

teorema

Vol. XXXVII/1, 2018, pp. 5-22

ISSN 0210-1602

[BIBLID 0210-1602 (2018) 37:1; pp. 5-22

The Role of Inner Speech in Self-Knowledge: Against Neo-Rylean Views

Dorit Bar-On and Jordan Ochs

RESUMEN

Se piensa que nuestro conocimiento de nuestros actuales estados mentales es, a la vez, característico y privilegiado. Gilbert Ryle es célebre por haber ofrecido una explicación deflacionista del auto-conocimiento, argumentando que nuestros *episodios de lenguaje interno* podrían servir como base evidencial privilegiada (y posiblemente también característica) para el auto-conocimiento de los estados mentales. La explicación de Ryle ha sido, en gran parte, rechazada. Sin embargo, diversos autores han intentado recientemente traer de nuevo a la vida la explicación de Ryle como un modo de dar cuenta del papel que desempeña el lenguaje interno en el auto-conocimiento. En este artículo, evaluamos críticamente dos de tales explicaciones “neo-ryleanas”, argumentando que son insatisfactorias, especialmente porque no pueden dar cuenta del carácter privilegiado de nuestro auto-conocimiento de los estados mentales. Extraemos de tal evaluación algunos *desiderata* que debe cumplir una teoría que pueda explicar adecuadamente la significación de los episodios de lenguaje interno para el auto-conocimiento privilegiado. Concluimos sugiriendo que esos *desiderata* favorecen un enfoque neo-expresivista para la comprensión del papel que desempeña el lenguaje interno en el auto-conocimiento.

PALABRAS CLAVE: *auto-conocimiento, lenguaje interno, estados mentales, neo-ryleanismo, neo-expresivismo, privilegio de primera persona.*

ABSTRACT

Our knowledge of our own present states of mind is thought to be both distinctive and privileged. Gilbert Ryle has famously offered a deflationary account of self-knowledge, arguing that our *inner speech episodes* (ISEs) could serve as a privileged (and possibly also distinctive) evidential basis for self-knowledge of mental states. Ryle's account has been, for the most part, rejected. However, several authors have recently attempted to revive the Rylean account by way of explaining the role of inner speech in self-knowledge. In this paper, we critically evaluate two such “neo-Rylean” accounts, arguing that they are unsatisfactory, especially since they cannot account for the privileged character of our self-knowledge of mental states. We extract desiderata for a theory that can adequately explain the significance of ISEs for privileged self-knowledge. We conclude by suggesting that these desiderata favor a *neo-expressivist* approach to understanding the role of inner speech in self-knowledge.

KEYWORDS: *Self-knowledge, Inner Speech, Mental States, Neo-Ryleanism, Neo-Expressivism, 1st-Person Privilege.*

I. INTRODUCTION

Consider the following utterances:

- a. “I would love a cup of tea.”
- b. “A cup of tea would be nice.”
- c. “Oh for some tea!”

In each case, the person who makes these utterances uses a linguistic vehicle to express her desire for tea. And, in each case, it seems the person (assuming her utterances are sincere) would also know that she has a desire for a cup of tea. But what, if anything, does the production of such utterances have to do with the *knowledge* one has that one is in a particular mental state?

One well-known proposal made by Gilbert Ryle was that we could observe our own behaviors and infer from them what states of mind we are in [Ryle (1949)]. So, if I am filling up the kettle with water and getting a mug from the cupboard I might infer, and hence know, that I’d like a cup of tea. In the same way, when I self-attribute a desire for tea, as in a., or express my desire in the manner of b. or c., or non-verbally, by eagerly reaching for the mug, I may infer, and hence know, that I’d like some tea. This model applies in the arena of inner speech episodes, too. So, even if I only profess a desire for tea in my head without any external vocalization, I can still observe my inner speech episode (ISE) and infer – and hence know – that I’d like a cup of tea. Our ISEs, then, constitute one type of *evidential basis* from which we can gain inferential knowledge of our states of mind. (The idea that there needs to be some such basis for knowledge is a common assumption held by epistemologists. Below, we refer to it as the Epistemic Basis Requirement.)

Ryle’s specific account has been (for the most part) rejected on various grounds. One strong reason against his account is that it cannot accommodate *first-person privilege* – the commonsense idea that subjects who are capable of having beliefs about their own states of mind are in a privileged position (as compared with others) to have knowledge of those states of mind. Yet several authors have recently tried to revive (versions of) the Rylean account in order to explain the role of inner speech in self-knowledge [Byrne (2011), Carruthers (2011), Cassam (2015), Roessler (2016)]. In sections II and III we critically evaluate two such *neo-Rylean* accounts, due to Alex Byrne and Johannes Roessler, respectively. In section 4, we extract desiderata for a satisfactory explanation of the epistemic significance of inner speech for privileged self-knowledge. We conclude by

suggesting that these desiderata favor a non-Rylean, *neo-expressivist* account of the role of inner speech in self-knowledge. (We leave the development of the latter account for another occasion.)

II. NEO-RYLEANISM, TAKE I: BYRNE

One recent account of our knowledge of our own thoughts, due to Alex Byrne (2011), takes a direct lead from Ryle's well-known proposal that inner speech constitutes a contingent yet relatively reliable *evidential* source or *basis* for our beliefs about our own minds. In our own case, Ryle suggested, we are not restricted – as we are in the case of knowledge of others' minds – to observations of (and inferences from) overt behavior. We are in a position to '*eavesdrop*' on our own inner speech. Byrne approves of the *economical* character of Ryle's account: the fact that it explains our knowledge of our own mind without invoking capacities that go beyond those we employ in coming to know about other matters [Byrne (2011), p. 109]. He also thinks that Ryle can accommodate the *distinctive* character of self-knowledge of thoughts,¹ insofar as only we ourselves have access to our own ISEs. But Byrne thinks that the Rylean account does not pay sufficient heed to the *privileged* character of self-knowledge – the fact that the beliefs we have regarding our own states of mind are *more apt to constitute knowledge* than others' beliefs about our states of mind. Byrne thinks this can be remedied by appealing to “an Evans-style ‘transparency’ procedure” [Byrne (2011), p. 110] that we have available to us when forming beliefs about our own thoughts. The idea behind such a procedure is that it enables us to infer something about our own mind by considering something outside our mind. So, for example, we can infer that we *believe* that there's a tree in front of us by just looking to see whether there's a tree in front of us. Insofar as this type of procedure can only be used to get secure beliefs about our own states of mind, and we cannot go wrong when employing it, invoking it can help explain the distinctive *and* privileged character of self-knowledge.

In this vein, Byrne maintains that we can be said to have privileged knowledge of a type of first-order mental state (belief that *p*, seeing *x*, intending to ϕ , wanting *y*, and so on) in virtue of following an *epistemically good*, transparent rule in reaching second-order beliefs about the first-order state. The rule takes the following schematic form:

If *C*, then believe that you are in *M*,

where C represents the evidential base from which the self-knowing subject *infers* (perhaps only tacitly) the second-order belief about her first-order state. The evidential basis is *neutral* insofar as it makes no reference to the self-knower's own mental states (or her knowledge thereof).² In the case of knowledge of what we are thinking, Byrne proposes the following rule, whose antecedent invokes inner speech:

THINK If the inner voice speaks about x, believe that you are thinking about x [Byrne (2011), p. 117].

It may appear at first blush that THINK fails to accommodate Byrne's own desideratum of neutrality. ISEs are paradigmatic mental events. If following THINK requires one to infer what she thinks from some of her states of mind, then the rule is clearly not neutral. (Knowing what you think would seem to depend on your ability to recognize some other states of mind you are in – namely, your ISEs.) Byrne claims that THINK's antecedent avoids this problem, since it does not mention the *rule-follower's* inner speech. To follow THINK, one only needs to consider *some* inner voice that is speaking; one needn't think of it as *her own* inner voice. Still, it would seem that, if inner speech is to be playing a role in the explanation of self-knowledge of thoughts, the rule-follower has to be credited with using an *inner* speech episode, which is a mental event, as the epistemic basis for her self-beliefs, and this threatens THINK's neutrality.

Now Byrne thinks that there actually are no inner voices; so strictly speaking, there are no inner *speech* episodes that one can consider in following THINK. That is, he believes there are no mental objects that are *silent utterances of sentences in a natural language*. Still, he maintains there are *phonological representations* of sentences, the activation and quasi-perception of which make it *seem* to us as though there are sentences uttered in inner speech. With this claim on the table Byrne explains that, since there is no such thing as inwardly *speaking* about anything, we cannot gain knowledge about our mental states by *literally* following the THINK rule. As he says: "If one follows THINK, one recognizes, hence knows, that the inner voice speaks about x. Since there is no inner voice, there is no such knowledge to be had, and one cannot follow THINK" [Byrne (2011), p. 117]. However, Byrne points out, one can *try* to follow the rule. And even merely *trying* to follow the rule will result in one's gaining knowledge of what they are thinking.

Setting aside these complications, what makes THINK a good rule of inference, the following of which can lead to distinctive and privileged knowledge of one's thoughts? Regarding distinctiveness, Byrne cites approvingly Ryle's observation that "I cannot overhear your silent colloquies with yourself"; presumably, THINK can yield distinctive self-knowledge, since its antecedent concerns something only I am in a position to recognize (viz. my own phonological imaginings). Regarding privilege, Byrne asks us to consider what he takes to be the third-personal version of the THINK rule.

THINKk: If Kylie speaks about x, believe that Kylie is thinking about x.

Byrne argues that, if Kylie herself were to follow the *first*-personal THINK rule in coming to know what she is thinking, she would be likelier to be right about her own thoughts than we are when following the third-personal version of the rule (THINKk), which involves inferring what Kylie is thinking about on the basis of her external utterance. This is because, in hearing another's external utterance, we may mishear *what* a speaker has said and even misidentify the speaker. Given that these possibilities of error arise in the third-personal case but not in the first-personal one, the beliefs we reach about our own thoughts using inner speech are likelier to constitute knowledge than the beliefs others could form about our mental states on the basis of our external utterances.

What Byrne says about both the peculiar (what we refer to as 'distinctive') and the privileged character of self-knowledge of thoughts is actually not easy to follow. When introducing the two notions [Byrne (2011) Section II], he characterizes peculiarity – following McKinsey – in terms of having access to our own mental states "that is available to no one else" and points out that Ryle "denies that we have *peculiar* access", though "he thinks that we have *privileged* access to our mental states", because we at least sometimes have better evidence about ourselves [Byrne (2011) pp.107f.]. But when he returns to privileged and peculiar access (in Section X), he cites Ryle as a plausible source of explanation of peculiarity (as noted earlier), and explains privilege by appealing to the fact that following the third-person version of THINK is not conducive to knowledge, whereas (even just trying to) follow THINK is likely to result in knowledge. Indeed, he suggests that we can improve on Ryle, by combining the idea that we use inner speech in learning of our own

thought with the Evans-inspired idea that we use a *transparent* rule of inference in forming our self-beliefs. Invoking such a rule “(arguably) explains *both* privileged and peculiar access” [Byrne (2011) p. 110].

For the most part, Byrne seems to portray THINK as Rylean in spirit, in that it attempts to capture Ryle’s idea that our ISEs serve as a (privileged) epistemic basis on which we form our beliefs about our own thoughts. However, at times, Byrne seems to point to a much stronger form of first-person privilege associated with following THINK. In the course of contrasting THINK and THINKk, he considers the possibility that one can mistake an (imagined) outer voice for one’s inner voice. He says:

[E]ven if she has mistaken an outer voice for her inner one (...), her belief that she is thinking about water is *still* true, and will likely be knowledge. (If you hear someone else speak about water, *the very act of understanding their words requires that you think about water.*) [Byrne (2011), p. 119, emphasis added].

The suggestion *seems* to be that even the mere act of trying to follow THINK will result in a true belief about one’s thought, *because* merely thinking that an (imagined) inner voice is speaking about x *ipso facto* results in thinking about x. But this suggests that the epistemic privilege associated with THINK comes from the fact that even just trying to follow THINK *guarantees* that one will come to have a true belief about what one is thinking. This is because the effort will inevitably result in one’s thinking about x, thereby making the self-attribution of thought true. This, however, would mean that the self-knowledge obtainable by following THINK is no more substantial than the knowledge one can have that one is speaking now, when one says: “I am speaking now”. Here, the very act of self-attribution brings into existence the state that one ends up self-attributing. And, similarly, the act of attributing to oneself a thought about x may seem sufficient for making the self-attributive thought – that one is thinking about x – true. If so, then THINK would seem to be *self-verifying*.³

Although the self-verification reading is strongly suggested by the above quotation from Byrne, he in fact says that “[t]he beliefs produced by trying to follow THINK are *not* absolutely guaranteed to be true” [*ibid.*]. And, in correspondence, he has said that THINK “is not supposed to be self-verifying because believing the antecedent doesn’t guarantee that you’re thinking, which I am taking to be episodic (although it does guarantee that you believe...)”. The reason Byrne provides for denying self-verification seems puzzling to us. Very briefly, it may be that believing

the antecedent does not guarantee that you're thinking it. But *following* the rule surely requires *episodically taking* it – or occurrently judging – that the inner voice is speaking about x. And this does seem to guarantee that one is thinking about x.

However, there are other good reasons to reject the self-verifying understanding of the epistemic success of Byrne's rule THINK. Such a reading strays far from the original Rylean idea – which, as we saw earlier, Byrne seeks to capture – that we can *obtain* self-knowledge concerning contingent mental goings-on *by attending* to our own ISEs as inner behaviors on the basis of which we can make (non-self-verifying) inferences about our own states of mind. By contrast, recall that on Byrne's view, we cannot really attend to our inner speech, because there is actually no inner speech. If we were to adopt the self-verification interpretation, it would turn out that we cannot make any abductive inference concerning our states of mind using the observation of our ISEs as our basis. Instead, all we could do is *try* to follow a certain inference rule, where the (putative) knowledge we gain is of a state that is in fact *produced by* the very attempt to follow the rule.

Relatedly, if the epistemic advantage that accrues to following THINK is exhausted by the fact of self-verification, there would seem to be nothing special about the use of, specifically, *inner speech* in obtaining privileged self-knowledge of thoughts. To see this, consider the following variant of THINKk:

THINKk': If Kylie speaks about x, believe that you are thinking about x.⁴

This rule shares important features with THINK; it can even be seen to improve on it, if we consider the desiderata that Byrne lays down for his own view. If a rule-follower takes (and even if she mistakes) Kylie (or any outer voice, or words on a page, for that matter) to be speaking about x, she will *ipso facto* be thinking about x, and thus her self-attribution of a thought about x is bound to be true. And this is only so in her own case. Moreover, since the antecedent of THINKk' makes *no* mention of the rule-follower's mental states, and is 'outward-looking', it is properly neutral and transparent. Finally, unlike THINK, THINKk' is a rule that we can actually follow (as opposed to only trying to follow).

Byrne might say that, as a matter of psychological fact, we *do* follow THINK but do not follow rules like THINKk'. But surely, we *do* sometimes follow these other rules. Say you see me gazing into space and you ask:

“What are you thinking about?” – when I’ve just been looking at a book. I may well say: “I’m thinking about the incompatibility of externalism and self-knowledge”. In giving the answer, I may simply be considering the paper I’ve just been reading (I may not have any ISEs). I just ‘move’ directly from what I have read to what my thoughts are about. Or, I ask you the same question, after you have just been listening to John’s rave review of a film he had seen. You might say: “I’m thinking about the movie John saw” (or just “I’m thinking about what John said”). Again, these cases fit very well Byrne’s pattern of transparency procedures (or ‘policies’, as he sometimes refers to them). If you are in a position to recognize the antecedents (that John spoke about *x*, that the text read was about *y*) you could securely move to a belief about your thoughts.

We take it that rules like ‘THINKK’ would *not* serve as good rules for a properly neo-Rylean account, insofar as they assign no specific role to *inner speech* (or even to *imagined inner speech*) in explaining self-knowledge of thoughts. To the extent that a self-verification reading of THINK opens the door for rules like ‘THINKK’ (consistently with Byrne’s own *de-siderata*), this counts against such a reading.

But now suppose that we steer away from the self-verification reading of THINK and go back to treating it as a substantive rule of inference. Byrne would then have to explain what renders THINK an especially good epistemic rule to follow in one’s own case in terms of its distinctive reliability. However, there is reason to question this reliability. As argued by Johannes Roessler (whose positive view we discuss in the next section), THINK – understood as a substantive rule taking one from episodes of imagined inner speech to beliefs about one’s own thoughts – is *not* reliably truth-conducive. Following THINK is apt to “produce false positives”, since “[n]ot every act of inner speech amounts to an episode of thinking” [Roessler (2016), p. 546]. Consider a case in which you internally recite a war poem. Your (imagined) inner voice has spoken of war. However, Roessler observes, the episode “does not necessarily involve thinking about war” [Ibid.]. This clearly suggests that directly basing one’s belief about one’s present thoughts on an episode of (imagined) inner speech is epistemically risky.⁵

III. NEO-RYLEANISM TAKE II: ROESSLER

Roessler himself thinks that Byrne’s view rests on a misguided starting point: the idea that we should look to inner speech to provide an *evidential source* for self-knowledge of what we think. Indeed, he altogether

rejects the Epistemic Basis Requirement – the claim that where there is knowledge there must be some distinctive epistemic basis on which the relevant beliefs are formed or arrived at [See Roessler (2015), pp. 157ff.]. If we are nevertheless to explain how we can be credited with distinctive and privileged knowledge of our own thoughts, Roessler thinks, we should not focus on the *spectatorial* aspect of the phenomenon of inner speech that features prominently in Byrne’s neo-Rylean account. Instead of focusing on ISEs as something we confront as inner hearers, we should be looking at ISEs as *mental acts* we perform *qua* inner *speakers*.

Roessler still thinks we can look to Ryle for insight concerning inner speech and self-knowledge. He points out that in addition to the more familiar Rylean observation-and-inference account that Byrne draws on, there is another (separate) idea to be found in Ryle’s work, namely, that we can also know our minds insofar as we know the *activities* we engage in as agents. According to Ryle’s second idea, one can know what one is thinking in virtue of being “alive to what one is doing” [Roessler (2016), p. 543]. Roessler understands the Rylean notion of ‘being alive to’ in terms of having practical knowledge of the kind described by Elizabeth Anscombe (1957) and Stuart Hampshire (1959) and proposes that we extend a ‘practical knowledge’ account to the case of *inner speech acts*. The result is an ‘ability-based’ explanation of self-knowledge. The ability at the heart of this explanation “is that of reasoning practically and in that way acquiring intentions ... it’s the intention informing an operation that makes the agent’s knowledge of what she is doing intelligible” [Roessler (2016), pp.550].

Roessler’s idea seems to be (at least roughly) this. Motivated by his rejection of Byrne’s seemingly alienated account, Roessler is trying to identify a more intimate epistemic relation we stand in to our ISEs. To begin with, Roessler suggests that we should draw a sharp distinction between ‘mere acts’ of inner speech and ‘inner speech acts’ [Roessler (2016) p. 548, following Green (2013)]. An example of the former is the above-mentioned case of reciting a poem in inner speech, which, according to Roessler, is not a case of *thinking* at all. By contrast, in many cases, what we do in inner speech is *asserting*, *conjecturing*, *wondering*, and so on. Roessler then proposes that “whether an act of inner speech ... amounts to an episode of thinking *depends on the intentions informing it*” [Roessler (2016), p. 549, emphasis added]. *Inner speech acts* (as opposed to mere acts of inner speech) are ones that are informed by intentions [Ibid., p. 547]. And only those ISEs that constitute inner speech acts should count as instances of

genuine thinking. But now we can apply the practical knowledge account to explain the distinctive and privileged knowledge we have of what we are thinking. On that account, when engaging in intentional acts, we have non-observational, non-inferential practical knowledge – ‘knowledge-in-intention’ – of *what we are doing*. So, insofar as the ISEs that are instances of (genuine) thinking constitute intentional activities we engage in, we can be said to have practical ‘knowledge-in-intention’ of *what we are thinking*.

It’s important to recognize that Roessler cannot offer knowledge-in-intention as a way of explaining *how we gain* self-knowledge in virtue of our ISEs (which is the question both Ryle and Byrne seem to want to answer). Appeal to the intentions with which one has done something, ϕ , is typically made by way of responding to the reason-giving question: *Why* did you ϕ ? For example, suppose you approach the door and begin to turn the handle, and are asked: Why are you turning the door handle? You would typically respond by identifying the intention that informs your action: “I’m opening the door” [Anscombe (1957) pp. 157ff.]. But knowledge-in-intention does not satisfy the Epistemic Basis Requirement – i.e., citing the relevant intention does not supply a distinctive *basis* or *source* which one *uses* to arrive at privileged self-knowledge. Indeed (as mentioned earlier) Roessler himself rejects the Epistemic Basis Requirement. His idea is that we do not *come to know* what we are thinking *by attending* to our inner speech. Rather, *in producing* inner speech acts, we can be said to have (*practical*) *knowledge-in-intention* of what we are doing.

One immediate difficulty with basing an account of the role of ISEs in self-knowledge on their character as *intentional acts* is this. At least some of our ISEs appear to have a passive character (e.g. being struck by a thought out of the blue, or having thoughts running through one’s mind as one falls asleep). But Roessler cautions against equivocating on the active/passive distinction. He explains that “(t)houghts with respect to which we are passive are said to be thoughts that do not occur by our own active doing”; but “a process may also naturally be labeled ‘active’ simply insofar as it is or involves an activity” [Roessler (2016) p.553]. Cases of passive inner speech have been wrongly classified as processes that do not involve activity simply on the grounds that they are not things we *intentionally initiate*. However, Roessler argues, being an activity does not require being initiated by a prior intention.

Of greater concern for Roessler’s proposal are cases of thinking that are not plausibly regarded as themselves intentional acts. Take *concluding* (or *realizing* or *discovering*) that p , for example. An episode of concluding that p can be token-identical with an event describable as

episode of *saying* that p. But, although the saying is an intentional act, there are reasons for denying that the episode under the description ‘*reaching the conclusion* that p’ is an intentional act. This is because to describe the episode in the richer terms involves imputing to the speaker propositional attitudes and processes that go well beyond those required for just *saying* that p, and that *as such* cannot themselves be understood as informed by the speaker’s intentions. Roessler acknowledges that this presents a major obstacle for his account. If episodes of concluding, realizing, etc. are not intentional acts, then how can we be said to have knowledge of them on this view?

To address this difficulty, Roessler explains that “some of the things we know are grounded in practical knowledge, even though our knowledge of them is not itself a case of practical knowledge” [Roessler (2016) p.550]. At least “[i]nsofar as thinking is conducted in words” an episode of, e.g., “reaching a conclusion... does involve an intentional action, viz. an ‘inner speech act’” [Roessler (2016) p.549]. It is true that, when one is engaged in an ISE that is a concluding that p, all that one has is non-inferential practical knowledge that one is *saying* in inner speech that p – in virtue of having knowledge-in-intention of what one is (inwardly) doing. Still, one can be entitled to re-describe the relevant episode in terms “laden with ascriptions of attitudes” [Roessler (2016) p.551]. And this should allow us to explain what ‘underpins’ our knowledge of what we are thinking when that thinking is an instance of concluding, discovering, etc.

But what is it that entitles one to such rich re-descriptions of inner sayings? Here Roessler appeals to our background knowledge and ‘appreciation’ of the status of an episode of saying that p, as, for example, the termination of a knowledge-conducive process. It is this that “enables us to know the event in question under thick doxastic and epistemic descriptions” [Roessler (2016) p.551]. Notice, however, that this lands Roessler in a partially *inferential* account of self-knowledge of thoughts (understood as inner speech acts). On the full account, we only have non-inferential, practical knowledge-in-intention of our inner speech *sayings*. Our knowledge of what we are doing *in* saying – the knowledge of the inner speech act performed – awaits inference that relies on background knowledge, memory, and recognition of the episode’s status. Roessler can still escape the standard charge of ‘alienation’ often leveled against Rylean and neo-Rylean inferentialist accounts – the charge that they fail to capture the first-person character of self-knowledge. This is

because inferential self-knowledge by Roessler's lights rests on a uniquely first-personal knowledge-in-intention of our inner sayings. However, in the end, accepting Roessler's account would require denying the appearance of the *groundlessness* of self-knowledge, as that knowledge is (in most cases) based on inference. And this, on its face, seems in tension with Roessler's rejection of the Epistemic Basis Requirement.

But even setting this aside, there is a difficulty that strikes at the heart of Roessler's explanatory strategy. The difficulty has to do with the very idea of knowledge-in-intention on which Roessler bases his account.⁶ It seems plausible to maintain that, if we have non-inferential knowledge-in-intention of our acts, this is at least in part due to the fact that we have non-inferential knowledge of *the intentions with which* we perform those acts. (This is not to say that we are constantly cognizant of those intentions as we act; it's just to say that if queried as to what we are doing – and why -- we are in a position to respond without reliance on any observation or inference.) Intentions, however, are states of mind, on a par with other states of mind such as hopes, desires, fears, and thoughts. Given that, knowledge of intentions falls directly under the scope of the explananda for an account of self-knowledge of the sort Roessler seeks. An account that tries to explain how we have first-person knowledge of our thoughts by appealing to first-person knowledge we have of our intentions would seem to be at risk of moving in too tight an explanatory circle. To put it differently, if knowledge-in-intention is to shed any light on self-knowledge of thoughts, *its* explanation had better not invoke self-knowledge of thoughts and thought-like states of mind. If we already have an understanding of the latter self-knowledge, then, at best, the appeal to knowledge-in-intention would be explanatorily idle.⁷

IV. THE ROLE OF INNER SPEECH IN SELF-KNOWLEDGE: SOME DESIDERATA

Let us take stock. Our critical evaluation of Byrne's and Roessler's neo-Rylean accounts has yielded certain desiderata for an account of the role that inner speech has to play in self-knowledge of our own thoughts. Such an account, we suggest, should preserve (at least) the following features:

- (1) In some way, such self-knowledge involves a genuine epistemic achievement. Self-knowledge of thoughts doesn't 'come for free', courtesy of the self-verifying character of self-attributions of thoughts.

- (2) Self-knowledge of thoughts does not depend on a contingently reliable introspective observation, or else on some kind of self-inference. In this sense, it is not *based on* anything.
- (3) Self-knowledge of thoughts that involves inner speech is distinctively first-personal. One has such self-knowledge in virtue of being the subject of the thoughts.
- (4) At the same time, the relation we stand in to our ISEs is not that of 'listening in' on – or otherwise witnessing – them. Instead, we have a more *active* relation to our ISEs – akin to the relation we have to our outer speech episodes. We *engage in ISEs as (inner) speakers* or producers.
- (5) However, the knowledge we have of our thoughts in virtue of being inner speakers is richer than the mere knowledge that we are *saying something* (in inner speech).
- (6) In addition, the explanation of such self-knowledge should not presuppose first-person knowledge we have of the intentions with which we produce the relevant ISEs.

Recall that Byrne's account fails on desiderata 1-4. On the self-verifying reading of Byrne's inference rule, the only knowledge we can get is insubstantial (contra 1). Yet, even if his THINK rule yielded substantial knowledge, that knowledge would be based on inference, which violates 2. On the self-verifying reading of Byrne, inner speech turns out not to play a distinctively first-personal role in explaining self-knowledge of thought (contra 3); and on the substantive reading, it is not clear what role it can play in giving us *knowledge*, given that (on Byrne's view) there *are* no such things as ISEs. And finally, the success of THINK (such as it is) depends on our being (putative) hearers of ISEs as events from which we infer something about what we are thinking, which violates 4. It should be noted that although Roessler rejects the Epistemic Basis Requirement, his view ultimately invokes inferential knowledge of inner speech acts, which makes it the case that his account also fails to meet 2. Since the only non-inferential knowledge possible on his account is knowledge that we are saying something in inner speech, he fails to accommodate 5. And, since his account is rooted in practical knowledge, it will presuppose knowledge of intentions, thereby violating 6.

V. SATISFYING THE DESIDERATA?

Developing a full account that preserves all the above features is far from trivial, and is something we cannot hope to undertake here. However, in this concluding section we would like to point to what we take to be a promising direction, by drawing on the *neo-expressivist* account of avowals and self-knowledge developed in Bar-On (2004) and elsewhere.⁸ Recall our opening examples:

- a. “I would love a cup of tea.”
- b. “A cup of tea would be nice.”
- c. “Oh for some tea!”

Although sentence a. differs semantically from b. and c. in that it involves a *self-attribution* of a mental state, a typical act of producing that sentence, like typical acts of producing b. or c., will constitute an *expressive act* in which a subject *speaks her mind*, giving direct voice to her current states of mind. Suppose you were to say “DB would like a cup of tea”. You would be employing a sentence that – like a. – *semantically* expresses the proposition that DB would like a cup of tea, and, if you were sincere, you would be giving voice – expressing in the *act* sense (a-express) your *belief* that DB would like a cup of tea. When uttering b., on the other hand, I use a sentence that semantically expresses a proposition that describes something as nice; but in uttering the sentence, I would typically also be giving voice to my desire for tea (whereas you can only use that sentence to express *your* desire for tea).

On the neo-expressivist account, avowals are different from evidential reports concerning states of mind (whether others’ or our own), in that they are acts in which we a-express *the very state picked out by the mentalistic term used in the linguistic vehicle employed*. This allows us to explain *avowals’ security* (viz., the special presumption of truth that governs them), and the *epistemic asymmetry* between avowals and (non-avowing) reports of the same states, yet *without* denying (as did traditional expressivist accounts) that avowals are truth-evaluable, just like moans and groans. The relevant contrasts hold *not* between *types of sentences with certain semantic contents*, but rather between *acts that directly express* one’s mental state and *reports* of that state, whoever produces them, and however reliably. (Notably, e.g. the *sentence* “I am upset at my mother” can be used by an individual not to *avow* her feeling upset, but rather to issue an informed

evidential report about her psychological state, say, at the conclusion of a therapy session.)

According to neo-expressivism, then, avowals considered as *acts* – like grimaces and groans – directly a-express the mental states avowed. For this reason, it is inappropriate to question an avowal, to ask the avower to justify it, to correct her, etc. This is the source of the epistemic asymmetries between avowals and other claims. However, considered as *products* (or in terms of the linguistic *vehicles* used in producing them), avowals and other linguistic expressions are importantly different from non-verbal expressions. For they are acts of *speaking* one's mind, using linguistic tokens with semantic contents, thereby rendering them semantically continuous with verbal utterances (including *non-avowing* self-reports of mental states).

Returning to inner speech, we note, first, that the neo-expressivist account is designed to apply to expressive utterances produced not only in overt speech but also in thought. This means that the account should be readily applicable to acts performed in inner speech. Our tentative proposal is that we should think of the ISEs that are relevant to self-knowledge as *acts of innerly speaking our mind*. Such acts are epistemically immediate; they are not based on any observation, inference, or self-interpretation; at least in that sense they are groundless. Moreover, avowals are suitably active, and non-spectatorial; and they are distinctively first-personal, thus satisfying desiderata 3 and 4. Yet avowals are not intentional acts in the rich sense canvassed earlier. Speaking one's mind (whether out loud or innerly) is something one does – it doesn't merely happen to one. When one does it in inner voice, one performs a mental act. But, as has been argued by several authors [see, e.g. Soteriou (2013)], not all mental acts constitute intentional actions – at least not if that means acts produced or even just sustained by a conscious intention, or else done with some communicative purpose or toward some end or to fulfill some desire. Therefore, the neo-expressivist account can accommodate desideratum 6.

The neo-expressivist understanding of acts of speaking one's mind – including acts performed in inner speech – seems to have the right overall shape for meeting at least some of the main desiderata mentioned earlier. Recall, however, that our consideration of inner speech was initially motivated by the question: How can we *know* our own thoughts in a distinctive and privileged way? And our criticisms of both Byrne and Ryle had to do, specifically, with their handling of the kind of knowledge

afforded by ISEs. Given the neo-expressivist understanding of the special security associated with inner acts of speaking one's mind, what can we learn about the self-knowledge such acts manifest? And, more specifically, how can the neo-expressivist improve on the two neo-Rylean accounts we have criticized?

Answering these questions goes beyond our scope here.⁹ However, it is important to recognize that, on the neo-expressivist account of inner speech's role in self-knowledge to be proposed, the relevant self-knowledge of thoughts (and other states of mind) would not require "innerly hearing" one's ISEs. In general, it should not be thought that a subject who speaks her mind (outwardly *or* inwardly) must await her expressive utterances in order to *come to know* what she feels, wants, or thinks. Note that the relevant knowledge on the neo-expressivist account would be knowledge of a mental state one is in, not merely what one is innerly saying (which satisfies desideratum 5). In keeping with the neo-expressivist's rejection of the Epistemic Basis Requirement, acts of speaking one's mind (whether out loud or silently) should not be taken to constitute an *evidential basis* that one uses to arrive at her self-beliefs (thereby accommodating desideratum 2).¹⁰

Given the rejection of the Epistemic Basis Requirement, however, the account faces the following explanatory task: to account for desideratum 1 by explaining how self-knowledge of thoughts can be due to the distinctive, expressive relation between ISEs and the states of mind they express, as opposed to being due to some epistemic relation between the self-beliefs and the first-order states they are about. Properly construed, the knowledge to be had should turn out to be distinctive and uniquely first-personal in that it is the prerogative of the *subject* of the mental state who is speaking her mind. It is knowledge that one can be said to have in virtue of being in a privileged position to give direct voice to one's thoughts, rather than in virtue of forming self-beliefs on an especially reliable epistemic basis. Providing an account of self-knowledge along these lines would, we believe, allow us adequately to capture the role inner speech plays in explaining the privileged knowledge we have of our own thoughts and other states of mind.

*Philosophy Department
University of Connecticut at Storrs
334 Mansfield Road
Storrs CT, 06269-1054 USA
E-mail: dorit.bar-on@uconn.edu*

*Philosophy Department
University of Connecticut at Storrs
334 Mansfield Road
Storrs CT, 06269-1054 USA
E-mail: jordan.ochs@uconn.edu*

NOTES

¹ Which Byrne describes as ‘peculiar’.

² Byrne characterizes neutrality in two different ways – in terms of the antecedent of the rule making no reference to the rule-follower’s mental states (e.g. p. 111 and 115), and in terms of whether recognizing that the antecedent obtains requires the capacity for self-knowledge (p. 116 and 117). The two ways are not equivalent. This will become important later on.

³ For relevant discussion, see Bar-On (2004a).

⁴ Rules along these lines multiply. Consider:

THINK^{speak}: If the outer voice speaks about x, believe that you are thinking about x.

THINK^{read}: If the words on the page are about x, believe that you are thinking about x.

And so on.

⁵ Roessler also worries that not every case of phonological imagining amounts to a case of inner speech. He considers someone imagining Margaret Thatcher saying to herself ‘there is no such thing as society’ [Roessler (2016), p. 545]. In this case, one has an episode of auditory imagination, but one is not licensed to infer that one is thinking about there being no such thing as society. Instead, one is thinking about someone *else* who has claimed that there is no such thing as society. But Roessler thinks this is a worry that Byrne may be able to address.

⁶ Roessler seems to recognize that his account depends on the availability of a proper explanation of the notion of practical knowledge: “Whether this notion of practical knowledge can be sustained (whether it is possible to make knowledge of what one is doing intelligible in terms of the agent’s intention without offering an account of how she knows what she is doing) are large issues I cannot adequately address in this paper” [Roessler (2016) pp.550].

⁷ On the other hand, if Roessler had to resort to a completely different account of our knowledge of intentions, this would introduce a new bifurcation in his epistemology of self-knowledge – one that would require substantial motivation.

⁸ E.g. Bar-On (2010), (2012), (2015), and Bar-On and Nolfi (2016).

⁹ We address these questions in “Speaking Your Mind in Your Mind: Inner Speech and Self-knowledge” (in progress).

¹⁰ As some interpreters have wrongly assumed – notably, Carruthers (2011).

REFERENCES

- ANSCOMBE, E. (1957), *Intention*, Oxford, Blackwell.
BAR-ON, D. (2015), ‘Transparency, Expression, and Self-Knowledge’, *Philosophical Explorations* (special issue), pp. 134-152.

- (2012) “Expression, Truth, and Reality: Some Variations on Themes from Wright,” in Coliva, A. (ed.), *Mind, Meaning, and Knowledge: Themes from the Philosophy of Crispin Wright*, OUP, pp. 162-192.
- (2010), “Neo-Expressivism: Avowals’ Security, and Privileged Self-Knowledge (Reply to Brueckner),” in Hatzimoysis, A. (ed.), *Self-Knowledge*, OUP, pp. 189-201.
- (2004), *Speaking My Mind: Expression and Self-Knowledge*, Oxford, Clarendon Press.
- (2004a), ‘Externalism and Self-Knowledge: Content, Use, and Expression’, *Nous*, 38, pp. 430-455
- BAR-ON, D. and NOLFI, K. (2016), ‘Belief Self-Knowledge’ in *Oxford Handbooks Online*.
- BYRNE, A. (2011), ‘Knowing That I’m Thinking’ in Hatzimoysis, A. (ed.) *Self-Knowledge*, Oxford, OUP, pp. 105-124.
- CARRUTHERS, P. (2011), *The Opacity of Mind*, Oxford, OUP.
- CASSAM, Q. (2015), *Self-knowledge for Humans*, Oxford, OUP.
- GREEN, M. (2013), ‘Assertions’ in Sbisa, M. and Turner, K. (eds.) *Pragmatics of Speech Actions, Handbook of Pragmatics Vol. II.*, Mouton-de Gruyter, pp. 387-410.
- HAMPSHIRE, S. (1959), *Thought and Action*, London, Chatto and Windus.
- ROESSLER, J. (2016) ‘Thinking, Inner Speech and Self-Awareness’, *Rev.Phil.Psych.* 7, pp. 541-557.
- (2015) ‘Self-knowledge and Communication’, *Phil. Explorations*, 18:2, pp. 153-168.
- RYLE, G. (1987), *The Concept of Mind*, New York, Routledge.
- SOTERIOU, M. (2013), *The Mind’s Construction: The Ontology of Mind and Mental Action*, Oxford, OUP.