-disagreement, self-trust, and epistemic injustice'. *Inquiry* DOI: 10.1080/0020174X.2019.1610051

Stark, Cynthia A. 2019. Gaslighting, misogyny, and psychological oppression. *The Monist* 102: 221 – 235.

Thomasson, Amie. 2017. Metaphysical disputes and metalinguistic negotiation. *Analytic Philosophy* 58 (1): 1 – 28.

Weber, Max. 1978. *Economy and society: An outline of interpretive sociology*. Berkeley: University of California Press.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford: Oxford University Press.

¹ There is no explicit mention of this distinction between kinds of gaslighting in the existing philosophical literature. Kate Abramson (2014) gets close. She offers a real world example – her sixth illustrative example out of eight – that looks like second order gaslighting: A female graduate student confronts someone who has sexually harassed her and is told that she is just being a ‘prude’. We can’t yet tell whether this counts as first or second order gaslighting. It could be revealed that they have the same concept of *sexual harassment* and this would ‘resolve’ the dispute; or it could be revealed that the reason why the man calls her a ‘prude’ is because he has a very narrow idea of what counts as sexual harassment. Cynthia A. Stark (2019) offers a similar example, except with the concept of *flirting*. This case is also under-specified. Other theorists who discuss the idea that concepts are implicated in forms of gaslighting are Elena Ruiz (forthcoming), and Nora Berenstain (forthcoming). However, Ruiz and Berenstain are primarily interested in gaslighting that occur at higher ontological levels, such as social structure and culture, not strictly about the metalinguistic dynamics of interpersonal gaslighting.

² For competing accounts, see Abramson (2014), Spear (2018, forthcoming), Stark (2019).

³ For detailed discussion on the relationship between gaslighting and self-trust see Spear (2018, forthcoming).

⁴ In a domain, or across domains.

⁵ Karen Jones (2012) has a really nice example that captures the role of affect in self-trust. She recounts an experience where she's at the airport constantly checking to see whether her passport is still in her pocket. She keeps checking her pocket even though she has no reason to doubt her perceptual and cognitive faculties. The feeling of doubt and anxiety is enough for her to disregard compelling reasons. To me, this shows that stakes are an important factor in affect and self-trust – after all, losing one's passport is a real nuisance when travelling in a foreign country. When it comes to gaslighting, stakes might also be an important factor. For example, if a wife catches her husband cheating, she might be more disposed to doubt herself if it means having to end the relationship. Given the power structures in play, leaving a marriage can be costly. Unfortunately I don't have room to explore this thought further.

⁶ And it would rob the notion of gaslighting of political usefulness.

⁷ Note that 'naïve' doesn't refer to the gaslighter, but to the definition.

⁸ We might want to add in a fourth condition that states that the intention to induce doubt in a hearer is *unwarranted*. This means that cases in which someone has good reason to convince another to doubt their interpretive abilities aren't counted as gaslighting. For example, one might get fed up with a conspiracy theorist who endorses the view that vaccines cause autism, and instead of offering arguments, one might simply tell the other person that they are 'crazy'. I want to keep open the question of whether this counts as gaslighting, though I admit it strays from typical cases.

⁹ This case of unintentional gaslighting is similar to a case discussed by Rachel McKinnon (2017) in which a trans woman is misgendered by an ally, and when she reports this to a friend, the friend says that she must have misheard the ally.

¹⁰ Even if the man did come to doubt himself after being dismissed, it would still not count as gaslighting. This is because the man is not subject to systematic credibility deflation due to him *being a man*.

¹¹ This disjunctive definition requires that *either* the intentional gaslighting conditions need to hold *or* the unintentional gaslighting conditions need to hold, in order for gaslighting to occur.

¹² Here I have in mind Sally Haslanger's (2012) notion of *schema* that undergirds practice, and where such practices constitutes social structure.

¹³ The phrase 'subject of gaslighting' is a mouthful. Others in the literature have used phrases such as 'target' or 'victim' of gaslighting. I disagree with the former on the grounds that it gives the impression that gaslighting is always intentional. And I wish to avoid the latter because the language of victimhood has negative social meaning attached to it.

¹⁴ This is not to be confused with the claim that the gaslighter and the subject of gaslighting aren't trying to coordinate on *some* concept – this I take to be what we're doing all the time. The claim is that the gaslighter and the subject of gaslighting aren't trying to coordinate on either of their preferred concept.

¹⁵ Gaslighting can occur without the subject of gaslighting offering testimony. For instance, someone might be subject to constant claims that they are 'crazy' throughout a relationship, even without saying a word, and this can prompt the recipient of such claims to doubt their interpretive abilities.

¹⁶ Sally Haslanger, an advocate of conceptual engineering, has said 'I don't know what concepts are, or even if there are any' (2020: 1). Herman Cappelen (2018) has expressed something similar. And apart from skepticism about the existence of concepts, there is also the problem that there are many theories about what concepts are. See Margolis and Laurence (1999).

¹⁷ See also Sally Haslanger (2012) and Christina Bicchieri (2017) on schemas/schemata/scripts.

¹⁸ Chris Barker (2002) calls this metalinguistic usage.

¹⁹ It needs to be made clear that the man believes that the event is *not* sexual harassment, and not that it is sexual harassment but there are excusing factors.

²⁰ Notice that this definition does not specify sufficient conditions.

²¹ I want to highlight the fact that 'sexual harassment' has a legal definition. Because of this, there are a couple of ways that we can understand the second order dispute. First, we can understand the dispute as the participants operating within the bounds of the legal definition of 'sexual harassment', but contesting some of its constituent terms. Second, we can understand dispute as concerning the legitimacy of the legal concept itself – the legal definition may not be just or accurate (or both). Thank you to an anonymous reviewer for this suggestion.

²² This thought is expressed by Mitchell Green (2016) who claims that when speakers recognize that contexts are defective (when one speaker is mistaken about what is common ground), they will do what they can to correct it.

²³ Note that the gaslighter and subject of gaslighting needn't take themselves to be having a disagreement about concepts. This awareness is too cognitively demanding. In a sense, *recognition* that the dispute is about concepts may be a first step for countermeasures against gaslighting. Thank you to an anonymous reviewer for raising the point.

²⁴ Thank you to an anonymous reviewer for pushing me on this.

²⁵ By 'consideration', I mean consideration to *some extent*. I don't have space in this paper to detail exactly the amount of consideration one ought to receive.

²⁶ For more on the distinction between *mere* epistemic injustice and *discriminatory* epistemic injustice see Fricker (2013) and Gerken (2019).

²⁷ This is close to the *power-over* concept of power. See Weber (1978).

²⁸ And perhaps take on the attitude that she is unreliable when it comes to picking out sexual harassment; and that she is better off deferring to the judgment of her colleague. Though, this is probably more likely if she suffers from perspectival subversion (see section 4.3).

²⁹ Conceptual obscuration can also occur when one possesses a *more* accurate concept, but in virtue of metalinguistic deprivation comes to adopt a *less* accurate concept. The idea is that the initial concept that is lost might not perfectly represent facts about the world, but it does a better job than the concept that is subsequently adopted.

³⁰ The primary function of many concepts isn't to reflect the world. As a result, they are not susceptible to standards of accuracy. Thus, for such concepts, conceptual obscuration is not a possibility. Thank you to an anonymous reviewer for this suggestion.

³¹ Another way to put this is that the woman would be in violation of (contested) knowledge norms if she accuses her husband of cheating. Because she doesn't 'know' that her husband cheated on her, she isn't licensed to: *assert* that he cheated on her (norm of assertion); use cheating in her *practical reasoning* (norm of action); and form *beliefs* about her husband cheating (norm of belief). For more on knowledge norms, see: Williamson (2000) and Hawthorne and Stanley (2008).

³² Including very mundane activities such as knowing how to catch a bus.

³³ How much one defers comes in degrees within domains.

³⁴ More than this, the wife might be scared to face the wrath of her husband if she accuses him of cheating (again). This isn't directly connected to gaslighting, but poses a barrier to the wife's ability to build confidence in making conceptual judgements. Thus, her perspectival subversion becomes more difficult to escape.