Philos Stud (2023) 180:745–772
https://doi.org/10.1007/s11098-022-01872-9

S.I. : PACIFIC APA 2020 & 2021

Dilemmatic gaslighting

Cameron Domenico Kirk-Giannini[1]

Accepted: 8 August 2022 / Published online: 15 September 2022

Abstract Existing work on gaslighting ties it constitutively to facts about the intentions or prejudices of the gaslighter and/or his victim's prior experience of epistemic injustice. I argue that the concept of gaslighting is more broadly applicable than has been appreciated: what is distinctive about gaslighting, on my account, is simply that a gaslighter confronts his victim with a certain kind of choice between rejecting his testimony and doubting her own basic epistemic competence in some domain. I thus hold that gaslighting is a purely epistemic phenomenon—not requiring any particular set of intentions or prejudices on the part of the gaslighter—and also that it can occur even in the absence of any prior experience of epistemic injustice. Appreciating the dilemmatic character of gaslighting allows us to understand its connection with a characteristic sort of epistemic harm, makes it easier to apply the concept of gaslighting in practice, and raises the possibility that we might discover its structure and the associated harm in surprising places.

Keywords Gaslighting · Epistemic injustice · Testimonial injustice · Social epistemology · Conceptual engineering

✉ Cameron Domenico Kirk-Giannini
camerondomenico.kirkgiannini@gmail.com

[1] Department of Philosophy, Rutgers University Newark, 175 University Avenue, 430 Conklin Hall, Newark, USA

1 Introduction

The term 'gaslighting' is usually introduced ostensively: it is the phenomenon paradigmatically exemplified by Gregory Anton's treatment of his wife Paula in the 1944 film *Gaslight*.[1]

> [Central Case]: Gregory seeks to rob Paula of her aunt's jewels, which are hidden in her attic. He routinely searches the attic, at which times the sound of his footsteps and the dimming of the house's gaslights are clearly perceptible to Paula. But when Paula discusses her observations with Gregory, he insists that she is merely imagining the footsteps and dimmings. Distressed, Paula begins to fear that she is losing her sanity.

This ostensive characterization is usually accompanied by an effort to explicate the concept by providing a definition that clarifies the range of possible cases which fall in its extension. But, as is often the case with ostensive definition, the subsequent project of explication is neither straightforward nor uncontroversial. Must gaslighting be intentional? Must it involve beliefs that are justified on the basis of perception or memory, as Paula's are in the 1944 film? Does gaslighting always occur along pre-existing gradients of social power?

In what follows, I develop my own characterization of gaslighting. My account draws on earlier characterizations by Abramson (2014), Ivy (2017), Stark (2019), Spear (2019, 2020), and Podosky (2021), but differs from them in that it focuses on understanding gaslighting neither in terms of the intentions or prejudices of the gaslighter nor in terms of the prior epistemic injustice experienced by his victim. Instead, my account understands gaslighting as a phenomenon individuated by the distinctive sort of epistemic dilemma with which it presents its victim—a certain kind of choice between rejecting the perpetrator's testimony, on the one hand, and doubting her basic epistemic competence in some domain, on the other.

Before one can argue for or against any particular characterization of gaslighting, however, one must address methodology. What should a satisfactory account of gaslighting be like? I will proceed on the assumption that a useful theoretical characterization of gaslighting will have two features: first, it will have in its extension a range of similar cases which includes most or all cases that intuitively count as gaslighting and excludes most or all cases that intuitively do not count as gaslighting; second, it will explain what is distinctive about this range of cases in a way that makes the concept of gaslighting a fruitful one with which to theorize.

Importantly, I do not conceive of my project in what follows as that of articulating the single correct analysis of the concept of gaslighting. I suspect that there are many reasonable ways of constructing a precise account of gaslighting from the tissue of our pretheoretical intuitions about [Central Case] and its kin. There may even be substantial differences between individuals regarding which cases count as paradigms of gaslighting, so that talk of "our pretheoretical intuitions" is an oversimplification. So, while I argue below that existing accounts

[1] The 1944 film is itself based on Patrick Hamilton's 1938 play *Gas Light*.

of gaslighting fail to categorize certain intuitive cases of gaslighting as such, I do not take this sort of argument to constitute a refutation of those accounts. The project is rather to construct an account of gaslighting which does as well as possible at satisfying the two desiderata of being theoretically fruitful and accommodating our pretheoretical intuitions. My contention is that the account of gaslighting I defend below is both faithful to our intuitions and useful in grouping together a set of cases which share a certain sort of epistemic structure and explaining why cases with this structure are apt to cause a characteristic sort of harm. I believe my proposal does better than existing accounts at capturing our intuitions, but I do not wish to exclude the possibility that those accounts are ultimately justified by their theoretical benefits in some domain. Just as there are different tools for different tasks, it may be that there are multiple theoretically important concepts of gaslighting—indeed, this seems likely given that the term was introduced into philosophical discourse ostensively—and I wish only to defend the position that my preferred account is one of them.[2]

In embracing theoretical fruitfulness as a desideratum on a theory of gaslighting, I mean to ally myself with a conceptual engineering approach to my subject matter. Following Cappelen (2018, 3), I understand conceptual engineering as the project of "assessing and improving our representational devices." As an example of conceptual engineering in philosophy, Cappelen points to Clark and Chalmers's (1998) argument that the most theoretically fruitful concept of belief will allow for beliefs to extend beyond the limits of the brain, so that, for example, an individual's beliefs can (quite literally) be recorded in a pen-and-paper notebook. Clark and Chalmers argue for this surprising conclusion by motivating the idea that there is no scientifically interesting difference between the functional role played by neural beliefs in normal humans and the functional role played by records in a pen-and-paper notebook in certain special cases.

Though I conceive of it as an exercise in conceptual engineering, my project differs from that of Clark and Chalmers in two respects. First, whereas Clark and Chalmers advocate a revisionary concept of belief on the basis of considerations of theoretical utility, I see the project of developing a theory of gaslighting as one of refinement rather than revision—we start with a large and diverse set of intuitions about particular cases and seek to develop a theoretically useful account of gaslighting which reveals some important aspect of their underlying structure. Second, for this reason I seek to articulate an account of gaslighting which remains faithful to our pretheoretical intuitions. Nevertheless, the idea of assessing and improving our representational devices is the same.[3]

[2] I am thus open to endorsing a pluralism about the concept of gaslighting akin to the gender-concept pluralism defended by Jenkins (2016).

[3] My discussion of conceptual engineering may call to mind a certain ameliorative tradition in social philosophy, as represented for example by Haslanger's (2000) treatment of gender and race and Dembroff's (2016) approach to the concept of sexual orientation. This tradition seeks to engineer concepts in order to achieve specific social and/or political goals. Though I do believe that the account of gaslighting I defend in what follows would, if widely adopted, have certain concrete positive social consequences, it is beyond the scope of this paper to argue for this conclusion. For this reason, I have

What makes my proposed account of gaslighting theoretically fruitful? I will answer this question in detail in Sect. 8 after presenting and clarifying my proposal. For now, I offer the following promissory note: the account of gaslighting I favor characterizes the structure of gaslighting in a way which clarifies why gaslighting is apt to cause a characteristic sort of epistemic harm while making it easier in practice to identify particular instances of gaslighting and opening up the theoretical possibility that this epistemic structure—and therefore the associated harm—might be discovered in surprising places.

In what follows, I begin by discussing existing characterizations of gaslighting, dividing them into two broad categories: those which tie gaslighting constitutively to the intentions of the gaslighter, and those which do not. Section 2 presents existing accounts in the first category and argues that they face an undergeneration problem. Section 3 considers existing accounts in the second category and argues that they, too, make unintuitive predictions. Section 4 introduces my preferred account of gaslighting, while Sects. 5 and 6 clarify that account and explain my view of the relationship between gaslighting, harm, and wrong action. Section 7 addresses the worry that my account overgenerates. Section 8 concludes by returning to the question of what makes my account theoretically fruitful.

2 Intentionalism

2.1 Intentionalism and anti-intentionalism

There is a trivial sense in which all gaslighting is intentional, since in order to gaslight his victim a perpetrator must do or say something to her, and doings-to and sayings-to are intentional. But it is frequently claimed that gaslighting is also intentional in a more demanding sense: that part of what makes an act an act of gaslighting is that the agent who performs it possesses certain intentions or desires to subvert or control his victim. Let us introduce the term *intentionalism* to describe accounts of gaslighting which endorse this claim. Correspondingly, let us introduce the term *anti-intentionalism* to describe accounts of gaslighting which deny that gaslighting as an act is partially individuated by the intentions or desires of the gaslighter.

Both intentionalism and anti-intentionalism have been represented in the philosophical literature on gaslighting since its inception. In the first major philosophical treatment of gaslighting, for example, Abramson (2014, 2) characterizes it as "a form of emotional manipulation in which the gaslighter tries (consciously or not) to induce in someone the sense that her reactions, perceptions, memories and/or beliefs are not just mistaken, but utterly without grounds," and goes on to argue that it should be individuated as a phenomenon partly in terms of the gaslighter's intentions and desires for control. Ambramson's early intentionalist

Footnote 3 continued
focused on its theoretical fruitfulness rather than its possible social benefits. Thanks to an anonymous referee for pressing me to clarify this point.

account contrasts with Ivy's (2017) early anti-intentionalist alternative: Ivy argues that there can be unintentional cases of gaslighting and seeks to understand gaslighting as a species of testimonial injustice involving disregard for the victim's first-person authority.[4] More recent authors remain divided between the two camps: Stark (2019) and Spear (2019, 2020) are intentionalists; Podosky (2021) is an anti-intentionalist.

As we will see, intentionalism and anti-intentionalism differ in terms of the sorts of cases beyond [Central Case] they recognize as gaslighting. In this section and the next, I will argue that while existing intentionalist and anti-intentionalist accounts do much to reveal the structure of the phenomenon, we have reason to embrace a novel proposal.

2.2 Notable intentionalist proposals

As we have seen, Abramson (2014) holds that gaslighting is constitutively tied to certain intentions and desires for control on the part of the gaslighter. She writes, "it [is] the destructive impulse that distinguishes the gaslighter's aim as the aims of a gaslighter: it's the intense anxiety and fear about challenge, the need to destroy that possibility that drives him to gaslighting" (12). Thus Abramson maintains that gaslighters attempt to insulate themselves from criticism by intentionally (whether consciously or unconsciously) undermining their victims' status as knowers so much that those victims are no longer in a position to offer criticism.

Two more recent intentionalist proposals should also be mentioned here. First, Stark (2019) develops an account of one type of gaslighting (what she calls *manipulative gaslighting*; this is to be contrasted with what she calls *epistemic gaslighting*, which is not constitutively intentional) that agrees with Abramson's account in characterizing gaslighting as a form of manipulation but differs from Abramson's account in holding that gaslighting is individuated by the methods the perpetrator uses rather than his motivations. Stark is explicit that her account of manipulative gaslighting is intentionalist: "Manipulative gaslighting is, by definition, intentional because manipulation is, by definition, intentional in the following sense: the manipulator always has an aim. He is attempting to get someone to do or to feel something." (2019, 223). Second, Spear (2019, 2020) maintains the core intentionalist commitment of Abramson but emphasizes that gaslighting also has an important epistemic dimension. I discuss this epistemic aspect of Spear's work in detail in Sect. 4; for now, it suffices to notice that he embraces Abramson's version of intentionalism.

2.3 Intentionalism and undergeneration

Recall that the first of our two desiderata for an account of gaslighting was that it have in its extension a range of similar cases which includes most or all cases that

[4] Here and elsewhere, I follow Ivy in using the term *testimonial injustice* as it is defined by Fricker (2007)—that is, so that a testifier experiences testimonial injustice just in case she experiences an "identity-prejudicial credibility deficit" (Fricker 2007, 28).

intuitively count as gaslighting and excludes most or all cases that intuitively do not count as gaslighting. In this subsection, I argue that this desideratum can be used to motivate an argument against intentionalism. The basic problem for intentionalism is that a range of cases which have come to be considered paradigms of gaslighting appear not to require that the gaslighter have the kinds of intentions which, according to intentionalists, are constitutive of gaslighting.

A number of authors have felt the force of this type of objection to intentionalism. Podosky (2021), for example, writes:

> "Think of a situation in which a man brushes up against a woman's bottom 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 about her testimony. Nevertheless, this looks like a cut-and-dried case of gaslighting. The lesson seems to be that not all cases of gaslighting are intentional, though perhaps a great many of them are." (2021, 210)

Though Ivy does not discuss the connection to intentionalism as explicitly, Podosky notes that much the same lesson could be taken from her discussion of a case involving a trans woman hearing her colleague repeatedly mispronoun her (Ivy 2017, 168). And it is perhaps because of the same difficulty with extending intentionalism to cover cases like these that Stark (2019) presents her intentionalist proposal as concerning only manipulative gaslighting rather than gaslighting *simpliciter*.

Even some cases which are explicitly offered by intentionalists as examples of gaslighting raise issues for intentionalism. Consider the following case, which Abramson presents as one of eight core cases of gaslighting:

> [Skeptical Peers]: "I moved out of one field of philosophy in grad school due to an overwhelming accumulation of small incidents... When I tried to describe to fellow grad students why I felt ostracized or ignored because of my gender, they would ask for examples. I would provide examples, and they would proceed through each example to 'demonstrate' why I had actually misinterpreted or overreacted to what was actually going on." (Abramson 2014, 5)[5]

I agree with Abramson and others that cases like [Skeptical Peers] are intuitively cases of gaslighting. But there are two related points I wish to emphasize about [Skeptical Peers].

First, the case as Abramson presents it is underspecified: it does not tell us anything about the intentions of the fellow graduate students. Second, there are various ways in which the case could be fleshed out. We can imagine that the perpetrators of the gaslighting in [Skeptical Peers] do indeed have the kinds of

[5] This case first appeared on the blog *What Is It Like To Be A Woman In Philosophy?*.

subterranean motivations Abramson regards as individuative of gaslighting. But we can also imagine that they do not. Perhaps they are all hardened misogynists, who genuinely believe that the victim is unable correctly to interpret her own experiences because she is female. Perhaps they have no real aversion to having their misogynistic worldview challenged; they just think, on the basis of their pre-existing misogyny, that the victim is not competent to assess whether she is being ostracized or ignored because of her gender.[6] Or perhaps they are wholly indifferent towards both the narrator in [Skeptical Peers] and those who have ostracized and ignored her, seeking to demonstrate that she has interpreted things incorrectly solely out of perverse and antisocial contrarianism.

The fact that we can identify [Skeptical Peers] as a case of gaslighting without knowing about the intentions of the gaslighters suggests that our judgment about the case is not sensitive to facts about those intentions. This conclusion is further suggested by the observation that our intuitive sense that the victim's fellow graduate students are gaslighting her persists when we fill out the case so that they lack an intention to subvert or control her. If this line of argument is sound, it must be possible for there to be gaslighting in the absence of the psychological features Abramson and other intentionalists identify, common or salient though those features may be. Indeed, the wide variety of ways in which it is possible to flesh out [Skeptical Peers] without undermining our intuition that it involves gaslighting suggests that a gaslighter need not have *any* particular intention vis-à-vis his victim beyond the intention to communicate whatever proposition or propositions constitute his gaslighting.[7]

To be clear, I do not take the above considerations to constitute a *refutation* of intentionalism. As I have emphasized above, it seems to me unlikely that our pretheoretical intuitions about gaslighting are clear or systematic enough to serve as the basis for a refutation of any existing proposal. Some readers may find that their initial intuitions about [Skeptical Peers] come along with substantive assumptions regarding the intentions of the victim's peers and disappear when those assumptions

[6] Why might such hardened misogynists take the time to "proceed through each example to 'demonstrate' why [the victim] had actually misinterpreted" what was going on? We might imagine that they are serial pontificators who love to hear the sounds of their own voices, so that their "demonstrations" are something like a form of recreation for them. Alternatively, we might imagine that the victim herself requests an explanation of why they do not believe her account of what has happened to her. Modifying the case in these ways does not affect the intuitive force of the judgment that it involves gaslighting.

[7] Ambramson (2014, 11) considers an objection along these lines:

> "...it may be less than entirely clear that all of the examples with which I began are examples of gaslighting... A single instance of one person saying to another, "that's crazy" may not appear—may not *be*—an instance of someone trying to destroy another's standing to make claims. But when that form of interaction is iterated over and over again, when counterevidence to "that's crazy" is dismissed, when nothing is treated as salient evidence for the possibility of disrupting the initial accusation, appearances shift." (emphasis in original)

In the two imagined scenarios I have sketched above, however, I do not think that moving from a one-off exchange to a pattern of interactions renders plausible the idea that the gaslighters have the kinds of motivations Abramson's account requires. In any case, as Podosky's example demonstrates, we can have a strong intuition to the effect that a one-off exchange constitutes gaslighting.

are not met. Nevertheless, for those of us who, like Podosky, Ivy, and myself, have no trouble imaging that [Skeptical Peers] and related cases might involve hardened misogynists or perverse contrarians, intentionalism fails to satisfy one of the key desiderata for a theory of gaslighting: it undergenerates by failing to characterize as gaslighting certain paradigmatic cases of gaslighting.[8]

While [Skeptical Peers] and related cases do not constitute a refutation of intentionalism, they give us theoretical motivation to develop anti-intentionalist accounts of gaslighting. If a theoretically fruitful and otherwise plausible anti-intentionalist proposal can be devised, the fact that it would also more easily be able to accommodate the intuition that [Skeptical Peers] is a case of gaslighting would give it a strong claim to be preferable all-things-considered to competing intentionalist proposals.[9] I present and defend an anti-intentionalist proposal of this type below.

3 Anti-intentionalism

3.1 Two anti-intentionalist proposals

The literature on gaslighting contains two major anti-intentionalist accounts. First, Ivy (2017) identifies what she calls a "subtle epistemic form" of gaslighting as that phenomenon which occurs, "often unintentional[ly]," when "a listener... raises doubts about the speaker's reliability at perceiving events accurately" (168). Ivy's primary focus is on cases in which the victim, but not the perpetrator, belongs to a marginalized group and is therefore better situated than the perpetrator to perceive events accurately; for this reason, she characterizes gaslighting as a form of testimonial injustice (that is, as a distinctively epistemic form of identity-based prejudice).

Podosky (2021) defends a disjunctive account of gaslighting. For Podosky, gaslighting must either be intentional, in which case it must work more or less along the lines Abramson identifies, or unintentional, in which case it must work more or less along the lines Ivy identifies. Because the account is disjunctive, it is a version of anti-intentionalism: it holds that gaslighting can occur in the absence of intentions to subvert or control the victim. Officially, Podosky's proposal is:

> "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

[8] Since anti-intentionalism differs from intentionalism in dropping a necessary condition from its account of gaslighting, one might reasonably worry that anti-intentionalist accounts will overgenerate rather than undergenerate. While it is difficult to evaluate this objection without looking at particular anti-intentionalist proposals, I address it in the context of my preferred account of gaslighting in Sect. 7.

[9] Or, if those competing proposals were also found to be theoretically useful in some domain, at least to be no less appealing than they are.

> doubt her interpretive abilities (vi) owing to the hearer being subject to systematic epistemic injustice that has disposed her to do so." (2021, 212)

Podosky and Ivy agree in holding that there is a constitutive connection between gaslighting and epistemic injustice, but differ on what this connection is: Ivy's account requires that the gaslighter's gaslighting itself constitute testimonial injustice, whereas Podosky requires that in unintentional cases the success of the gaslighter's gaslighting be explained in some way or other by the victim's prior experience of epistemic injustice. Despite this important difference, however, I argue that tying gaslighting to epistemic injustice in either of these ways leads to an undergeneration problem for Ivy and Podosky similar to the one faced by the intentionalist proposals discussed above.[10]

3.2 A problem case

Though I agree with Ivy and Podosky that gaslighting often involves the privileged targeting the marginalized and that it is at its most problematic when it is bound up with prior epistemic injustice, I think it is intuitively clear that there are cases of gaslighting in which epistemic injustice of the sorts envisioned by Ivy and Podosky is not implicated because there is no pre-existing social power differential between victim and perpetrator. For example, the film *Gaslight* could have been written about a same-gender couple without losing the distinctive character of its plot.

Suppose for illustration that we replace Paula in [Central Case] with Paul, imagining that Gregory and Paul are two cisgender gay men who have never experienced significant prejudicial treatment on the basis of either their gender or their sexual orientation and who harbor no identity-based prejudices against each other. Still, Gregory's treatment of Paul seems to fall squarely into the extension of the concept of gaslighting. And this is so despite the fact that it is implausible (i) that Gregory disregards Paul's testimony about the gaslights because of identity prejudice of some kind (rather, he disregards it because he is a thief and a scoundrel), and (ii) that Paul is led to doubt his interpretive abilities after interacting with Gregory because he has been subject to systematic epistemic injustice (rather, he doubts them because someone he trusts has called them into question).

Similarly, in the version of [Skeptical Peers] involving perverse and antisocial contrarianism, we can imagine that the gaslighters harbor no identity-based prejudice against the victim, and that she is led to doubt her perceptions because they have been directly challenged by her peers rather than because she has been subject to systematic epistemic injustice. So neither intentional gaslighting nor unintentional gaslighting appears to be constitutively connected to epistemic injustice in the way Ivy and Podosky imagine.[11]

[10] As we will see in Sect. 5, my preferred account of gaslighting also holds that it is intimately connected to epistemic injustice—though not in a way that leads to undergeneration worries.

[11] Stark (2019, 223) agrees that gaslighting need not occur along pre-existing gradients of social identity power.

Indeed, in more recent work, Ivy has distanced herself from the claim that gaslighting must involve identity-based prejudice (and possibly also the claim that it must involve epistemic injustice) for exactly this reason: "I'm no longer convinced that the source of the credibility deficit being an identity-based prejudice is a necessary condition for testimonial injustice (and thus gaslighting, if gaslighting is properly understood in terms of testimonial injustice)" (Ivy 2019, 288–9).

So existing anti-intentionalist proposals seem to me to enforce the wrong kind of connection between gaslighting and epistemic injustice, which causes them to do poorly when judged against our first criterion of adequacy for accounts of gaslighting: including a range of similar cases which contains most or all cases that intuitively count as gaslighting. This is not, however, an essential feature of anti-intentionalism—as I argue below, it is possible to construct an anti-intentionalist account of gaslighting which is not subject to this problem.

Before introducing my preferred account, however, I would like to emphasize that I think that Ivy's and Podosky's characterizations of gaslighting are illuminating in two important respects. First, they allow that gaslighting can occur in the absence of any particular intention or set of motivations on the part of the gaslighter. Second, Ivy argues that "gaslighting... constitutes a failure to afford the first person (epistemic) authority of disadvantaged speakers [its] appropriate epistemic weight" (2017, 170), and both Ivy and Podosky characterize gaslighting as (at least sometimes) individuated in terms of primarily epistemic criteria. As we will see in the next section, the first-person epistemic authority of individuals over their perceptual and certain other beliefs—understood in terms of a presumption that individuals possess basic epistemic competence in forming such beliefs—figures centrally in the anti-intentionalist, purely epistemic account of gaslighting I favor.

4 The dilemmatic account

Recall again the exchange between Gregory and Paula regarding the gaslights. Paula believes that the brightness of the gaslights has fluctuated. She believes this on the basis of direct visual inspection of the lights over the period during which their brightness fluctuated. Nevertheless, Gregory testifies that the brightness of the lights has remained constant. So Gregory's testimony presents Paula with an unpleasant menu of options: either she is mistaken about her own perceptual states, or her perceptual states do not reliably track the brightness of the lamps, or Gregory has said something false. On either of the first two options, Paula is an unusually defective epistemic agent: it is part of a sighted individual's basic epistemic competence to be able to form and preserve in memory correct beliefs about the brightness of everyday objects in her immediate environment. So Gregory's testimony in effect forces Paula into a dilemma: she must choose between herself

and Gregory. If she trusts him, she is epistemically defective; if she is not epistemically defective, he is not to be trusted.[12]

The idea that the structure of gaslighting is intimately connected with this sort of dilemma has been explored by Spear (2019, 2020). Spear emphasizes that the victim of gaslighting

> "...must adjudicate the question of whether her gaslighter's behavior and say-so constitute sufficient defeating reasons for her to downgrade or abandon entirely her own epistemic self-trust (her confidence in her cognitive abilities), or whether his claims instead constitute grounds for trusting *him* less and so downgrading her confidence in him" (2020, 232; emphasis in original).

Yet no philosophical work on gaslighting has yet explored the possibility that this dilemmatic structure can, by itself, serve as the core of an account of gaslighting—as we have seen, Spear himself accepts Abramson's intentionalism, holding only that that the dilemmatic nature of gaslighting indicates that "*all* gaslighting involves issues of epistemic status and trust" in addition to manipulation and control (2019, 8; emphasis in original).

The account I propose departs from existing work in that it takes the dilemmatic character of gaslighting, properly understood, to be individuative of the phenomenon: what it is for a perpetrator to gaslight his victim is, roughly, for him to put her in a position where she must either reject his testimony or believe that she is epistemically defective. Unlike existing intentionalist accounts, then, the account I propose individuates gaslighting in purely epistemic terms; unlike existing anti-intentionalist accounts, it does not enforce a constitutive connection between gaslighting and testimonial injustice (Ivy) or a victim's prior experience of systematic epistemic injustice (Podosky).[13]

I have employed the qualifications *properly understood* and *roughly* above because simplistic versions of this kind of purely epistemic, dilemmatic proposal make unintuitive predictions. Whenever one agent's testimony conflicts with a

[12] Two clarifications regarding the sense in which Paula faces a dilemma: First, levels of confidence are graded. One might rationally respond to a forced epistemic choice between believing p and believing q (an epistemic dilemma) by maintaining full confidence is one of the two propositions and disbelieving the other, or one might respond by assigning a subjective probability of 0.5 to each proposition. The crucial thing is that, when faced by such an epistemic dilemma, one cannot rationally maintain full confidence in both p and q: one has a certain amount of probability to distribute between the two options—enough to assign probability 1 to one of the two or probability 0.5 to both—and one can choose how one distributes it. Second, gaslighting presents the victim with a *normative* dilemma. That is, if she is forming beliefs rationally, she must choose how to assign subjective probabilities to the proposition that she lacks basic epistemic competence and the proposition that the gaslighter is telling the truth in such a way that those probabilities add up to 1 (certainty). But if she is forming beliefs irrationally, she might not do this. She might, for example, become confident both that she lacks basic epistemic competence and that the gaslighter is lying. But this doesn't mean that the original choice situation is not a dilemma. In the same way, someone who is faced with a practical dilemma (e.g. "Your money or your life!") might, if not acting rationally, choose to give up both. But this does not indicate that the choice situation is not a dilemma.

[13] When I say that my account of gaslighting is *purely epistemic*, I mean that it holds that all of the conditions which must be satisfied for an act of intentional communication to count as gaslighting are epistemic conditions—conditions involving knowledge, justification, testimony, and so forth.

second's beliefs, the second must make a decision about whether to keep those beliefs and disregard the testimony or accept the testimony and abandon the beliefs, in which case she may be led to reassess the reliability of whatever processes she used to form them. Without refinement, then, a dilemmatic account of gaslighting risks implausibly classifying almost all verbal disagreements as instances of gaslighting. This is perhaps why even those who, like Spear (2019, 2020), have emphasized the connection between gaslighting and peer disagreement have not explored the possibility that it could be analyzed in terms of its dilemmatic structure.[14]

What must be added to the bare structure of the epistemic dilemma faced by Paula in [Central Case] to yield a plausible anti-intentionalist account of gaslighting? First, it has seemed intuitive to many that the victim of gaslighting must take the perpetrator's testimony seriously. Spear articulates this intuition in terms of the concept of trust:

> "Even a stranger could gaslight someone, as long as the person being gaslighted believed that the stranger was essentially informed and sincere in his claims: as long as she trusted him. But without this basic threshold of trust, gaslighting simply wouldn't get off the ground. If the victim doesn't trust the person who is attempting to gaslight her, doesn't think that he is a basically sincere source of accurate information or evaluation, then the gaslighting project gets no traction." (2020, 232)

This intuition strikes me as basically correct, though I think talk of trust is potentially distracting—it does not matter whether the victim has a trusting attitude toward the perpetrator in general, as long as she assigns significant weight to the particular testimony which constitutes the gaslighting. So we must add to our account as a necessary condition that the victim of gaslighting assign significant weight to the perpetrator's testimony.

Second, though much previous work on gaslighting has tended to be relatively unspecific about what sorts of beliefs or faculties can be targeted by gaslighting, intuition suggests that only some are potential targets. Thus, while Podosky (2021, 212) writes broadly of the victim of gaslighting "doubt[ing] her interpretive abilities" and Spear (2019, 10) invokes the sweeping notion of self-trust in one's cognitive faculties, it seems to me that it is not possible (for example) for one paleontologist to gaslight another by suggesting that her considered view about what caused the extinction of the dinosaurs is implausible. This is because the kinds of interpretive abilities and cognitive faculties on which paleontologists rely in deriving their considered views from large and heterogeneous bodies of evidence are too sophisticated for calling them into question to lead to the kind of harm characteristic of gaslighting. Instead, it seems to me that Abramson gets things right when she explains that in gaslighting, "the accusations are about the target's *basic* rational competence—her ability to get facts right, to deliberate, her basic

[14] Thanks to Jeremy Goodman for pressing me to clarify how gaslighting on my account differs from mere disagreement.

evaluative competencies and ability to react appropriately" (2014, 8; emphasis added). So we must clarify that the relevant dilemmatic structure constitutes gaslighting only if it calls into question the victim's basic epistemic competence in some domain.

Third, it seems to me that there are a number of important further constraints on the epistemic states of both victim and perpetrator. This is an issue which has not received sufficient attention in existing work on gaslighting. For example, it is not intuitive to classify as gaslighting cases in which one agent correctly and with knowledge-level justification calls into question the basic epistemic competence of another. If Smith is known to be prone to hallucinatory episodes and remarks to Jones that the golden bat in the room is singing beautifully, Jones does not gaslight Smith by denying that this is the case. Since it will be easiest to discuss these conditions once they have been precisely formulated, I now present my proposed account of gaslighting:

> **(Dilemmatic Gaslighting):** For all persons A, B, and propositions p: A gaslights B with respect to p iff (i) A intentionally communicates p to B, (ii) B knows (and A is in a position to know) that if p is true, then B has good reason to believe that she lacks basic epistemic competence in some domain D, (iii) A does not correctly and with knowledge-level doxastic justification believe p, and A does not correctly and with knowledge-level doxastic justification believe that B lacks basic epistemic competence in D, and (iv) B assigns significant weight to A's testimony.

Here condition (iv) corresponds to the first point made above, and the appeal to the notion of basic epistemic competence is responsive to the second. Conditions (ii) and (iii) articulate the further epistemic constraints just mentioned.

Conditions (i)-(iii) of (Dilemmatic Gaslighting) demand further explication. As I understand condition (i), it is neutral between the many ways in which A might get p across to B. A could testify that p—but he could also insinuate p, conversationally or conventionally implicate p, or introduce p into the common ground through presupposition accommodation. Indeed, A might intentionally communicate p to B simply by refusing to accept or respond appropriately to B's testimony, as when a victim of sexual harassment describes her experience to a colleague who responds with a dismissive 'That's interesting' or 'I'm sorry you feel that way' (in such a case, p is something like *the victim's belief that she was sexually harassed lacks justification*). Despite the permissiveness of condition (i), however, for simplicity I will use the term *testimony* in what follows to refer to whatever intentional communicative act is at issue in the cases of gaslighting I discuss.

The purpose of condition (ii) is to exclude cases in which the connection between the proposition communicated by B and A's epistemic competence is not sufficiently clear to those involved. It might be, for example, that the universe is deterministic, and that if some proposition q concerning its initial conditions is true, then B lacks basic epistemic competence in some domain. Nevertheless, if neither A nor B knows of the connection between q and B's epistemic competence, it does not intuitively seem that A gaslights B by intentionally communicating q. It should be noted, however, that (ii) does make room for cases in which A fails to know that the

proposition he intentionally communicates to *B* bears on *B*'s basic epistemic competence in virtue of never entertaining the relevant conditional proposition. It is thus no escape from the charge of gaslighting to manifest so little concern for one's victim that one does not even consider her epistemic competence.[15]

There are many ways in which a gaslighter can satisfy condition (iii): He might believe both *p* and the proposition that *B* lacks basic epistemic competence in *D* falsely, or without knowledge-level doxastic justification, or both falsely and without knowledge-level doxastic justification; or he might fail to believe these two propositions altogether; or he might adopt different attitudes towards each of them, so long as he does not believe either correctly and with knowledge-level doxastic justification. In each of these cases, I suggest, the intuitive verdict is that his communicating *p* to *B* is gaslighting. If he believes neither *p* nor the proposition that *B* lacks basic epistemic competence in *D*, or if his beliefs in these propositions lack knowledge-level doxastic justification, he confronts *B* with a choice between his testimony and her basic epistemic competence without even the excuse that he reasonably takes his testimony to be true.[16] Even if he believes one of the two propositions with knowledge-level doxastic justification, if that proposition turns out to be false, intuition does not excuse him from the charge of gaslighting.[17] It thus seems that the only way to escape the charge of gaslighting is for one's relevant beliefs to both be true and have knowledge-level doxastic justification.[18]

[15] I maintain that *B* must know of the connection between *p* and her epistemic competence (rather than merely being in a position to know) because merely being in a position to appreciate the dilemma characteristic of gaslighting is not a state apt to cause the characteristic harm associated with the phenomenon.

[16] The same point applies if he fails to believe one proposition and believes the other without knowledge-level doxastic justification.

[17] Robin Dembroff suggests the following case with this structure, which motivates the intuition that knowledge-level doxastic justification is not enough to escape the charge of gaslighting: suppose you witness a hit-and-run accident caused by a blue car, and testify to this effect. Suppose further that you have an identical twin who is blue-to-green color blind. Someone who knows this confuses you and your twin, forming the justified false belief that you lack basic epistemic competence in forming beliefs about the colors of cars. If they insist that the car was green, not blue, they are gaslighting you. I take it that in this case, we are supposed to imagine that the person who has confused you with your twin lacks knowledge-level doxastic justification for the proposition that the car was green. If we instead assume that they have knowledge-level doxastic justification (for example, because they formed their false belief that the car was green on the basis testimony from a reliable source), intuitions are somewhat murkier. Readers who have strong intuitions about this modified case may be inclined to complicate (Dilemmatic Gaslighting) accordingly.

[18] Why not revise condition (iii) so that what is required is a lack of knowledge rather than a lack of true belief with knowledge-level doxastic justification? My inclination is to say that when knowledge fails for Gettier-type reasons, an individual is excused from the charge of gaslighting. This is, however, not an issue on which I wish to take a particularly strong stand.

5 Gaslighting, harming, and wronging

No account of gaslighting is complete without a discussion of why gaslighting is both harmful and wrong. In this section, I discuss how (Dilemmatic Gaslighting) explains these features of gaslighting.

(Dilemmatic Gaslighting) makes it clear how gaslighting causes a characteristic sort of epistemic harm to its victims: the victim of gaslighting must choose between regarding the perpetrator as an unreliable (and possibly malicious) informant and suspecting that she herself is a defective epistemic agent. If she chooses the latter path, her access to knowledge in the domain targeted by the gaslighting is likely to be radically undermined. This harm is different from the primary epistemic harms associated with various related phenomena, including testimonial injustice (Fricker 2007), willful hermeneutical ignorance (Pohlhaus 2012), epistemic violence (Dotson 2011), contributory injustice (Dotson 2012), and hermeneutical injustice (Fricker 2007). For example, Fricker (2007) argues that the victim of testimonial injustice is harmed in that she is wronged in her capacity as a knower because she is not believed. But the harm associated with gaslighting is a more pernicious one, for the victim of gaslighting who comes to doubt her own epistemic competence is debilitated as an author and possessor of knowledge and not just as a transmitter of knowledge.

Similar points can be made about other notions which are *audience-centered* in the sense that they focus on how a speaker or group's epistemic contributions are received by others: willful hermeneutical ignorance involves knowers in situations of social privilege refusing to take seriously the insights of knowers in subordinate social situations; epistemic violence involves hearers refusing to "communicatively reciprocate" due to pernicious ignorance (Dotson 2011, 238); contributory injustice involves dominant knowers' willful hermeneutical ignorance leading to them "maintaining and utilizing structurally prejudiced hermeneutical resources" (Dotson 2012, 31). In each of these cases, the primary epistemic harm to the victims is that their contributions—whether the the form of particular propositions conveyed through testimony or in the form of conceptual or other epistemic resources developed to understand the world from non-dominant standpoints—are not given the uptake they deserve. The harm characteristic of gaslighting is not audience-centered in this way; it affects whether agents themselves are able to use their basic epistemic competence to generate knowledge in the first place.

Fricker's concept of hermeneutical injustice arguably comes closest to gaslighting in terms of the primary harm it causes. For Fricker, hermeneutical injustice is "the injustice of having some significant area of one's social experience obscured from collective understanding owing to persistent and wide-ranging hermeneutical marginalization" (2007, 154). Like the harm characteristic of gaslighting, then, the primary harm done by hermeneutical injustice is a harm which prevents epistemic agents from coming to know certain significant propositions. Yet the primary harm caused by hermeneutical injustice also differs from the harm characteristic of gaslighting. The victim of gaslighting who trusts the testimony of her gaslighter does not simply miss out on knowledge of certain important propositions—she is

actively undermined in a fundamental way, losing both knowledge she already possesses and the opportunity to gain more knowledge in the future.

I have been careful to write of the *primary* harms caused by testimonial injustice and related phenomena because, as Fricker emphasizes, they are apt indirectly to cause a range of other harms. For example, the characteristic epistemic harm caused directly by gaslighting is similar to what Fricker (2007, 47-8) identifies as one species of secondary (i.e. indirect) harm associated with testimonial injustice: "...someone with a background experience of persistent testimonial injustice may lose confidence in her general intellectual abilities to such an extent that she is genuinely hindered in her educational or other intellectual development." Fricker (2007, 163) makes a similar point about hermeneutical injustice ("the sorts of epistemic disadvantages at stake [in cases of hermeneutical injustice] are the very same as those we discussed at some length in respect of testimonial injustice, for they once again stem most basically from the subject's loss of epistemic confidence"), and indeed similar points could be made about willful hermeneutical ignorance, epistemic violence, and contributory injustice.

If the secondary harms associated with testimonial injustice and related phenomena resemble the primary harm which—according to (Dilemmatic Gaslighting)—gaslighting is apt to cause, does it follow that my claim that this harm is characteristic of gaslighting is false? No, for two reasons. First, the claim that a harm is characteristic of gaslighting is not as strong as the claim that it can occur only as a result of gaslighting. It is rather like the claim that a certain set of symptoms is characteristic of a particular disease: when the disease is present, the symptoms are often there; when the symptoms are there, the disease is often present.

Second, the fact that the secondary harms associated with testimonial injustice and related phenomena closely resemble the primary harm caused by gaslighting may suggest that, under certain circumstances, these phenomena can *constitute* gaslighting. Consider the case of testimonial injustice: under certain circumstances, persistent testimonial injustice can in fact constitute gaslighting as it is understood in (Dilemmatic Gaslighting), since treating a person's testimony as unworthy of credence can be a way of intentionally communicating to that person that they lack basic epistemic competence in the domain relevant to their testimony. Similar remarks apply to other audience-centered notions.

We can even plausibly understand the secondary harm associated with hermeneutical injustice in terms of gaslighting. In her discussion of this topic, Fricker remarks that "When you find yourself in a situation in which you seem to be the only one to feel the dissonance between received understanding and your own intimated sense of a given experience, it tends to knock your faith in your own ability to make sense of the world, or at least the relevant region of the world" (2007, 163). Here Fricker is imagining not just a lack of understanding, but an active dissonance between one's own understanding and "received understanding." It is easy to imagine that this dissonance manifests itself in interactions with dominantly situated epistemic agents who dismiss the victim's attempts to reject the received understanding, thereby communicating to her that she lacks basic epistemic competence when it comes to interpreting her own experience. An exchange of this sort might well constitute gaslighting according to (Dilemmatic Gaslighting). So it

does not seem to me that the similarities between the primary harm gaslighting is apt to cause and the secondary harms caused by testimonial injustice and related phenomena undermine my claim that the harm in question is characteristic of gaslighting.

It is worth noting that, in focusing on the characteristic harm gaslighting is apt to cause, I do not mean to suggest that gaslighting does not also cause other harms. As Podosky (2021) argues, gaslighting which calls into question the accuracy of the victim's concepts (*second-order gaslighting*) can cause additional harms including *discriminatory metalinguistic deprivation*, in which an individual is unjustly denied the capacity to help determine which concept is expressed by a word, and *conceptual obscuration*, in which an individual loses knowledge as a result of being pressured by her gaslighter to adopt a less accurate concept.[19,20]

So much for the nature of the primary harm characteristic of gaslighting as (Dilemmatic Gaslighting) understands it. I turn now to the questions of what it means to claim, as I do, that gaslighting is *apt* to cause this harm, and of why it is always or nearly always morally wrong to gaslight.

To say that gaslighting is apt to harm its victims by causing them falsely or unjustifiedly to believe they are epistemically defective is not to say that it *invariably* leads victims to doubt themselves in this way. The claim is rather that gaslighting makes it *substantially more likely* that they will. (Dilemmatic Gaslighting)'s requirement that the victim assign the perpetrator's testimony significant weight plays an important explanatory role here: the victim of gaslighting cannot nonchalantly choose to reject the perpetrator's testimony, since she takes seriously the possibility that he is correct and she is epistemically defective.

On my view, then, gaslighting is apt to cause harm in much the same way as a physician's urging her patient to undergo chemotherapy would be apt to cause harm if the physician lacked medical justification for suggesting that treatment. Such a course of action on the part of a physician would make it substantially more likely that her patient would experience the harm of undergoing the acute and chronic physical and mental consequences of chemotherapy without thereby curing, preventing, or delaying any more serious condition. But the physician's urging would not *inevitably* cause the harm—patients are under no obligation to act in accordance with their physicians' recommendations. It is rather that, given that patients generally assign their physicians' testimony about the best medical course

[19] Podosky also introduces the term *perspectival subversion* to describe cases in which "subjects of gaslighting are targets of persistent conceptual challenges over time such that they come to doubt their ability to make conceptual judgments" (2021, 223). This harms strikes me as so similar to the primary harm characteristic of gaslighting that it should not be included in a list of secondary harms.

[20] Thanks to an anonymous referee for pressing me to clarify the relationship between the characteristic harm of gaslighting and the harms caused by other varieties of epistemic injustice.

of action significant weight, a patient in this situation would be quite likely to proceed with the treatment and experience the harm.[21]

That gaslighting is apt to cause its characteristic epistemic harm explains why we take it, considered as an act-type, to be always or nearly always morally wrong, and why we take victims of gaslighting to have been wronged. Given the harm it is apt to cause, the wrongness of gaslighting is perhaps clearest when it is done intentionally, since in such cases the gaslighter deliberately attempts radically and unjustifiedly to undermine his victim's belief in her basic epistemic competence.[22] But I think it is clear that gaslighting as it is understood in (Dilemmatic Gaslighting) can be wrong even in unintentional cases. Return to the medical analogy. Unintentionally urging a patient to undergo chemotherapy without medical justification is negligent: it reveals that a physician has failed in her duty to be responsive to evidence and reason carefully about what is in the best interest of her patient. It is wrong for a physician to act negligently towards a patient, and a patient who is treated negligently by her physician is thereby wronged. In much the same way, the unintentional gaslighter who lacks knowledge-level doxastic justification for the proposition he asserts and for the proposition that his victim lacks basic epistemic competence fails in his moral and epistemic duty to be responsive to evidence and reason carefully about what is in the best interest of his interlocutor. In such cases—which constitute the vast majority of cases of unintentional gaslighting—it is wrong to unintentionally gaslight, and a victim of unintentional gaslighting has been wronged. So (Dilemmatic Gaslighting) is able to explain why both intentional and unintentional gaslighting are held to be morally wrong.[23,24]

[21] This medical analogy also helps to clarify the sense in which the harm associated with gaslighting according to (Dilemmatic Gaslighting) is characteristic, given that one can face the choice between trusting someone's testimony and coming to doubt one's basic epistemic competence without experiencing gaslighting (as when the testifier correctly and justifiedly believes that the addressee lacks basic epistemic competence). In the medical case, the patient might experience the very same choice of whether to act on the physician's recommendation in a case where the physician did have a compelling medical reason to suggest chemotherapy. But in such a case, if the patient acted on the physician's recommendation, we would not think of her as being harmed by the treatment. In just the same way, if an individual faces the choice between trusting someone's testimony and coming to doubt her basic epistemic competence, but the individual whose testimony has presented her with the dilemma believes correctly and with justification that she lacks basic epistemic competence, then her coming to believe that she lacks basic epistemic competence is not an epistemic harm: it is forming a justified and true belief. Doubting one's own basic epistemic competence in a domain harms one only if this is not the epistemically responsible thing to do.

[22] Abramson (2014) contains a helpful discussion of the ways in which intentional gaslighting is wrong.

[23] Matters are somewhat more complicated in cases where the gaslighter has knowledge-level doxastic justification for his beliefs (though they are incorrect). Even in such cases, however, I think there is a sense in which the gaslighter has acted wrongly and his victim has been wronged. Compare, for example, a case in which a physician believes that her patient has a life-threatening bacterial infection in her leg on the basis of a usually reliable diagnostic test which is subject to an uncharacteristic error, and she decides to amputate. Even if the physician is responding appropriately to her evidence, given that the amputation is not in fact necessary, there is a sense in which she is acting wrongly in amputating her patient's leg, and there is a sense in which her patient has been wronged by having her leg amputated. Her action may be *excusable*, but this does not make it *right*. In just the same way, I think that when a gaslighter has appropriate justification for his beliefs, his action might be excusable though it is wrong and harmful. In

If we follow Fricker (2007, 1) in understanding epistemic injustice as "[any] wrong done to someone specifically in their capacity as a knower," it follows from the discussion so far that gaslighting is a form of epistemic injustice: the victim of gaslighting is wronged in that she is forced into a situation which increases the probability that she will form false or unjustified negative beliefs about her own basic epistemic competence. It should be emphasized, however, that the relationship between gaslighting and epistemic injustice according to (Dilemmatic Gaslighting) differs from that relationship as it is understood by Ivy (2017) and Podosky (2021): Ivy holds that gaslighting constitutes testimonial injustice specifically, while Podosky holds that epistemic injustice is implicated in gaslighting because it explains why victims of unintentional gaslighting come to doubt their interpretative ability. (Dilemmatic Gaslighting) allows us to understand the relationship between gaslighting and epistemic injustice while avoiding the objections to Ivy's and Podosky's views discussed in Sect. 3.

6 Clarifying the account

A few further notes. First, it bears reiterating that I believe those who have previously worked on gaslighting are correct in holding that it is especially theoretically interesting when it is either done intentionally (in which case there are questions to ask about the psychology of habitual gaslighters, many of which are addressed by Abramson) or tracks gradients of social power (in which case there are questions to ask about how it contributes to the reinforcement of these gradients). But I think that it is an advantage of (Dilemmatic Gaslighting) that it understands gaslighting in isolation from issues of abuse and marginalization. For it is only by understanding gaslighting independently of these issues that we can hope to *explain* why it is so commonly bound up with them. On the one hand, because the characteristic harm gaslighting is apt to cause to its victim simultaneously undermines her as an epistemic agent and makes her dependent on the testimony of the perpetrator, gaslighting is an effective way for the latter to establish control over the former. On the other hand, gaslighting is most likely to be successful (that is, the

Footnote 23 continued
any case, it is worth remembering that very few cases of gaslighting involve gaslighters who have appropriate justification for their beliefs. Thanks to an anonymous referee for pressing me to clarify why unintentional gaslighting is morally wrong.

[24] In her discussion of testimonial injustice, Fricker (2007, 22) argues that genuine testimonial injustice cannot result from an ethically innocent but epistemically culpable error, for example forming an unjustified and false view about a testifier's credibility on the basis of a careless web search. This is because "an ethically non-culpable mistake cannot undermine or otherwise wrong the speaker" (2007, 22). Fricker's position here may make sense for audience-centered concepts like testimonial injustice. When it comes to gaslighting, however, it seems clear that the victim of gaslighting can be wronged and undermined even if the perpetrator's error is epistemic rather than ethical. It follows that even unintentional gaslighters can be culpable. Thanks to an anonymous referee for pressing me to clarify this point.

victim is most likely to accept the perpetrator's testimony) when the perpetrator occupies a position of social power vis-à-vis the victim.

Second, because it does not specify anything about the relationship between victim and perpetrator, and because it allows single instances of communication to constitute gaslighting, (Dilemmatic Gaslighting) does not predict that gaslighting is essentially a cumulative or diachronic phenomenon, or that it must occur between people who have a pre-existing relationship of love or trust. Again, I do not think these facts count significantly against (Dilemmatic Gaslighting). Indeed, some intuitive cases of gaslighting lack both features. Abramson, for example, includes in her list of examples of gaslighting one in which an undergraduate student on a panel discussion argues that her institution needs to continue to work to address racism and is told afterwards by members of the audience, "Don't be crazy," "You're being a little sensitive," and "You made the panel really uncomfortable" (Abramson 2014, 4). Here there need not be any pre-existing relationship at all between the victim and the members of the audience, and, while Abramson's case involves the cumulative effects of three different exchanges, the intuition that gaslighting is occurring is no weaker if we instead imagine just one ("Don't be crazy").

Even if it does not give us reason to append further necessary conditions to our account of gaslighting, however, the fact that [Central Case] and others involve close relationships between victim and perpetrator and diachronically extended patterns of interaction suggests that our account of gaslighting ought to be able to explain why these characteristics should be present in many of the most intuitively forceful examples of gaslighting. With respect to the first observation, the friend of (Dilemmatic Gaslighting) can offer the thought that a pre-existing relationship of love or trust between victim and perpetrator is likely to lead the victim to assign greater credibility to the perpetrator's testimony, thus simultaneously ensuring that condition (iv) is satisfied (so that gaslighting does indeed occur) and increasing the probability that the victim will ultimately accept the perpetrator's testimony and suffer the characteristic epistemic harm caused by gaslighting. With respect to the second observation, the friend of (Dilemmatic Gaslighting) can note that the epistemic harm associated with gaslighting is cumulative, so that extended patterns of gaslighting are liable to undermine victims more than isolated episodes.

Third, though I have formulated (Dilemmatic Gaslighting) so that it appeals to *being in a position to know* and *knowledge-level doxastic justification*, I do not wish to nail my flag to any particular relationship between the concept of gaslighting and the concept of knowledge. It may turn out that the most theoretically fruitful concept of gaslighting will appeal to having and lacking *significant* doxastic justification rather than to being in a position to know and lacking knowledge-level doxastic justification. I would regard this outcome as a vindication of (Dilemmatic Gaslighting) rather than a refutation of it.[25]

[25] For an account of gaslighting stated in terms of justification simpliciter rather than knowledge-level justification, see Stark (2019).

Finally, I have said nothing so far about what constitutes basic epistemic competence or what kind of propositions are such that, if they are true, agents have reason to believe that they lack it. I regard it as a strength of (Dilemmatic Gaslighting) that it does not pronounce on the nature of our basic epistemic competence, since this is a substantive question in epistemology. There are central cases of basic epistemic competence, such as our competence to form everyday perceptual beliefs reliably, and these correspond to central cases of gaslighting. But there could in principle be gaslighting about claims in many domains. I return to this issue briefly in the conclusion.

It may help to preclude certain objections to (Dilemmatic Gaslighting), however, to note that there are some domains in which our beliefs are not plausibly regarded as formed on the basis of any basic epistemic competence. First, there are beliefs about theoretical domains like advanced mathematics, the natural and social sciences, philosophy, and so forth. If you demonstrate that I have made some mistake in a complex calculation involving the physics of lasers, I do not thereby gain a reason to doubt any basic epistemic competence of mine. The same can be said about most areas of philosophy. I am not revealed to lack basic epistemic competence in any domain if my theory of the parthood relation is shown to entail some non-obvious contradiction, or if I have the wrong view about whether any two states of affairs must be such that one is at least as good as the other. This is not to say that theoreticians are always immune to charges of gaslighting: there may be cases in which conclusions drawn on the basis of theory do call into question the basic epistemic competence of individuals, and these cases may constitute gaslighting if they satisfy further conditions. But most theoretical disagreements are not of this kind.

Second, there are beliefs which, while they do not belong to theoretical domains, are formed on the basis of evidence which is subtle or otherwise difficult to interpret. Suppose our colleague Professor Plum gives us a cryptic smirk at the department colloquium. You think he means to indicate that he has a devastating objection to the speaker's theory; I think he means to indicate that he has once again succeeded in pilfering one of the bottles of wine meant for the reception. Even if your belief is correct, my insisting on my wine hypothesis does not call into question any basic epistemic competence of yours—Plum's smirk was, after all, cryptic.

Finally, there are disagreements over matters of taste. If you believe that "Window of Appearances" from Philip Glass's opera *Akhnaten* is the greatest musical work of the past hundred years and I believe that that distinction belongs to "Like a Surgeon" from "Weird Al" Yankovic's album *Dare to Be Stupid*, our disagreements—heated though they may be—do not constitute gaslighting. If there is such a thing as objectively bad taste, to have it is not to lack basic *epistemic* competence in any domain.

7 Overgeneration

Does (Dilemmatic Gaslighting) overgenerate, classifying too many cases as cases of gaslighting?

Consider first a case involving an awkward social blunder:

> [NAME TAG]: You are famed space adventurer Barbarella, and it says so on your name tag. Nevertheless, space knave Ruprecht approaches you at the space conference mixer and greets you by saying, "Ah, Arabella, it's so nice to finally meet you." It is obvious to you that he has not looked at your name tag and has mistaken you for someone else.

In [NAME TAG], Ruprecht intentionally communicates to Barbarella that she is Arabella by addressing her with a referential expression ('Arabella') which is felicitous only if she is Arabella. Barbarella knows that if what Ruprecht communicates is true, she must lack basic epistemic competence in forming correct beliefs about her identity. Moreover, Ruprecht is in a position to know this, since if he simply glances at Barbarella's name tag, he will learn that she is not Arabella and therefore that what he communicates to her, if accepted, would give her reason to doubt her basic epistemic competence. And Ruprecht lacks knowledge-level justification for both the claim that Barbarella is Arabella and the claim that she lacks basic epistemic competence.

So [NAME TAG] meets many of the conditions for being an instance of gaslighting according to (Dilemmatic Gaslighting). But it fails to meet the final condition, for we cannot but imagine that Barbarella, confident in her belief about who she is and recognizing that this is the kind of awkward social blunder which is liable to happen at space conference mixers, will not assign significant weight to Ruprecht's testimony. This point generalizes to a range of cases: when a speaker communicates a proposition which calls into question a hearer's basic epistemic competence, but it is clear to the hearer that the speaker does so only because he has made some common and trivial epistemic mistake, the exchange will fail to be gaslighting in virtue of failing to satisfy the final condition of (Dilemmatic Gaslighting).

What about cases of disagreement involving perceptual or inferential justification? Consider:

> [BIRD]: You are gazing out the window with your veterinary acupuncturist. Both of you see a bird alight on a branch and then fly away. It seems to you that the bird was red. Your veterinary acupuncturist, however, asserts that it was brown. In fact, the bird was red.
>
> [BILL]: You and your veterinary acupuncturist are calculating the tip for lunch. You agree that the tip should be 25% of the total on the bill. You do your mental arithmetic and form a belief about the amount of the tip. Your veterinary acupuncturist, however, announces a number for the tip which is different than yours. In fact, your calculation is correct.

Because whichever party gets things wrong in such disagreements (here the veterinary acupuncturist) does not correctly believe either the proposition they

assert or the proposition that their interlocutor lacks basic epistemic competence in the relevant domain, it might at first seem that (Dilemmatic Gaslighting) predicts that cases like these usually or always involve gaslighting. But this conclusion about (Dilemmatic Gaslighting) is premature. Three points are relevant here.

First, if your veterinary acupuncturist speaks first without knowing whether you agree or disagree, the cases are not correctly characterized as gaslighting, since he is not in a position to know that his testimony will give you reason to doubt your basic epistemic competence in any domain.

Second, if conditions are not ideal for determining the colors of birds (if there is fog, or if the bird is far away, or if it is in near-constant motion), or if the bill calculation is not straightforward, then getting things wrong will not call into question one's *basic* epistemic competence in any domain. So in order to make [Bird] and [Bill] cases of gaslighting according to (Dilemmatic Gaslighting), we must imagine that the bird is large, the branch is very near the window, and the lighting is ideal, or, in the case of the tip, that the calculation is very simple (as it would be, for example, if the bill was for exactly $100).

Third, even in cases where the disagreement plausibly involves a belief formed using a basic epistemic competence, your veterinary acupuncturist's testimony can only constitute gaslighting if you assign significant weight to it. Though matters are not as clear-cut as in [Name Tag] because your disagreement cannot be explained away by appealing to any common or epistemically trivial mistake, it is by no means obvious that you will do this. If you have correctly calculated that the tip on the $100 bill is $25, and he claims that it is in fact $35, you might well respond with the incredulous stare rather than coming to doubt your basic epistemic competence vis-à-vis simple arithmetic.

So we are left with, for example, a version of [Bill] in which you have correctly calculated that the tip on the $100 bill is $25, your veterinary acupuncturist claims (knowing that your answer is $25) that it is in fact $35, and you assign his testimony significant weight, seriously wondering whether you have somehow miscalculated despite the apparent obviousness of the answer.[26] (Dilemmatic Gaslighting) does characterize this unusual sort of case as gaslighting. Yet, when we spell out the details in this way, the verdict is not counterintuitive.

Here is a way to think about cases like these: when two epistemic agents exercise the same basic epistemic competence in some domain to answer the same question and then find that they disagree, they can either suspend judgment concerning the answer, in which case they avoid the risk of gaslighting, or they can stand their epistemic ground. The second strategy is risky: if an agent is in the right, standing her ground preserves her knowledge of the answer and her confidence in her basic epistemic competence and, if he does not also stand his ground, potentially helps her interlocutor come to grips with whatever defect in his basic epistemic competence

[26] Why might your veterinary acupuncturist make this claim even after learning that you have calculated $25? Perhaps he endorses steadfasting in cases of peer disagreement and is misguidedly attempting to share what he takes to be his knowledge with you.

led him to the wrong answer.[27] If she is in the wrong, on the other hand, standing her ground precludes her from recognizing the defect in her own basic epistemic competence and, if he does not also stand his ground, potentially leads her interlocutor to believe falsely that he is a defective epistemic agent. But to lead someone to believe falsely that they are a defective epistemic agent just is to cause the harm characteristic of gaslighting, and so it is no objection to (Dilemmatic Gaslighting) that it characterizes cases with this structure as cases of gaslighting.[28]

It is worth emphasizing, however, that very few cases of gaslighting involve direct competition of this sort between the basic epistemic competences of two agents. Instead, most cases of gaslighting challenge the basic epistemic competence of the victim on the basis of a proposition which is not believed at all (as in [Central Case]) or believed on some basis other than the exercise of a basic epistemic competence (as in [Skeptical Peers]). Whatever sympathy we might feel in [Bill] for your veterinary acupuncturist, who, though he gaslights you, at least does so out of deference to the (flawed) deliverances of his basic epistemic faculties, should not extend to the perpetrators in other cases.

I have claimed that the verdict that your veterinary acupuncturist is gaslighting you in certain versions of [Bird] and [Bill] is not counterintuitive. This is not to claim that it is particularly intuitive—in fact, intuitions about strange and complex cases of this kind strike me as quite weak. For this reason, our verdicts about such cases should plausibly be determined by our best theory of gaslighting: to the victor go the spoils. There is thus an important dialectical difference between cases like [Bird] and [Bill] and cases like [Skeptical Peers]. Whereas existing accounts' difficulties with capturing the intuition that certain versions of [Skeptical Peers] involve gaslighting give us reason to hope for an account which does better, the fact that (Dilemmatic Gaslighting) classifies certain versions of [Bird] and [Bill] as gaslighting does not indicate that it struggles to capture our intuitions in the same way.[29]

A final question concerns various forms of skepticism in epistemology. On the assumption that those who defend skepticism lack knowledge-level justification for their views, does their testimony to the effect that skepticism is true constitute

[27] Here I am in agreement with epistemologists like Lackey (2010a, b), who holds that in disagreements over simple arithmetical questions the agent who knows the answer does not have a reason to reduce her confidence.

[28] It is worth pointing out in this connection that other accounts of gaslighting make similar predictions. For example: if, convinced of the correctness of his answer, your veterinary acupuncturist in [Bill] intends (somewhat paternalistically) to gently convey to you that you lack basic epistemic competence in arithmetic, then he is gaslighting you according to Podosky's account. Similarly, if he maintains his confidence in his answer after learning that you disagree in part because of identity prejudice, then he is gaslighting you according to Ivy's account. So (Dilemmatic Gaslighting) is not alone in predicting that gaslighting can occur during disagreements involving perceptual or inferential justification in which both parties take themselves to be correct.

[29] Some readers might disagree with me here, finding (Dilemmatic Gaslighting)'s verdicts about [Bird] and [Bill] quite counterintuitive. To them I once again offer the olive branch of conceptual pluralism about gaslighting. Perhaps we need more than one concept of gaslighting to account for the full range of our intuitions about cases. Thanks to an anonymous referee for pressing me to clarify the dialectical significance of [Bill] and related cases.

gaslighting? For concreteness, suppose our colleague Professor Plum tells one of his undergraduate students that she does not know whether she has hands because for all she knows she is a brain in a vat. Can we reasonably criticize him for gaslighting?

I do not think (Dilemmatic Gaslighting) predicts that skeptics are gaslighters. For what the external-world skeptic is doing is offering an especially modest proposal about what human basic epistemic competence amounts to: human basic epistemic competence does not extend to such exotic propositions as that one has hands. It follows that Professor Plum does not gaslight, for neither he nor his student knows or is in a position to know that if what he says is true, she lacks basic epistemic competence; rather, both know that if what he says is true, her basic epistemic competence (like everyone else's) does not suffice for knowledge of the proposition that she has hands.

8 Conclusion

I have argued that an account of gaslighting should be judged according to the twin criteria of faithfulness to our intuitive verdicts about cases and theoretical utility. With respect to the first of these criteria, I have highlighted the difficulties faced by existing characterizations of gaslighting in capturing our intuitions about cases of unintentional gaslighting and cases that do not involve identity prejudice or pre-existing systematic epistemic injustice, and I have emphasized that (Dilemmatic Gaslighting) succeeds where they do not.

I would like to conclude by saying a bit more about (Dilemmatic Gaslighting) and the second criterion. What makes (Dilemmatic Gaslighting) a theoretically useful account of gaslighting? Here I will make three points.

First, as I have argued in Sect. 5, (Dilemmatic Gaslighting) makes it clear why gaslighting is apt to cause a characteristic sort of epistemic harm, and therefore why we regard it, considered as an act-type, as always or nearly always morally wrong. To my mind, explaining the connection between gaslighting and harm is one of the central theoretical goals of an account of gaslighting. Surprisingly, however, because they fail to incorporate conditions like condition (iii) of (Dilemmatic Gaslighting), some existing accounts struggle to explain how gaslighting is connected to harm. For example, Podosky's (2021) disjunctive account, which individuates gaslighting either in terms of the intentions of the gaslighter or in terms of the effects of epistemic injustice on his victim, predicts that one agent can gaslight another by telling her a *true* proposition about her epistemic competence as long as he intends her to form a negative attitude toward her interpretive abilities. I think it is far from clear that in all such cases the victim has been harmed epistemically or otherwise—if our account of gaslighting predicts that a psychiatrist gaslights her patient when she correctly diagnoses him with schizophrenic hallucinations, then it also predicts that there is no constitutive connection between gaslighting and harm.

Second, (Dilemmatic Gaslighting) makes it easier than other accounts to determine whether gaslighting has occurred in real-world situations. While we can

often do little more than speculate about others' intentions or hidden prejudices, it is often quite clear what they are in a position to know or believe with knowledge-level justification. Thus, while intentionalist accounts of gaslighting require us to know about the motivational structure of individuals in order to conclude that they are gaslighters, and while anti-intentionalist accounts that tie gaslighting to testimonial injustice similarly require us to know about their inner prejudicial attitudes, (Dilemmatic Gaslighting) asks only what they are in a position to know and what they believe correctly and with knowledge-level justification. For example, consider Podosky's case of inappropriate workplace touching, which was discussed in Sect. 2. On an intentionalist account of gaslighting, it is not at all clear whether the victim's male colleague is gaslighting her—to find out, we would need to know facts about his inner life about which he might remain silent or lie. It follows that, if gaslighting is understood along intentionalist lines, it would be quite difficult in practice to hold him accountable for gaslighting. A similar point can be made about proposals, like Ivy's, according to which gaslighting requires that the perpetrator harbor identity-prejudicial attitudes. Even on Podosky's account, we must somehow decide whether the victim's doubt about whether her belief in the inappropriate touching is justified is causally explained by systematic epistemic injustice, or whether it is produced by some other causal process.

(Dilemmatic Gaslighting) asks much simpler and more straightforward questions: Did the gaslighter correctly and with knowledge-level justification believe that the victim was not touched inappropriately? Did he correctly and with knowledge-level justification believe that the victim lacked basic epistemic competence in forming beliefs about how she was touched? Was he in a position to know that his dismissal of the victim's interpretation would give her good reason to doubt her basic epistemic competence? Given that the gaslighter was not present when the event happened, in the absence of some special information such as that the victim is prone to hallucinations, the answers to these questions seems quite clearly to be *no*, *no*, and *yes*: gaslighting has occurred in this case. So if one of our goals in developing a theory of gaslighting is to successfully hold actual gaslighters accountable for their actions, it seems to me that (Dilemmatic Gaslighting) is more useful than other accounts.

Third, in showing how gaslighting can be apt to cause significant harm even when it is not understood in terms of the intentions or prejudices of the perpetrator, (Dilemmatic Gaslighting) opens up the possibility that the distinctive dilemmatic structure of gaslighting might be found in surprising places. It is often assumed, for example, that because academic discussions of questions related to socially significant topics like race, gender, disability, and sexual ethics are abstract, and because they are conducted in a spirit of inquiry rather than with the intention of undermining or manipulating anyone, it is not possible for the claims made in them to constitute gaslighting. According to (Dilemmatic Gaslighting), this sort of reasoning is a non sequitur: gaslighting can occur in any domain where beliefs are formed on the basis of a basic epistemic competence, and it may well be that certain parties in these discussions form their beliefs about, for example, what it is like to be disabled, or whether it is good for them to be involved in a queer romantic relationship, on the basis of their basic epistemic competence. It is well beyond the

scope of the present discussion to argue that gaslighting does in fact sometimes occur in these sorts of contexts, but one important theoretical upshot of (Dilemmatic Gaslighting) is that we must consider the possibility that the dilemmatic structure of gaslighting can be found in exchanges quite different from [Central Case]. If it is a mark of the theoretical fruitfulness of an account of a phenomenon that it opens up questions about the structural similarities between paradigm cases of that phenomenon and cases in other domains, the fact that (Dilemmatic Gaslighting) does this is a mark of its success.

If my arguments so far are successful, they show (i) that (Dilemmatic Gaslighting) fares better than existing accounts in terms of capturing our intuitions about which cases involve gaslighting and (ii) that it is a useful account of gaslighting for theoretical purposes. It follows that (Dilemmatic Gaslighting) articulates a concept of gaslighting worth having, whether uniquely or as one of several concepts of gaslighting which prove theoretically useful in different domains.

Acknowledgement I am grateful to Kate Abramson, Emma Atherton, Ben Bronner, Robin Dembroff, Kenny Easwaran, Carolina Flores, Jeremy Goodman, Chris Hauser, Veronica Ivy, Kelsey Laity-D'Agostino, Jonathan Schaffer, Cynthia A. Stark, and two anonymous referees for *Philosophical Studies* for their comments on earlier drafts of this paper. The ideas here also benefitted from discussion at the New York City Minorities and Philosophy Workshop Series in 2019 and the Pacific Meeting of the APA in 2021.

References

Abramson, K. (2014). Turning up the lights on gaslighting. *Philosophical Perspectives, 28*, 1–30.

Cappelen, H. (2018). *Fixing language: An essay on conceptual engineering*. Oxford: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*, 7–19.

Dembroff, R. (2016). What is sexual orientation? *Philosophers' Imprint, 16*(3), 1–27.

Dotson, K. (2011). Tracking epistemic violence, tracking practices of silencing. *Hypatia, 26*, 236–257.

Dotson, K. (2012). A cautionary tale: On limiting epistemic oppression. *Frontiers: A Journal of Women Studies, 33*, 24–47.

Fricker, M. (2007). *Epistemic injustice*. Oxford: Oxford University Press.

Haddock, A., Millar, A., & Pritchard, D. (Eds.). (2010). *Social epistemology*. Oxford: Oxford University Press.

Haslanger, S. (2000). Gender and race: What are they? What do we want them to be? *Noûs, 34*(1), 31–55.

Ivy, V. [McKinnon, R.] (2017). Allies behaving badly: Gaslighting as epistemic injustice. *In Kidd, Medina, and Pohlhaus Jr., 2017,* 67–174.

Ivy, V. [McKinnon, R.] (2019). Gaslighting as epistemic violence: "Allies," mobbing, and complex posttraumatic stress disorder, including a case study of harassment of transgender women in sport. *In Sherman and Goguen, 2019*, 285–302.

Jenkins, K. (2016). Amelioration and inclusion: Gender identity and the concept of *woman*. *Ethics, 126*, 394–421.

Kidd, I. J., Medina, J., & Pohlhaus, G., Jr. (Eds.). (2017). *The Routledge handbook of epistemic injustice*. England: Routledge.

Lackey, J. (2010). A justificationist view of disagreement's epistemic significance. *In Haddock, Millar, and Pritchard, 2010*, 298–325.

Lackey, J. (2010). What should we do when we disagree? *In Szabó Gendler and Hawthorne, 2010*, 274–293.

Podosky, P. M. C. (2021). Gaslighting, first- and second-order. *Hypatia, 36*, 207–227.

Pohlhaus, G. (2012). Relational knowing and epistemic injustice: Toward a theory of willful hermeneutical ignorance. *Hypatia, 27*, 715–735.
Sherman, B. R., & Goguen, S. (Eds.). (2019). *Overcoming epistemic injustice*. New York: Rowman and Littlefield.
Spear, A. D. (2019). Epistemic dimensions of gaslighting: Peer disagreement, self-trust, and epistemic injustice. *Inquiry*. https://doi.org/10.1080/0020174X.2019.1610051
Spear, A. D. (2020). Gaslighting, confabulation, and epistemic innocence. *Topoi, 39*, 229–241.
Stark, C. A. (2019). Gaslighting, misogyny, and psychological oppression. *The Monist, 102*, 221–235.
Szabó Gendler, T., & Hawthorne, J. (Eds.). (2010). *Oxford studies in epistemology* (Vol. 3). Oxford: Oxford University Press.