Canadian Journal of Philosophy (2020), 50: 6, 735–749
doi:10.1017/can.2020.22

ARTICLE

Lit from Within: First-Person Thought and Illusions of Transcendence

Léa Salje

University of Leeds, Leeds
Email: l.c.salje@leeds.ac.uk

Abstract
Philosophical treatments of the self in a range of different traditions have positioned it outside the realm of ordinary worldly objects. This paper argues that part of the explanation for this seemingly widespread and persistent temptation to mystify the self is that the epistemic properties of *I*-thought are apt to give rise to an illusion of transcendence about their objects—that is, about ourselves.

Keywords: The self; first-person thought; *de se*; cognitive illusions; transcendence of the self

> When a child, like Kim, having no theoretical commitments or equipment, first asks himself "Who or What am I?", he does not ask this from a desire to know his own surname, age, sex, nationality, or position in the form. He knows all this ordinary personalia. He feels that there is something else in the background for which his "I" stands, a something which has still be described after all his ordinary personalia have been listed. (Ryle 1994, 32)

Philosophical treatments of the self in a range of different traditions have positioned it outside the realm of ordinary worldly objects. The self—which is to say the *true* self, or the self *as it is in itself*—is not something easily apprehended or described. In the pretheoretical arena, too, there is a tenacious instinct for the idea that we are really, at our most fundamental, something much more exotic and otherworldly than, say, socially situated animals.

These positions are not unmotivated. There are historically influential arguments that support them, religious and cultural frameworks that normalise them, and a number of philosophical attempts to diagnose the underlying drive towards them. In the latter spirit, for instance, Ryle argued that we find ourselves mystifying the self for a somewhat unmysterious reason: self-consciousness involves bearing higher-order states and attitudes toward oneself, resulting in a maddening sense of systematic elusiveness. As he explains, "the more the child tries to put his finger on what 'I' stands for, the less does he succeed in doing so. He can catch only its coat-tails; it itself is always and obdurately a pace ahead of its coat-tails" (1994, 32). More recently, Christopher Peacocke has argued that the "persistent impulse amongst thinkers about the self to postulate a transcendental subject of experience and thought" arises from *representationally independent* uses of the first person—first personal beliefs formed in the absence of underlying first personal representational content (1999, 263). More recently still, Annalisa Coliva, with echoes of Sydney Shoemaker before her, has argued that we should understand this drive towards transcendental construals of the self as yet further part of the intrigue of the phenomenon of *immunity to error through misidentification* attaching to first-person thought (Coliva 2012; Shoemaker 1968).

This paper offers a different diagnostic explanation. In common with the accounts just mentioned, I locate the explanation among the semantico-epistemic features of self-conscious

thought—or thought in which we represent ourselves *as ourselves*. The relevant kind of thought here is first-person thought, or *I*-thought: self-conscious thought of a kind most naturally expressed in language with first person pronouns, and whose pattern of reference is governed by a rule of thinker-reflexive reference. We are inclined to mystify the self, so the common idea between these accounts goes, because of some oddity about the way we have of *thinking* about the self we are each most intimate with—namely, ownself. The specific proposal of this paper is that there are epistemic conditions inherent to episodes of first-person thinking that are apt to give rise to a *cognitive illusion of transcendence about the self* in first-person thinkers. If this is right, then it offers a new explanation for our seemingly pervasive temptation to mystify the self.

The plan for the paper is as follows. The short section that comes next is setup (section 1); in it I clarify the general notion of a cognitive illusion (with examples), set out in more detail the task of characterising a particular cognitive illusion, and say what the content is of the cognitive illusion that will be of interest to us. The central argument of the paper comes in the three sections following the setup. There are two separate epistemic conditions of first-person thought to be identified (section 2 and section 3 respectively), which are then brought together in section 4 for a full characterisation of the proposed illusion. That illusion, I argue, emerges as a way of alleviating the epistemic discomfort that results from being, at one and the same time, in a position of unusual epistemic privilege with respect to one's own *I*-thought, and yet not thereby positioned to know anything substantive about the self it refers to.

1. The cognitive illusion

In this section, I clarify three things: the general notion of a cognitive illusion, the specific task ahead of explaining a given cognitive illusion, and the content of the cognitive illusion that pertains to the stated task of this paper.

The notion of a cognitive illusion is most clearly brought out by reference to its more familiar counterpart, the *perceptual* illusion. On a standard view, to say that a perceptual experience is illusory is to say that the subject misperceives one or more of the properties of a mind-independent target of a perceptual experience; the object is perceived, but it is perceived as having properties other than it in fact has.[1] Such misperceptions can be caused by diverse features of the conditions in which the perceptual experience is had. To undergo the classic bent-stick-in-water illusion, for example, is to have a perceptual experience in which a straight stick is misperceived as bent because of facts about the respective refractive properties of water and air. The illusion is systematic inasmuch as it can be reliably recreated by recreating those perceptual conditions.

To say that a cognitive episode (call it a *thought*) is illusory is to say that the object of thought is represented as having properties other than it really has, and that this misrepresentation is an arational result of the conditions in which the thought is had. (It's important that the misrepresentation is caused arationally—e.g., not as the result of an inference—because without this restriction *all* false beliefs would count as cognitive illusions.) An example of a cognitive illusion is the *belief perseverance bias*, in which a subject, whose original grounds for a given belief have been discredited, nevertheless perseveres in believing it.[2] While it's not entirely clear why the illusion arises when it does, the subject's initial impressions of the reliability of the original grounds apparently leave the belief-content with a lasting stamp of epistemic credibility *even after those grounds have been rationally undermined*. In this case, then, the illusion's target—the belief-

[1]See, e.g., Macpherson and Batty (2016, 264); illusions are standardly contrasted with *veridical perceptions* on one hand (in which a mind-independent object is accurately perceived as having only the properties it in fact has), and *hallucinations* on the other (in which there is no mind-independent object perceived at all, but the experience is subjectively similar to an experience of veridical perception).

[2]Ross, Lepper, and Hubbard (1975).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

content—appears to be epistemically credible even after the subject has reason to believe otherwise, and this misrepresentation seems to be caused otherwise than by a rational process.[3]

Illusions, both perceptual and cognitive, are conscious and personal-level states. A perceptual illusion requires the subject to consciously perceive the object; a cognitive illusion requires consciously thinking about it.

Now, if the task of this paper was to characterise a given perceptual illusion, then the aim would be to identify the conditions in which the illusion systematically occurs and to say something about why it is apt to occur in those conditions. Relevant to the task would be mention of such factors as the health and functioning of the perceptual organ, the acoustic or lighting conditions and their like, the conduction properties of the medium through which the sound or light waves must travel, and so on. That, of course, isn't our task; the task of this paper is to characterise a certain *cognitive* illusion. But in much the same way, the aim is to identify the conditions in which the illusion systematically occurs, and to say something about why it is apt to occur in those conditions.

There are at least two more similarities between cognitive and perceptual illusions worth mentioning (though there are, no doubt, also many important differences). First, in the perceptual case, one can readily undergo an illusion without forming the corresponding judgment. I can visually experience the stick as bent without judging it to be so. The same is true in the cognitive case. To give another example, a well-known cognitive illusion arises in connection with the familiar *anchoring* effect, in which exposure to an initial number serves as an influential reference point (or "anchor") for subsequent judgments formed under uncertainty about numerical values.[4] Even after learning about the anchoring effect, it is easy to feel a pull toward, and yet resist, the judgment that one's answer has been formed free of influence from an initial anchor. In this case, one undergoes a cognitive illusion—as of the number not being causally influenced by an initial reference point—but one knows well enough not to go on to form the corresponding judgment. This judgment-independence of cognitive illusions is key to the task of this paper, which is not to identify premises used by thinkers about the self to argue their way into a considered judgment about the transcendence of the self. Rather, the task is to show that just as anyone in the right perceptual conditions, whether or not they know it, will be apt to undergo an illusory experience of the stick as bent, any thinker of a first-person thought in the identified conditions will be apt to undergo a cognitive illusion of transcendence about the self. And the aim of the paper certainly isn't to recommend the corresponding judgment of transcendence about the self.

The second similarity worth highlighting is that neither kind of illusion provides a way of finding out what the illusory state's object is really like. The most obvious reason for this is that taking illusions at face value will normally lead to faulty judgments. But even once the illusion has been unveiled and the corresponding judgment resisted, it bears emphasising that being aware of the illusion does not contribute positively to what is known about its object. Becoming wise to the fact that I am undergoing a perceptual illusion as of a bent stick does not produce any new evidence about the stick's shape properties.[5] Correspondingly, establishing that I am undergoing a cognitive illusion as of a transcendental self does not provide me with any new positive evidence about my self's real nature. The proposal of this paper concerns a systematic distortion in how the intentional targets of *I*-thoughts cognitively appear to their thinkers. In this, it says nothing positive about what a clear-sighted account of the nature of the self would be.

[3]There are good questions about whether how exactly to understand the relevant notion of (mis)representation here, whose answers will largely depend on theoretical preferences concerning the metaphysics of thought that are orthogonal to the present discussion.

[4]Tversky and Kahneman (1974). For example: subjects were asked to spin a rigged "wheel of fortune" before guessing what percentage of African countries are in the U.N. The wheel always landed on either 10 or 65. The average answer for subjects whose wheel spin landed on 10 was 25%; for those who landed on 65 it was 45% (Kahneman 2012, 119).

[5]Using language from Jim Pryor (1999), this is because finding out one is undergoing an illusion only gives one *undercutting* evidence, whereas to find out something new about the object would require *additive* evidence.

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

Isn't there an exception to this general point—that is, doesn't undergoing the bent-stick-in-water illusion at least provide me with grounds to think that the stick is straight, and doesn't undergoing an illusion of transcendence about the self give me grounds to think that the self is *not* transcendent? This challenge raises a wrinkle in the above characterisation of the general notion of an illusion that needs addressing. Both standard formulations of the notion of an illusion and archetypal illustrative cases typically involve error: the object is (perceptually or cognitively) represented as having one or more properties it doesn't really have. All the examples above were of this kind. Notice, however, that there is nothing in principle to rule out cases in which the conditions of perception or thinking cause there to be more than one misrepresentational effect whose combination results in a personal-level conscious representation that happens to be accurate. Consider, for instance, a perceptual case in which I have severe jaundice and am looking at a white wall. Without adding any further details to the case, I would see this white wall as yellow. Suppose, however, that we do add a further detail: the lighting conditions are atypical such that to a nonjaundiced viewer the wall would appear pink, but to a jaundiced viewer it would appear white (indeed, the very shade of white a nonjaundiced viewer would see in ordinary lighting conditions). My visual experience in this case involves not one but *two* levels of misrepresentation—one caused by my jaundice, the other by the lighting conditions. But it does not involve an error in the properties I visually represent the wall as having. The wall is white, and I see it as white. A parallel case involving a luckily accurate cognitive illusion can be devised using the above case of belief-perseverance bias. Imagine that from the subject's perspective the case runs as described above, but it turns out that the original grounds for the belief *really were* in fact reliable, and the later undercutting grounds were faulty. In this version of the case, the persevering believer would not be making a mistake—the belief content really is epistemically credible, and she represents it as such. But this accuracy is nothing more than a matter of luck resulting from one misrepresentation (i.e., of the belief as credible in the face of contrary evidence) being corrected for by another (i.e., of the [in fact, unreliable] undercutting grounds as reliable). Cases of this kind have a structure familiar from discussions of Gettier cases.

We have two options once we recognise the possibility of such cases. The first is to insist that illusion must involve error, and so rebrand these cases as having an interestingly close, but nevertheless distinct, epistemic status. The second is to allow a more inclusive definition of illusion that includes both inaccurate and luckily accurate cases. While the first option might be called for by some theoretical purposes, the differences between these two sorts of case will not be important for ours, so I will leave them undifferentiated under the term "illusion." This strictly calls for the following adjustment to our operational definition of a cognitive illusion.

Cognitive illusion $=_{df}$, a thought whose object's properties are *either*:

(i) inaccurately represented as an arational result of the conditions in which the thought is had; *or*
(ii) accurately represented as an accidental arational result of the conditions in which the thought is had.

Readers who prefer the first option are welcome to reframe the stated purpose of this paper as that of showing that the epistemic conditions of first-person thought are such that they *either* give rise to a cognitive illusion of transcendence, or to the relevant nearby state of accurate representation. In either case, we will have an explanation of the inclination to characterise the self as transcendent that disqualifies it from carrying any evidential force with respect to the question of what the self is really like, and so that can be used to clear the ground for less mysterious views of the self.

That clarifies the general notion of a cognitive illusion and the task ahead. What, finally, is the content of the cognitive illusion that I am going to argue for? The proposal of this paper is that the

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

inherent epistemic conditions of first-person thought are apt to systematically produce an illusion in which the state's intentional object—the self—appears to be *by its very nature substantively unknowable*. This would make it a mysterious object indeed. I use the label *transcendental* as an openly ahistorical shorthand for this property. I take it that establishing that the referent of *I*-thoughts systematically appears to have this property would provide a plausible explanation (at least in part) of the draw towards the kind of mystification of the self as described in the introduction.

2. What we know

In the remainder of this paper, I am going to argue that there are two separately characterisable epistemic conditions inherent to first-person thought which, when undergone together, are liable to produce this cognitive illusion of transcendence of the self. These two conditions are constituted by distinct epistemic perspectives had by first-person thinkers on their own *I*-thoughts. Talk of "epistemic perspective" here is, of course, metaphorical: I use it as shorthand for the range and restrictions on what can be known by a subject in certain circumstances—in the particular case that interests us, the circumstance of being the thinker of an *I*-thought. The first perspective (to be characterised in this section) involves a certain sort of insight into the referent's existence; the second (to be characterised in the next) involves ignorance of the referent's nontrivial properties. In section 4, I will argue that the juxtaposition of these two perspectives is apt to give rise to a cognitive illusion of transcendence of the self because that illusion is a way of relieving the uncomfortable epistemic situation we find ourselves in wherein we seem to have unusual insight into a domain in which, when we look, we find we know very little.

The aim of this section, then, is to propose, motivate, and defend the claim that first-person thinkers have the first of these two epistemic perspectives on their own *I*-thoughts. I propose it as follows:

> *Special Insight*: In virtue of being the thinker of a conscious *I*-thought, a subject has privileged noninferential epistemic grounds for first-personal knowledge that the referent of their thought exists.

Epistemic grounds are sensitive to the hyperintensional content of the beliefs formed on their basis. Special Insight says that a first-person thinker will thereby (i.e., in virtue of being a first-person thinker) have grounds to know that the referent of their thought exists, where that referent is thought of in a first personal way—to know, that is, that *I exist*. It does not say that the first-person thinker will thereby have grounds to know that the referent of their thought exists, where that referent is thought of in any other way.

To see the initial case for Special Insight, it may help to think a first-person thought—the thought *I am reading a philosophy paper*, perhaps. Assuming you've just done this, an episode of conscious *I*-thought just occurred in the room you're in. Of all the people in the room, however, you alone have a distinctive way of knowing that it occurred, and what's more, of knowing that it had a thinker that it successfully referred to (you're it!) that didn't rely on ordinary sensory evidence, inference, or testimony. That's because merely in virtue of thinking the thought, you had a special way of knowing that you were thinking the thought, and so a fortiori, of knowing that you exist. If this correctly captures what just happened, then Special Insight is true.

Here is the more detailed motivating case for Special Insight. When a subject tokens a conscious thought of any kind (say, *the library closes at eight*), she thereby has noninferential epistemic grounds for judging that she, thought of first personally, is thinking it (*I am thinking: the library closes at eight*) that isn't available to others. This is a maximally weak version of a familiar privileged access, or epistemic asymmetry thesis. It is weak because it makes no claims to the self-intimating nature, infallibility, incorrigibility, or even the superiority of the thinker's side of this asymmetry

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

and is largely silent on how best to understand the introspective mechanisms underpinning the epistemic grounds. It says only that there are distinctive grounds for first-personal knowledge that one is the thinker of a given conscious thought that are had *in virtue of being* that thinker, and that those grounds are non-inferential. Just in virtue of being the thinker of a conscious thought I have privileged non-inferential grounds for knowing that I am thinking it. And in having grounds for knowledge that I am thinking a given thought, I thereby also have grounds to know that I exist. So just in virtue of being the thinker of a conscious thought, I have a distinctive non-inferential way of knowing that I exist.

Now, this epistemic asymmetry plausibly holds for *all* conscious thoughts, regardless of their content. When the thought in question is a first-person thought, however, something special happens. That's because for these thoughts in particular—as contrasted, for example, with thoughts about water, or the library, or about LS—there's a coincidence between what I have a distinctive non-inferential way of knowing to exist (me, thought of first personally) and the object of my thought (me, thought of first personally). So in the special case of first-person thought, I am in an exceptionally powerful epistemic position with respect to the object of my thought. I can know, just by thinking a conscious thought of the right kind, that the referent of my thought exists. And that is what Special Insight says.

There are at least four potential objections to Special Insight, whose discussion might help to bring out the distinctive epistemic achievement it involves, and to distinguish it from neighbouring acheivements and guarantees.

The first objection is Lichtenbergian in spirit. There is daylight, the objection goes, between the occurrence of a conscious thinking on the one hand, and the presence of a thinking subject on the other, and it is really only the former that is directly epistemically supported by introspection. So it is not true that we conscious thinkers have privileged grounds for first-personal knowledge that we are thinking conscious thoughts—the most we can say is that we have distinctive grounds for knowledge of *the occurrence of conscious thinkings*. But that wouldn't be enough to count as having distinctive grounds for knowledge that an *I*-thought's referent exists as Special Insight claims for us.

The problem with this style of objection is that it collides head-on with a highly natural way of understanding what a conscious state is that would be hard to give up: viz, that conscious states are states of a subject. Just as certain physical states (*being warm, being tall, having crossed legs*) are ways that a subject can be with respect to its physical properties, conscious states (*aimlessly daydreaming, mentally planning the route home, worrying about Trump*) are ways that a subject can be with respect to its conscious properties. One comes to be in physical states as the result of physical changes—an increase in body temperature, growth, movement of the limbs. Likewise, one comes to be in conscious states by undergoing episodes of active mentation, broadly construed to include both deliberate and nondeliberate changes in conscious properties. (This way of understanding things, note, is silent on the nature of the thinker. It could be just as true of an animal as of a Cartesian ego.) So long as we insist on this dimension of continuity between physical and conscious states—that they are all *states that a subject can be in,* or *ways that a subject can be*—then there really isn't any daylight between the occurrence of an episode of thinking and the presence of a thinker in whom the thinking occurs; to be aware of a given conscious state *just is* to be aware of a way that the subject is with respect to its conscious properties.[6]

It might be responded that the critical question here isn't whether there is *in fact* a potential metaphysical gap between the occurrence of a conscious thought and the presence of a thinking subject—even if it is agreed that there isn't, the real question is whether this gaplessness is *introspectively manifest* to the subject such that the epistemic status accorded to beliefs about the occurrence of a conscious thinking formed on the basis of introspection carries over to

[6]This is an adapted version of a response strategy to Lichtenberg-style objections found in O'Brien (2007, 2015), Coliva (2012, 2017), and, in stronger form, in Peacocke (2012).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

corresponding beliefs about the existence of the thinking subject. If we are serious about conscious states being states of a subject, however, then we should reject this challenge as being unfairly constructed on Lichtenbergian terms. That's because we no longer have reason to *begin* with the idea that introspection supplies grounds for knowledge about the occurrence of conscious thinkings that stand in need of justified expansion if they are to accommodate supplementary singular knowledge about the existence of a thinker. Indeed, given the first-personal language in which it is natural to directly express what one comes to know through introspection, it is much more plausible that introspection supplies grounds for singular knowledge in the first instance (*I am thinking that* p), and existential knowledge about conscious thinkings only derivatively (*there is thinking that* p). So *pace* Lichtenberg, my introspective awareness of occurrent conscious thinking plausibly constitutes distinctive noninferential grounds for first-personal knowledge that the thought's thinker exists.

The refusal to reify conscious states displayed in the last two paragraphs is, of course, unlikely to convince a committed Lichtenbergian. But as is familiar from other sceptical domains, it isn't always appropriate to appease the sceptic on their own terms. When isn't it? In *The Possibility of Knowledge* (2007), Quassim Cassam provides a helpful heuristic for settling this question: to uphold a sceptical worry, one ought to be more certain that the sceptically driven epistemological requirement is one that must be met than one is of the commonsensical epistemological commitment that would be overturned by its failure. If not, the putative epistemological requirement should be abandoned.[7] As he writes, "acceptable epistemological requirements mustn't have unacceptable epistemological consequences" (32)—so wherever we find a trade-off between a seemingly plausible epistemological requirement and a seemingly perverse epistemological consequence, our question must be whether the consequence is strictly unacceptable. For this, he proposes the following approach:

> [I]n any serious investigation of the conditions of knowledge we start off with the idea that there are certain things that we know, or certain kinds of knowledge that we actually have. We regard some of the knowledge that we take ourselves to have as negotiable, and some as more or less non-negotiable. An unacceptable epistemological principle would be undermining of knowledge in the latter category. (32)

Cassam's example of nonnegotiable knowledge is perceptual knowledge that my coffee cup is chipped. My introspective knowledge that I exist is even more obviously a basic case of nonnegotiable knowledge. I am decidedly more certain that I have introspective knowledge of my own existence than I am of the epistemological requirement that I must rule out the sceptical possibility of thinkerless thoughts before I can know it. According to Cassam's rule, then, we ought to give up that putative epistemological requirement.

Even if it does not stand up as an objection, this objection-response pair helps bring out both an important point of contrast and an important point of similarity between Special Insight and the epistemic achievement involved in the *Cogito*. The Cogito is a performative existence proof in which a subject infers her own existence from the self-ascription of an occurrent conscious state. The knowledge in Special Insight, by contrast, is not inferential. What it describes is our privileged noninferential way of knowing about our own conscious states. The point of contrast, then, is that only one of them is inferential. The point of similarity is that both of these epistemic achievements exploit the connection between reference success and referential existence to produce knowledge of the existence of the intentional target of one's first-person thoughts. In both cases, one enacts a special "from the inside" way of knowing that there is an extant referent at the end of one's *I*-thoughts.

[7]Cassam (2007, 31–32); he derives this principle from McDowell 1998, and the approach is clearly Moorean (Moore 1939).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

A second objection is a worry about what to say about thinkers who lack the theoretical inclination to accept the epistemic achievement captured by Special Insight. Ryle (1994), for instance, denied a version of the privileged access thesis. Does it follow that Special Insight is false for him as a special case? It doesn't. Ryle had the epistemic grounds described in Special Insight, even if he was disinclined to rely on them, or was sometimes able to override them by philosophical argument. A dedicated (but misguided) external-world sceptic still has epistemic grounds for existence claims about objects around her in virtue of perceiving them, even if she sometimes manages to prevent herself from making use of them to form the corresponding existence judgments.

Objection three questions the novelty of Special Insight. Isn't this just an overly complicated way of redescribing the familiar phenomenon of *guaranteed reference*, or the thesis that first-person thought is immune from failures of empty reference? It isn't, though the two phenomena are clearly related. Guaranteed reference is a semantic rather than an epistemic property of first-person thought—it concerns only limitations on the ways in which the thought's referential properties can go wrong, not limitations on the thinker's potential knowledge of the existence of the thought's referent. A seed of truth in this objection, however, is that Special Insight is arguably the epistemic face of guaranteed reference: while the guaranteed reference thesis says that *I*-thoughts cannot fail of reference, Special Insight says that the *I*-thinker always has epistemically distinctive grounds to *know* that her conscious *I*-thought successfully refers.

A fourth objection is that Special Insight does not make first-person thought so different from thoughts of other kinds. Let's say that it's right that, in virtue of being their thinker, I can know that the referent of my *I*-thoughts exists. Something comparable is surely also true of our *water*-thoughts. Merely by consciously thinking a thought of the right kind, a suitably conceptually competent subject can know that their *water*-thoughts refer to water (if they refer at all), by redeploying their unitary *water* concept at both levels of the intentional hierarchy—a trick, notice, that can only be pulled off reflexively for a subject's own first order states. As Shoemaker explains:

> [T]he contents of mental states are fixed holistically, [. . .] whatever fixes the content of the first-order belief I express by saying "There is water in the glass" also fixes in the same way the embedded content in the second-order belief I express by saying "I believe there is water in the glass." (1994, 260n7)

There is nevertheless a crucial difference between what Special Insight claims for *I*-thought and this climb up the intentional hierarchy in the case of *water*-thoughts. In the case of a *water*-thought, one can know that if the reference conditions have been met at the first order, then the same reference conditions are met at the second order and beyond. But the claim in the case of an *I*-thought is that its thinker is positioned to know that the first-order reference conditions have been satisfied *tout court.* It's not that I can know that my *I*-thought refers to me *if it refers at all.* Rather, in the case of *I*-thought, I am in a uniquely privileged position to know that my *I*-thought successfully refers—that I, its referent, exist.

Still, an opponent might respond, there's no great difference here with our *water*-thoughts. If causal theories of reference are in good standing, according to which concept acquisition proceeds via appropriate causal contact with the concept's referent, then my very capacity to *use* the *water*-concept makes available a sort of transcendental argument for the existence of its referent. (Roughly, water must exist in my world, because I am a competent *water*-concept user, and that requires me to have had sufficient causal interactions with it.)[8] If that's right, then merely in virtue of being the thinker of a *water*-thought I am in a privileged position to know that the reference conditions on my *water*-thought have been met *tout court*—that water exists—just as I have been claiming for *I*-thoughts.

[8]Putnam (1981, chap. 1); see also Warfield (1997) and (1998).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

But this version of the objection also misses its mark. Even if those arguments are sound (which must surely be in question), what they offer is a way of *inferring* a conclusion about the existence of water in the local world from representational facts about *water*-thoughts. Special Insight, by contrast, claims that we have a noninferential way of knowing "from the inside" that the conditions on one's *I*-thoughts have been met. In this, Special Insight captures something exceptional about the inbuilt epistemic conditions of first-person thought in particular.

That concludes the proposal, motivation, and defence of the claim that *I*-thinkers have this first epistemic perspective on their own *I*-thoughts. I turn now to the second.

3. What we don't know

Special Insight characterises an unusual dimension of epistemic insight had by *I*-thinkers into the referential success of their own first-person thoughts. A next question we might ask is, what can these thinkers know from this position of privilege about what their *I*-thoughts refer *to*? It is in answering this question that the second epistemic perspective had by *I*-thinkers on their own *I*-thoughts comes into view.

A true but unsatisfying answer in my own case is that *I can know that my I-thoughts refer to me.* To know this, however, is not to know very much. As Anscombe writes of the linguistic counterpart to this claim:

> [T]he explanation of the word "I" as "the word which each of us uses to speak of himself" is hardly an explanation!—at least, it is no explanation if that reflexive has in turn to be explain in terms of "I." (1975, 142)

Crucially, it's not to know any nontrivial truths about the referent of my first-person thoughts. I cannot, merely in virtue of thinking a conscious *I*-thought, discover any of my Rylean ordinary personalia—my age, my sex, nationality, my position in the form. Neither can I know whether I'm referring to a whole person or merely to a person time slice, to an animal, or to a Lockean person with psychological survival conditions. I can't thereby know whether my undetached fingernails, my fetus, or my wooden leg get to count as proper parts of my referential target. Perhaps I can know that I am the sort of thing that has the current capacity for conscious thought. But even then, I can't thereby know whether I would be any the less myself without it. Now, none of this is to say that I can't find out about these substantive properties of the referent of my *I*-thoughts in other ways (by philosophical argument, perhaps). But I cannot know them merely in virtue of thinking a conscious *I*-thought. The social, normative, physical, and metaphysical contours of my referent are hidden from view from this otherwise privileged perspective on my own I-thoughts.

This brings us to the second proposed epistemic perspective first-person thinkers have on their own *I*-thoughts:

> *Ordinary Ignorance*: A subject does not, in virtue of being the thinker of a conscious *I*-thought, have privileged noninferential grounds for knowledge about the nontrivial properties of the referent of their thought.

I repeat the first epistemic perspective from section 2, to have them side-by-side:

> *Special Insight*: In virtue of being the thinker of a conscious *I*-thought, a subject has privileged noninferential epistemic grounds for first-personal knowledge that the referent of their thought exists.

Unlike Special Insight, Ordinary Ignorance contains a negative claim stating our epistemic limits "from the inside." That is why it is a claim about ignorance. The ignorance is ordinary because it has nothing special to do with first-person thought—rather, it's a special instance of a general principle:

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

that being a singular-thought thinker of any kind doesn't come with any distinctive epistemic advantage with respect to what the thought's referent is substantively like. Being the thinker of a *water*-thought won't, by itself, help you uncover whether you are thinking about H_2O or *XYZ*. That is an empirical question, answering only to empirical investigation. Likewise, being the thinker of an *I*-thought won't, by itself, help you to uncover whether you are a Cartesian ego, an animal, or something else altogether.

It is noteworthy that even if the form of ignorance captured in Ordinary Ignorance is quite general, the temptation to overlook it seems especially inviting in the case of *I*-thoughts—unlike, say, the case of *water*-thoughts, where we are little inclined to claim privileged insight into the nature of its extension "from the inside." This is importantly like the criticism raised by Kant against the rational psychologists in "The Paralogisms." Making these parallels explicit will help highlight the epistemic limitation captured by Ordinary Ignorance, and to partly forecast its interaction with Special Insight ahead of the next section.[9]

Across the two editions of the first critique, Kant delivers a series of attacks on various attempts to derive substantive knowledge about the self from facts about the representation *I* given in self-consciousness. We cannot, he argues, know anything about the simplicity, the substantiality, or the diachronic identity of the self, its freedom, or its relations to other objects in space merely by reflecting on the *I* of self-consciousness. What a thinker *can* know from the privileged position of self-consciousness is that the target of the self-conscious thought exists; "[t]here can be no question that I am conscious of my representations; these representations and I myself, who have the representations, therefore exist" (A370).[10] But it would be a mistake to think that as the conscious thinker of these representations, I can exploit this perspective to gain knowledge of my substantive properties—that is one of the limits of pure reason. It is for this reason that Kant sometimes describes "I" as a merely formal or bare mechanism of self-reference: "the 'I', which is simple solely because its representation has no content [. . .] and for this reason seems to represent, or (to use a more correct term) denote, a simple object" (A382). The "I" of self-conscious thought must refer, but from the position of the self-conscious subject, one cannot know to what.

These passages from the paralogisms contain repeated articulations of claims corresponding to both Special Insight and Ordinary Ignorance, which Kant draws on to establish the transcendence of the self of apperceptive self-consciousness.[11] For instance, the claim quoted above that there can be no question whether the "I" of self-conscious thought denotes an object shares with Special Insight a commitment to the idea that from the position of the self-conscious thinker one can be assured of one's own existence. (Or here again: "I think" [. . .] is already involved in every thought. [. . .] [T]he Cartesian inference *cogito*, *ergo sum*, is really a tautology, since the *cogito* (*sum cogitans*) asserts my existence immediately" (A354–55). In the same passage, however, he stresses that we "have no right to transform it into a condition of knowledge." He writes:

> It is obvious that in attaching an "I" to our thoughts we designate the subject of inherence only transcendentally, without noting in it any quality whatsoever—in fact, without knowing anything of it either by direct acquaintance or otherwise. (A355)

[9]Thanks to Anil Gomes for raising this comparison with Kant.

[10]The A and B numbers refer to the page numbers of the first (1781) and second (1787) German editions of Kant's *The Critique of Pure Reason* respectively.

[11]The same themes emerge earlier, too, in the B transcendental deduction, when he first introduces the notion of apperceptive self-consciousness; e.g., "[A]lthough my existence is not indeed appearance (still less mere illusion), the determination of my existence can take place only in conformity with the form of inner sense, according to the special mode in which the manifold, which I combine, is given in inner intuition. Accordingly, I have no *knowledge* of myself as I am but merely as I appear to myself. The consciousness of the self is thus very far from being a knowledge of the self" (B157/8).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

Claims like this are, of course, construed in theoretical language proper to the Kantian system that would make straightforward assimilation to another framework improper. But the epistemic limitation expressed in such passages seems to be closely related to the one captured by Ordinary Ignorance: to know that the *I* of self-conscious thought refers is not to gain knowledge of any of its substantive properties. Kant uses these claims to drive a view of the self of self-consciousness as transcending the limits of our knowledge: the "'I', does not contain in itself the least manifold" (A354). In the next section, I offer a related proposal, that being subject to both the epistemic achievement captured by Special Insight and the epistemic limitation captured by Ordinary Ignorance is apt to produce in us first-person thinkers an illusion of transcendance about the self.

4. A cognitive illusion of transcendence about the self

It will be helpful to recap where we've gotten to so far. According to Special Insight (section 2), *I*-thinkers have a distinctive source of insight into the referential success of their own *I*-thoughts. Merely in virtue of being the thinker of an *I*-thought, one can know that the referent of the thought exists. According to Ordinary Ignorance (section 3), *I*-thinkers do not have corresponding insight into the nontrivial properties of the referent of their *I*-thoughts. One cannot, merely by virtue of being the thinker of an *I*-thought, know anything substantive about the thought's referent. Now, to co-occupy these two perspectives is to take up a highly uncomfortable epistemic position: one's epistemic orientation towards the referent of one's thought is exceptionally powerful, and yet one can thereby know very little about it. To give an analogy, this would be a bit like seeing an object in optimised viewing conditions without being able to say anything substantive about what the seen object is like.[12] The cognitive discomfort of this situation is alleviated, I will now argue, by undergoing an illusion of transcendence about the self.

The psychological mechanism underlying the proposed cognitive illusion is the one posited by *cognitive dissonance theory* in psychology, according to which experienced cognitive inconsistencies of sufficient importance give rise to a state of psychological discomfort ("cognitive disonance") that motivates its own reduction or elimination. Three core commitments of the theory survive largely unchanged from Leon Festinger's seminal formulation (1957):[13]

1. There may exist disonant or "nonfitting" relations among cognitive elements.
2. The existence of dissonance gives rise to pressures to reduce the dissonance and to avoid increases in dissonance.
3. Manifestations of the operation of these pressures include behaviour changes, changes of cognition, and circumspect exposure to new information and new opinions. (31)

A quick word on each: (1) invokes a relation of *dissonance* or *nonfittingness* between cognitions. This relation need not involve strict logical contradiction, but experienced inconsistency in a weaker sense: "the relation between the two elements is dissonant if [. . .] the one does not, or would not be expected to, follow from the other" (15), where those elements bear some relevant thematic

[12]Clearly, there would also be disanalogies between these two cases, too, and I have not here committed to any particular view about how we ordinarily come to know about things and their properties by seeing them. Such a view would plausibly involve mention of additional factors beyond merely seeing the object, including, e.g., additional cognitive capacities, conceptual (and specifically sortal) competence, cognitive penetration effects, favourable externalist epistemic conditions, etc. The force of this quick heuristic depends only on the intuitive claim that it would be *unexpected* to be in optimised viewing conditions, and yet to know nothing substantive about the object. Many thanks to an anonymous reviewer for raising this point.

[13]Though see Fazio and Cooper (1984) and Cooper (2007, chap. 4) for disagreement about whether recognised inconsistencies between cognitions *cause* or are merely reliable proxies for the negative affective state and motivational effects of cognitive dissonance; there is also ongoing debate about *why* the state has the motivational element that it does; see next footnote.

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

relation to each other. Number (2) is the claim that such states of cognitive dissonance motivate their own reduction or elimination. In this, Festinger takes cognitive dissonance to be like hunger (or think, too, of thirst, lust, or tiredness) in which a state of felt discomfort motivates its own removal or reduction: "[c]ognitive dissonance can be seen as an antecedent condition which leads to activity oriented toward dissonance reduction just as hunger leads to activity oriented toward hunger reduction" (3). It is now widely thought that cognitive dissonance should be understood as an *experienced* state of negative affective arousal, and that it is this experienced negative affect that drives the state's motivational component (Cooper 2007, chap. 3).[14] Number (3) raises a number of strategies for the reduction or removal of cognitive dissonance. The range of strategies now known are many, varied, and often striking; we are, it seems, surprisingly flexible in the cognitive contortions we are prepared to undergo in the effort to promote cognitive harmony. Relevant to present purposes, however, is only the schematic suggestion included above that cognitive dissonance can be alleviated by changes of cognition. That's because the proposal of this section is that a cognitive illusion of transcendence of the self is a cognitive change motivated by a state of cognitive dissonance produced by the co-occupation of Special Insight and Ordinary Ignorance.

Consider again the position the first-person thinker finds herself in. Just in virtue of being an *I*-thought's thinker, a subject can know that the referent of her thought exists. This is an exceptionally powerful epistemic position to be in. Despite this, however, she cannot know anything substantive about the thought's referent. Now, this combination does not involve any strict contradiction. But as we have seen, cognitive dissonance does not require a contradiction, only that one element of the subject's cognitive profile *would not be expected to follow* from the other. And this is, I suggest, exactly the relation we find between Special Insight and Ordinary Ignorance. Given the exceptionally powerful epistemic perspective on one's own *I*-thought captured by Special Insight, one would *expect* to be in a position to know about its referent's substantive properties. (Recall the analogy of viewing something in idealised conditions—one would likewise expect to thereby know something about the object's substantive properties). But in the case of *I*-thought, one doesn't. When salient, this frustrated expectation produces a state of cognitive dissonance in first-person thinkers. And this state of cognitive dissonance triggers a cognitive response aimed at relieving itself.

How might the dissonance be resolved? In principle, we might concoct a number of different strategies that could help here. One option, for instance, would be to adjust one's representation of the object of one's *I*-thought so that it *is* represented as having some special nontrivial property that is knowable in virtue of being the thought's thinker. Put this together with the epistemic perspective captured by Special Insight, and the result would be a much more "fitting" set of cognitive facts.[15] There is nothing to rule out a priori that there are individual thinkers whose particular cognitive characters dispose them to resolve the tension this way. What research on this question has shown, however, is that which sort of resolution-strategy is favoured in the face of cognitive dissonance tends to be the one that requires the least overall cognitive change (the "path of least resistance").[16] And there is, I suggest, another available strategy that requires much less overall change: the thinker can undergo an illusion of transcendence about the self—an illusion, recall, in which the referent of one's *I*-thought is cognitively represented as something that is *by its very nature substantively unknowable.* Here the strategy involves minimal change, because it simply externalises the contents of the experienced dissonance onto the portion of reality it concerns. Rather than confront an

[14]Current hypotheses about the motivational element of the state include construing it as a conative drive towards self-consistency or self-affirmation, a drive away from aversive consequences, or a pull towards effective action-guidance. For self-consistency hypothesis see Aronson (1968, 1969,1992, 1999); for self-affirmation see Steele (1988), Steele and Liu (1983), Steele, Spencer, and Lynch (1993); for aversive consequences see Cooper and Fazio (1984); for action-guidance see Harmon-Jones (2002).

[15]Many thanks to an anonymous reviewer for this journal for suggesting this option.

[16]See Harman-Jones (2002, 100) and Festinger (1957, 24).

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

uncomfortable set of cognitive facts according to which one is positioned to know about something's existence, but not its substantive properties, the thinker projects that unknowability onto the nature of the object concerned. By the lights of the illusion of transcendence of the self, the combination of Special Insight and Ordinary Ignorance is perfectly fitting given the metaphysical exoticism of the object they concern—that is, something whose substantive properties *just are* fundamentally unknowable.

Of course, this strategy doesn't exactly demystify the scene. We are left with an inclination to think of selves as mysterious, otherworldly objects. But by projecting this mystery outward beyond the limits of the subject's cognitive profile, it succeeds in alleviating the cognitive dissonance that we might otherwise predict to manifest in first-person thinkers—especially at times when the co-occupation of Special Insight and Ordinary Ignorance is rendered salient, as, for example, when any of us asks ourselves the question raised by Kim in the earlier quote from Ryle: "Who, or What am I?" The proposal of this paper is that the temptation to answer such questions by treating the self as a mysterious or otherworldly object is, at least in part, explained by a cognitive illusion of transcendence about the self that arises as a way of alleviating the felt cognitive dissonance that is otherwise apt to arise from the distinctive epistemology of *I*-thought.

A pair of closing comments will help bring out the intended force of the proposal. First, nothing in the epistemic conditions of first-person thought, the emerging state of cognitive dissonance, or in the proposed resulting cognitive illusion of transcendence of the self as I have characterised them involves the performance of an inference on the part of the thinker. This is a bit of a repetition, but it is important. The proposal is not that first-person thinkers are inclined to *infer* that the self is transcendental. The proposal, rather, is that just as enjoyers of perceptual illusions need only find themselves in the right conditions to be potentially subject to the illusory state, first-person thinkers will simply find themselves apt to be struck by illusions of transcendence about themselves in conditions in which the relevant epistemic conditions of first-person thought are sufficiently salient. This should help with lingering worries about overintellectualisation.

By way of the second comment, recall that I initially introduced the task of this paper as mirroring the task of identifying and explaining the conditions in which the perceptual bent-stick-in-water illusion arises. We are now in a position to see a significant difference between the cognitive transcendence-of-the-self illusion and the perceptual bent-stick-in-water illusion. The difference is that we sometimes see sticks (as such) out of water, but we can never think self-consciously about the self outside a first-person thought, whose conditions of illusion are inbuilt. Of course, we have plenty of ways of thinking about what is in fact oneself outside a first-person thought—I can think of myself using a description, for instance, or a mental demonstrative, or as LS. But these are not ways of thinking about myself *as* myself. For that, I must think a first-person thought. And to think a first-person thought, the proposal has been, is to enter into the conditions in which the cognitive illusion of transcendence about the self might potentially arise.[17]

The upshot of these two observations is that, if the proposal of this paper is correct, then we should expect illusions of transcendence about the self to be (i) widespread and (ii) tenacious. That's because (i) all first-person thinkers will be subject to them, no matter what their powers of inference or conceptual sophistication, and (ii) no alternative nonillusion-inducing way of thinking about the self *as* the self is available to correct the cognitive distortion. Here we come full circle. It was precisely the widespread and tenacious appeal of transcendental views of the self with which we began.

Of all there is to understand in the world, we seem to be among the greatest mysteries to ourselves. As Ryle said, our selves seem to be always eluding us—"always and obdurately a pace ahead." The proposal of this paper has been that this is, at least in part, a kind of cognitive trompe l'oeil produced by the inherent epistemology of *I*-thought. Illusions of transcendence give us a way of reconciling our powerful internal epistemic perspective as thinkers of these thoughts, with our

[17]Kant makes a comparable point about the inescapability of the perspective of the self-consciousness in A364.

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

lack of substantive knowledge from this privileged viewpoint about their objects. If this is right, there is a pleasing contrariness to it. It is precisely because we exceptionally know so much about the referents of our *I*-thoughts that we end up mystifying them.

Acknowledgments. For comments on earlier drafts of this paper, thanks to Daniel Morgan, Lucy O'Brien, Matthew Soteriou, Joe Saunders, and two anonymous referees for this journal. Thanks also to audiences at a number of talks, including a UCL graduate seminar, for exceptionally helpful feedback.

Léa Salje is a philosopher of mind at the University of Leeds. She received her PhD from University College London in 2016. Her primary research interests are first-person thought and bodily self-awareness. In addition, she has published on the format of thought, second-person thought, the significance of nonindexical thought, and immunity to error through misidentification.

References

Aronson, E. 1968. "Dissonance Theory: Progress and Problems." In *Theories of Cognitive Consistency: A Sourcebook*, edited by R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P.H. Tannenbaum. Chicago: Rand McNally.

Aronson, E. 1969. "The Theory of Cognitive Dissonance: A Current Perspective." In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, vol. 4. New York: Academic Press.

Aronson, E. 1992. "The Return of the Repressed: Dissonance Theory Makes a Comeback." *Psychological Inquiry* 3 (4): 303–11.

Aronson, E. 1999. "Dissonance, Hypocrisy, and the Self Concept." In *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*, edited by E. Harman-Jones and J. Mills. Washington, DC: American Psychological Association.

Anscombe, G. E. M. (1975) 1994. "The First Person." In *Self-Knowledge*, edited by Quassim Cassam. Oxford: Oxford University Press.

Cassam, Quassim. 2007. *The Possibility of Knowledge*. Oxford: Oxford University Press.

Coliva, Annalisa. 2012. "Which 'Key to All Mythologies' about the Self? A Note on Where Illusions of Transcendence Come from and How to Resist Them." In *Immunity to Error through Misidentification*, edited by Simon Prosser and François Recanati. Cambridge: Cambridge University Press.

Coliva, Annalisa 2017. "Stopping Points: 'I', Immunity and the Real Guarantee." *Inquiry: An Interdisciplinary Journal of Philosophy* 60 (3): 233–52.

Cooper, J. 2007. *Cognitive Dissonance*. Los Angeles: Sage.

Cooper, J., and R. H. Fazio. 1984. "A New Look at Dissonance Theory." *Advances in Experimental Social Psychology*, vol. 17, edited by L. Berkowitz, 229–66. New York: Academic Press.

Festinger, L. 1957. *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Harman-Jones, E. 2002. "A Cognitive Dissonance Theory Perspective on Persuasion." In *The Persuasion Handbook: Developments in Theory and Practice*, edited by James Price Dillard and Michael Pfau. Thousand Oaks, CA: Sage.

Kahneman, Daniel. 2012. *Thinking Fast and Slow*. London: Penguin.

Kant, Immanuel. (1787) 2003. *The Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: Palgrave MacMillan.

Macpherson, Fiona, and Clare Batty. 2016. "Redefining Illusion and Hallucination." *Philosophical Issues* 26 (1): 263–96.

McDowell, John. 1998. "Singular Thought and the Extent of Inner Space." In *Meaning, Knowledge and Reality*, edited by John McDowell. Cambridge, MA: Harvard University Press.

Moore, G. E. 1939. "Proof of the External World." *Proceedings of the British Academy* 25 (5): 273–300.

O'Brien, Lucy. 2007. *Self-Knowing Agents*. Oxford: Oxford University Press.

O'Brien, Lucy. 2015. "Ambulo Ergo Sum." *Royal Institute of Philosophy Supplement* 76: 57–75.

Peacocke, Christopher. 1999. *Being Known*. Oxford: Oxford University Press.

Peacocke, Christopher. 2012. "Defending Descartes." *Aristotelian Society Supplementary Volume* 86 (1): 109–25.

Pryor, James. 1999. "Immunity to Error through Misidentification." *Philosophical Topics* 26 (1/2): 271–304.

Putnam, Hilary. 1981. "Brains in a Vat." *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Ross, L., M. R. Lepper, and M. Hubbard. 1975. "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm." *Journal of Personality and Social Psychology* 32 (5): 880–92.

Ryle, Gilbert. 1994. "Self-Knowledge." In *Self-Knowledge*, edited by Quassim Cassam. Oxford: Oxford University Press.

Shoemaker, Sydney. 1968. "Self-Reference and Self-Awareness." *Journal of Philosophy* 65: 555–67.

Shoemaker, Sydney. 1994. "Self-Knowledge and 'Inner Sense': Lecture 1: The Object Perception Model." *Philosophy and Phenomenological Research* 54 (2): 249–69.

Steele, C. M. 1988. "The Psychology of Self-Affirmation: Sustaining the Integrity of the Self." In *Advances in Experimental Social Psychology*, vol. 21, edited by L. Berkowitz. New York: Academic Press.

Steele, C., and T. J. Liu. 1983. Dissonance Processes as Self-Affirmation. *Journal of Personality and Social Psychology* 45 (1): 5–19.

https://doi.org/10.1017/can.2020.22 Published online by Cambridge University Press

Steele, C., S. J. Spencer, and M. Lynch. 1993. "Self-Image Resilience and Dissonance: The Role of Affirmational Resources." *Journal of Personality and Social Psychology* 64 (4): 885–96.

Tversky, A., and D. Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 27 (4157): 1124–31.

Warfield, Ted A. 1997. "Externalism, Privileged Self-Knowledge, and the Irrelevance of Slow Switching." *Analysis* 57 (4): 282–84.

Warfield, Ted A. 1998. "A Priori Knowledge of the World: Knowing the World by Knowing Our Minds." *Philosophical Studies* 92 (1/2): 127–47.

Cite this article: Salje, L. 2020. Lit from Within: First-Person Thought and Illusions of Transcendence. *Canadian Journal of Philosophy* 50: 735–749, doi:10.1017/can.2020.22