

Speaking My Mind

Dorit Bar-On
University of North Carolina, Chapel Hill

INTRODUCTION

Certain self-ascriptions of thoughts, feelings, and sensations, such as "I am thinking about Paris right now," "I am scared of that dog," or "I am very thirsty," appear to have rather unusual and philosophically interesting features. On their face, typical self-ascriptions of occurrent mental states—"avowals" as they are sometimes called—concern contingent matters of fact; they ascribe some particular happenings to a particular individual at a particular time. Semantically speaking, an avowal such as "I am scared of that dog" identifies an individual—the speaker, or thinker—and ascribes to her a property—here, being scared of the dog. And it will share truth conditions with any ascription that identified that same individual and ascribed to her the same property at the same time (e.g., "*She* is scared of the dog," or "*DB* was scared of the dog yesterday"). It can also serve as a premise in humdrum logical inferences, such as "I am scared of the dog, and so are you; so that makes two of us." Avowals thus exhibit continuity of logico-semantic structure and are truth-conditionally interchangeable (in context) with ordinary, unproblematic statements. I call this claim Semantic Continuity.¹

However, epistemically speaking, avowals seem radically different from their semantic close cousins. Avowals are not based on any ordinary observation of, or evidence or inference or memory about, a particular object and

its features or states. Moreover, they are not normally subjected to certain kinds of epistemic assessment. There is a strong presumption that the avowals of a sincere subject are true; we do not expect the subject to have reasons in order for her self-ascription to be warranted; and we ordinarily do not stand ready to correct avowals. Avowals thus exhibit notable epistemic asymmetries with a wide variety of ordinary empirical ascriptions, including all nonmental self-ascriptions. I call this claim Epistemic Asymmetry.

An adequate account of the secure status assigned to avowals should respect both Semantic Continuity and Epistemic Asymmetry. Yet it may be more difficult than it seems to come by an account that does not compromise one or the other claim.² Of particular interest is what I shall refer to as the *epistemic-basis approach* to explaining the security of avowals. On this popular approach, we explain the distinctive security of avowals by identifying a different way of knowing, a special epistemic access, route, or basis associated with avowals.³ The epistemic-basis approach seems tempting when we focus on the non-evidential character of avowals; i.e., their immediacy and lack of reliance on ordinary inference, evidence, or observation. Since avowals appear more secure than semantically equivalent ascriptions that *are* evidential, we are encouraged to suppose that there must be a special epistemic way each of us has for determining certain things about what is currently going on with her.

The epistemic-basis approach is well placed to accommodate Semantic Continuity, for it portrays avowals as representing judgments about the states of a certain individual (oneself), which judgments could be made by others as well. What distinguishes avowals from other ascriptions is simply the fact that they are *arrived at* in an epistemically special, highly secure way.

However, it is not clear how the epistemic-basis approach can accommodate at least one aspect of Epistemic Asymmetry. I have in mind a striking contrast in security, acknowledged by many, between avowals and a certain range of nonmental self-ascriptions. Present-tense proprioceptive reports such as "My legs are crossed" and self-ascriptions of position (e.g., "I am sitting down") and orientation (e.g., "I am in the middle of the room") are as non-evidential as avowals: they are not normally based on inference, evidence, or ordinary observation. In this case it seems very plausible to suppose that there is a distinctive way of knowing associated with such self-ascriptions. After all, no one can tell that my legs are crossed, or that I am standing up, and so on, in the same way that I can, and normally do. Yet avowals seem strikingly more secure than these non-evidential nonmental self-ascriptions, in ways that will be explored below. Once this contrast in security is fully appreciated, it will be recognized that the epistemic-basis approach to avowals requires compromising aspects of Epistemic Asymmetry, so its appeal will seem more limited.

It is important to distinguish in this connection between two different questions one might raise regarding avowals.

- (i) How can avowals, understood as true or false ascriptions of contingent states to an individual, enjoy *unparalleled security*?

This is the question that will primarily concern me in this paper. In answering it, I will be offering a *non-epistemic* account of the special status of avowals; that is, an account that does *not* attribute their distinctive security to the operation of a special way of knowing, or to a special epistemic access or route or basis associated with avowals. However, there is a second question, intimately connected with question (i), that is often of most interest to philosophers studying avowals, namely:

- (ii) What allows avowals to represent items of privileged *self-knowledge*?

Question (ii) directly concerns the epistemic status of avowals, *qua* judgments about one's mental states that constitute items of a special kind of knowledge. It asks what explains the special claim to knowledge subjects have with respect to their present states of mind, which knowledge is articulable in avowals. One may deny that we possess privileged self-knowledge, thus denying the presupposition of (ii), *without* thereby denying that avowals *qua* performances, in speech or in thought, enjoy a special status, in that they are not subjected to ordinary epistemic scrutiny.⁴

The difference between (i) and (ii) will become clearer as I develop my account of avowals. For now, I would like to point out that, by offering a non-epistemic answer to (i), I do not mean to rule out substantive answers to (ii). Since I believe subjects do have a special claim to knowledge of their own present mind, I seek an account of avowals that is at least consistent with a substantive answer to (ii). Developing such an answer requires a properly epistemological investigation of the nature and status of so-called self-knowledge, an investigation which goes beyond my scope here. However, in the final section, I will indicate why I think the account offered in this paper does allow for a substantive account of self-knowledge.⁵

I. THE SECURITY OF SOME "I"-TALK: IMMUNITY TO ERROR THROUGH MISIDENTIFICATION

My concern is with the special security of a subclass of self-ascriptions of occurrent mental states such as thoughts, beliefs, desires, emotions, feelings, sensations—what I am calling "avowals." The strikingly high degree of security that avowals enjoy relative to other ascriptions has led many to describe them as infallible, indubitable, and incorrigible. But even if we find

such descriptions hyperbolic, we may regard the fact that avowals are treated so differently from other ascriptions interesting and worthy of explanation. Of particular interest to me here will be the contrast between avowals, which involve present-tense self-ascriptions of occurrent *mental* states, and certain present-tense *nonmental* self-ascriptions, such as "My legs are crossed," or "I am standing in the middle of a room."

In the normal case, these nonmental self-ascriptions, much like avowals, are not based on evidence or inference or ordinary observation. More crucially, as with spontaneous mental self-ascriptions such as "I am thirsty," or "I think I need to lie down," these bodily self-ascriptions do not seem to require the self-ascriber to ascribe the relevant property to herself based on recognition of some individual *as* herself. Except under unusual circumstances, it makes little sense to raise the question: "*Someone's* legs are crossed, but is it me?" or: "*Someone* is standing in the middle of the room, but is it me?"

On a diagnosis offered by Sydney Shoemaker and Gareth Evans,⁶ the reason for the oddity is as follows. In the normal case, a present-tense ascription of limb position or location to myself, *like* a spontaneous self-ascription of a sensation or feeling or thought, etc., does not involve picking myself out from among potential candidates as the "right" subject of my ascription. When I make such an ascription, I cannot be in error as to who is the individual to whom I ascribe a particular property. Such bodily self-ascriptions do not rely on any substantive identification of oneself and for that reason, like avowals, they are, in Shoemaker's and Evans's terminology, *immune to error through misidentification* (IETM, for short).⁷

Suppose I make an ordinary perceptual judgment that Sheila is sitting on a chair. Presumably, my judgment is based on my recognizing in some way the presence of an individual whom I take to be Sheila, as well as on my perception that this individual is sitting on a chair. I may be victim to at least two kinds of error in this case. First, I may think Sheila is sitting on the chair, when she is really only leaning against it in a peculiar way, or I may think it is a chair she is sitting on, when in fact it is a small table. But, second, I may also think that it is *Sheila* who is sitting on the chair, when in fact it is Dinah (or maybe no one at all!). Now consider my judgment that *I* am sitting on a chair. *In the normal way* of making this judgment (see below), I have no reason, grounds, or basis for judging that *someone* is sitting on a chair, over and above, or separately from whatever grounds my judgment that *I* am sitting on a chair. This means that, although my judgment may well be in error—I may not be sitting on a chair but on a table, or I may not be sitting but squatting—my error will *not* be one of misidentifying *who* it is that is sitting on a chair. This is what it means to say that my judgment is IETM. What protects me from such error is the fact that, in an important

sense (to be explained below), my judgment does not rely on an identification of the judgment's subject, namely, myself.

Two important observations. First, unlike certain other Wittgensteinian proposals,⁸ the diagnosis offered by Shoemaker and Evans for the security of certain "I"-ascriptions does not rely on denying that "I" is an expression whose semantic function in such ascriptions is to refer to a particular individual. Instead, they deny that successful reference always requires an act of recognition, or "thick" identification (as we might put it) on the part of the user of the referring expression.⁹ When I say, or think: "I am sitting on a chair," under normal circumstances, I do not identify myself as the "right" subject of predication to whom I ascribe the property of sitting on a chair. Nonetheless, I am saying, or thinking, something *about* a particular individual. My ascription, in speech or in thought, may still succeed in "hitting" the right target of reference, namely—myself. For, arguably, successful reference does not require "thick" identification.

Second, as both Shoemaker and Evans emphasize, ascriptions are IETM only insofar as they are arrived at in a certain way. We have noted that ordinary empirical ascriptions are not IETM. To vary the example, consider my judgment that John is walking fast. This judgment has a distinct identification component: it relies on the judgment that *someone* is John. Since the judgment constituting this identification component can clearly fail—say, if I mistakenly take Jack to be John—my judgment that John is walking fast is subject to error through misidentification.¹⁰ Importantly, many judgments about myself also involve an identification component. My judgment that I have insufficient funds in my checking account will rely on the judgment that I am the person about whose account I am receiving certain information, say, through the ATM screen. And the same is true for ordinary judgments about my distance from some place, or about my current weight, a bump on my forehead, etc. Even "I"-ascriptions that are IETM when made in the normal way can have semantic equivalents that are *not* IETM when arrived at in an unusual way. Thus, whereas normally "I am about to faint!" will be considered IETM, it will not be IETM, if I say or think it after catching a glimpse of my reflection in the mirror.

What is characteristic of "I"-ascriptions that are not IETM is that, like "John is walking fast," they contain an identification component: a judgment of the form "I = r," where "r" stands for some way that the self-ascriber uses for representing herself to herself as part of her basis for judging that *she* herself satisfies the predicate used in the ascription. For instance, my ascription "I have insufficient funds in my checking account" may rely on my judgment that I am the one whose account information is shown on the ATM screen. Or my judgment that I am about to faint may on an occasion rely on the judgment that I am the person reflected in the mirror.

What renders the "I"-ascription open to error of misidentification in such cases is the fact that it can fail due to the falsity of the judgment comprising the identification component.

Whether or not an ascription is IETM will depend on the self-ascriber's way of gaining knowledge that the relevant property is instantiated. The immunity does not attach to the ascription abstractly conceived. Nor is it guaranteed by the subject's employing a "direct" referential device to pick herself out. In the above cases, my use of the pronoun "I" is "direct," or epistemically unmediated, in the relevant sense: it is not backed up by "thick" epistemic methods for identifying the referent, or ascertaining who it is. However, my judgment that it is I myself who instantiate the relevant property, *is* epistemically mediated. It is based on some kind of observation or evidence or inference and is thus open to potential errors of misidentification.

It is worth noting that to say that an "I"-ascription is *not* IETM is not to say that its use of "I" fails to refer to the self-ascriber. It is a distinctive feature of the semantic functioning of "I" that its uses are guaranteed to have reference. And one would be hard pushed to come up with a convincing case in which one's use of "I" fails to refer to the "right" individual—i.e., to the very individual who produces "I" (in speech or in thought). Indeed, in the above non-IETM cases, success in self-reference is precisely what would allow us to reject the self-ascriptions as false: "*You* have three hundred dollars in your bank account; it's someone else who doesn't have sufficient funds." Or: "*You* are just fine; it's your twin sister who's about to faint." We take the ascriptions to be *false of the self-ascriber*, which presupposes that the self-ascriber successfully referred to *herself* with her use of "I." In making the relevant self-ascription, our subject did refer to herself. But her self-ascription relied on an identificatory judgment: she took herself, erroneously, to be the person whose bank account information was displayed on the ATM screen or whose reflection she caught in the mirror.

By the same token, we should recognize that the immunity of self-ascriptions that *are* IETM goes beyond the guarantee of self-reference associated with the use of an "I" device. Immunity to error through misidentification is a distinctive kind of security. The anti-Cartesian insight offered by Shoemaker and Evans is that this security does not derive from the special ability of the self-ascriber to identify without failure the subject of her ascription, namely, her Self. Rather, it is a matter of immunity to error, which is a negative form of security: it derives from the *absence* of any epistemic effort of identification on the part of the self-ascriber. A self-ascription such as Wittgenstein's famous "The wind blows my hair about" is clearly not absolutely infallible or incorrigible. Nor is it tempting to suppose that it concerns a Cartesian Ego. However, when made in the normal

way, such an ascription does not involve deploying any epistemically "thick" means of identifying oneself as the subject of the ascription. It does not depend on an identificatory judgment concerning oneself (e.g., I am the person who . . .). Thus, if this self-ascription is mistaken, the mistake will not be one of "getting hold" of the wrong subject. This is what it means to say that the self-ascription is immune to error *through misidentification*.

II. THE SPECIAL SECURITY OF AVOWALS

Avowals present one type of paradigm examples of ascriptions that are IETM. Indeed, it is sometimes argued that the immunity to error through misidentification of nonmental "I"-ascriptions must ultimately be inherited from the immunity to error through misidentification of some underlying avowal. If my nonmental self-ascription "The wind is blowing my hair about" is IETM, this is due to the fact that it rests on a mentalistic judgment regarding how, e.g., my scalp and face feel, which is in turn IETM.¹¹ Even if this priority claim is denied, however, it should be recognized that the security of avowals goes beyond the fact that they are IETM. This can be appreciated when we compare avowals with *nonmental* self-ascriptions that are IETM.

Consider again the proprioceptive report: "My legs are crossed," or the self-ascription of relative position: "I am spinning around the table," or the self-ascription of location: "I am standing in the middle of the room." As normally made, these self-ascriptions are IETM. However, they are completely and straightforwardly open to denial or correction by potential observers on the basis of their own observation. Looking at me, you might simply say, "No, your legs are *not* crossed," or "You're not spinning around, you're standing still," or "You're in the corner, not in the middle of the room." In contrast, avowals are not straightforwardly open to correction simply on the strength of an observer's contrary judgment. Note that in both types of cases there is a first-person/third-person contrast between the ways the relevant ascriptions are arrived at. But in the case of the nonmental ascriptions, the observer's verdict typically carries more weight than the self-ascriber's, even though it is rendered on a basis which is causally less direct (i.e., sense perception, as opposed to proprioception, or kinesthetic sense). It is not easy to imagine in the case of avowals anything coming to play a role analogous to that played by your direct observation to the effect that, e.g., my legs are not crossed. I submit that neither the putative telepath nor the expert neuro-scientist would be granted this kind of power simply to trump a subject's avowal, or to cancel the weight it initially carries.

It would be an overreaction to declare that avowals, unlike other "I"-ascriptions, are absolutely incorrigible, as Rorty proposed at one time.¹² For

we do sometimes find reason to question an avowal. It seems perfectly conceivable that even a sincere avowal of pain should be false. Sitting in a dentist chair afraid of what is to come, I may sincerely utter, or think: "Ow! It hurts!" before the instruments properly touch the afflicted tooth. The seasoned dentist could sensibly doubt that my tooth really hurts (and I myself may take back the avowal after a moment). Still, I think it can be argued that avowals enjoy a distinctive form of security. They do not seem open to brute epistemic error; they are especially resilient against epistemic criticism; and they are much less open to doubt or correction than other ascriptions.¹³ All in all, avowals carry a special weight relative to other ascriptions.

Bear in mind, however, that to recognize a distinctive form of security avowals enjoy is not necessarily to accept any particular form of explanation of that security. In particular, it is not to adopt an epistemic-basis approach—one that appeals to a special epistemic basis or some privileged epistemic access associated with avowals. This can be appreciated when we recall the Shoemaker-Evans explanation of the security of certain "I"-ascriptions. Their explanation, we saw, did not proceed by postulating a special epistemic means that each of us can use for picking herself out. Instead, they pointed out that in the case of certain "I"-ascriptions, there is simply *no room* for errors of misidentification, because such ascriptions do not rely on any act of identification. As we shall later see, this idea can be profitably generalized, so as to point the way to a *non-epistemic* explanation of the special security of avowals.

It is common, in this connection, to identify a special privilege of avowing subjects. This privilege, often dubbed "first-person authority," has a very restricted domain and statute of limitation: it applies only to *occurrent mental* states of mine and only to states of my mind that are *contemporaneous* with my ascription.¹⁴ At least as ordinarily conceived, the privilege in question has the following features. First, the privilege is reflexively anchored to self-ascribers. It has to do with the fact that the states ascribed by my avowals are *my own* and are ascribed to me *by myself* as *being my own*. Second, the privilege cannot be inherited by another person. Perhaps we could imagine another person becoming highly reliable at producing non-evidential reports of my current mental states. I submit that we would still not surrender to this "mind-reader" the privilege I have regarding my current mental states. If disagreement broke out between us over what is now going on in my mind, her consistent past success in reading my mind would not be sufficient ground for taking her word over mine. It seems, rather, that we would take the disagreement as signaling the waning of her mind-reading powers. Third, the privilege cannot be easily "detached" from subjects of mental states. Though we can imagine someone producing a false avowal, it is not clear how to give sense to the idea of a subject with

mental states who is "chronically unreliable" in all the pronouncements she makes about her present mental states. It is not clear that we could find reason to doubt everything someone says about what she wants, or intends, for example, without taking at face value at least some of her self-ascriptions of thoughts (in particular, her self-ascriptions of thoughts about what she wants and intends).

We can sum up the above features by saying that the privilege associated with avowals is essentially *first-person* and is both *non-transferable* and *inalienable*.¹⁵ There is no comparable first-person privilege associated with the production of nonmental self-ascriptions. If so, then any account that was geared to assimilate the security of avowals to that of contingently secure nonmental "I"-ascriptions would fail to provide a fully satisfactory explanation of the special status ordinarily reserved for avowals. Such an account would require rejecting Epistemic Asymmetry in its full scope, inasmuch as it would involve denying the appearance of the unique status of avowals and of an associated special privilege we enjoy with respect to them.

If it is accepted that avowals are more secure than other, nonmental "I"-ascriptions that are IETM, then it follows that the security of avowals goes beyond their immunity to error through misidentification. The epistemic asymmetries between avowals and nonmental "I"-ascriptions that are IETM suggest that the special security of avowals must partly come from their *ascriptive* part. When avowing, a self-ascriber is secure not only in the identification of herself as the subject of the ascription, but also in the ascription of the occurrent mental state to herself. And this should help explain why, in contrast even with nonmental self-ascriptions that are IETM, my avowal is not open to straightforward correction.

It may seem natural to take the ascriptive security of avowals to be a matter of subjects' special epistemic access to their occurrent mental states, or else to seek some other special epistemic basis on which avowals are made. This would be analogous to taking the security we enjoy in picking *ourselves* out as subjects of certain "I"-ascriptions to be a matter of a special epistemic relation we have to our Selves. Now, we saw that there is an alternative to this way of seeing things: on the Shoemaker-Evans diagnosis outlined above, the security of picking yourself out as the subject of certain ascriptions is a matter of immunity to error through misidentification. Similarly, I shall propose that the security of ascribing a present mental state to yourself, which is the distinctive mark of avowals, is a matter of *ascriptive immunity to error*.

Notice that, in the normal case, as I say, or think: "I am feeling terribly thirsty," it would seem as out of place to suggest: "I am feeling *something*, but is it thirst?" as it would seem to question whether it is *I* who is feeling

the thirst. In the case of avowals with intentional content, such as "I'm really mad at you," this ascriptive immunity divides into two. "I am mad at *someone*, but is it *you*?" and "I'm in *some* state, but is it being mad?" would both be as odd as "*Someone* is mad at you, but is it *I*?" when I simply *avow* being mad at you (as opposed to making a conjecture about my own state of mind, for example). Similarly for avowals of propositional attitudes, such as "I am wondering whether it's time to leave." "I am wondering *something*, but am I wondering *whether it's time to leave*?" as well as "*Something* is going on in me, but is it a *wondering*?" seem out of place. By contrast, while I may be immune to error in identifying myself as the subject of certain bodily ascriptions, I am not immune to error in ascribing a bodily condition, or location, etc., to myself, even as I use the normal way of gaining that type of information about myself. Both "I am doing *something* with my arm, but am I putting it on the chair?" (as well as "I'm putting something on the chair, but is it my arm?") and "I am putting my arm on *something*, but is it a chair?" could make perfect sense, as I think to myself: "I'm putting my arm on the chair" in the normal way.

To say that it would be out of place to raise the above-mentioned questions in the case of avowals is only to deny that the relevant self-ascriptions are subject to certain kinds of epistemic error regarding the ascriptive part. It is *not* to say that these self-ascriptions are guaranteed success, in the sense of being absolutely infallible. As we shall see later, an avowal *can* be false. Moreover, the ascriptive immunity to error offered here by way of explaining the security of avowals does not even issue in absolute incorrigibility. Observers may sometimes have reasons to suspect the falsity of an avowal, and offer a correction. The strong presumption of truth governing avowals can thus be overridden. To say that an avowal is immune to ascriptive error is only to say that, if it fails, the failure is not an epistemic one. It is to deny that, when an avowal fails, its failure should be seen as a result of the self-ascriber's mistaking one state (or content) for another or as the failure of an epistemic mechanism in the self-ascriber which is designed to track the presence and character of mental states in him. The analogous point regarding self-ascriptions that are IETM was that, when they fail, the failure is not one of the subject misidentifying herself, say, by mistaking someone else for herself. In both cases, the idea is that the special security, which manifests itself in protection from epistemic challenge, criticism, doubt, or error may not be due to a guarantee of epistemic success, but may rather be due to there being no room for epistemic failure.

We can perhaps see the distinctive security of avowals as a joint product of the immunity to error through misidentification of the subject part *and* immunity to error in the ascriptive part. In the case of intentional avowals, the latter immunity will in turn break down into immunity regarding the presence of the occurrent mental state and immunity regarding the state's

intentional object/propositional content (henceforth "intentional content," for short). This combined immunity can serve to mark the difference between avowals, on the one hand, and nonmental "I"-ascriptions that are IETM, on the other. Whereas the latter may enjoy immunity to error through misidentification of their subject, the former enjoy, in addition, immunity to error in the ascription of the relevant state to the subject (and of the relevant intentional content when it has one).

As observed earlier, the immunity to error through misidentification associated with certain "I"-ascriptions is a consequence of the fact that such uses do not rely on any "thick," or substantive identification of oneself as the subject of the "I"-ascription. Nonetheless, we may insist that such uses of "I" do involve genuine semantic reference to a particular object (oneself). Similarly, I propose to see the ascriptive immunity to error of avowals as a consequence of the fact that avowing does not involve deploying any epistemic procedure for ascertaining the presence or character of the ascribed mental state. And still avowals can be seen as involving genuine ascriptions. In avowing, a subject ascribes to herself being in a certain mental state—e.g., feeling tired, or thinking about the weather. But the ascription need not be seen as underwritten by, nor does it require, recognition or detection of the presence of the ascribed state or of its intentional content. To the extent that the avowing subject is not seen as using any epistemic means for recognizing the presence of the mental state (or its content), her avowal will be regarded as immune to epistemic errors of ascription. Being immune to ascriptive errors, avowals will emerge as epistemically different from other ascriptions, nonmental IETM "I"-ascriptions included. Since they do, however, involve genuine ascriptions of mental conditions to the self-ascriber, they can, at least for some purposes, be interchanged with other ascriptions that share truth conditions with them. We can thus capture Epistemic Asymmetry while preserving Semantic Continuity.

III. THE SECURITY OF SELF-VERIFYING "I"-ASCRPTIONS

Avowals are immune to ascriptive error, I am claiming, insofar as they involve no use of epistemic means for recognizing the presence (or character) of the ascribed mental state. The question now arises how, in the absence of such epistemic mediation, a person avowing a mental state can still succeed in genuinely ascribing a mental state to herself. If it is true that avowals as such are not made on the basis of some kind of recognition or determination of the presence and character of the avowed condition, how can the avower be genuinely saying or thinking *that* she is in such-and-such mental condition?

Consider a special case of avowals, exemplified by the Descartes-inspired self-ascription: "I am thinking about being deceived by an evil demon." Understood as ascribing to me an entertaining of a thought, rather than a positive affirmation, such a self-ascription is maximally secure, for it cannot fail on any of the scores we have identified so far. In making it, I cannot misidentify the subject of my ascription; I cannot fail to ascribe to myself some thought or other that is crossing my mind; and I cannot fail to ascribe to myself a thought *about being deceived by an evil demon*. Such an ascription is *self-verifying*—it makes itself true.¹⁶ This is because any time I ascribe to myself the entertaining of a thought with a certain content (not in the sense of making some judgment, or assertion) I, *ipso facto*, entertain a thought with that content. The very act of making the thought ascription summons up the ascribed thought, as it were; it inevitably involves the thought "passing through" my mind, thereby rendering my ascription true. Self-ascribing a presently entertained thought using "I" is in this respect like deciding to think about the thought. In the normal way of doing it, you cannot decide to think *about* a particular thought without thereby entertaining that very thought.¹⁷

But now consider self-ascriptions of intentional states other than the pure entertaining of thoughts: "I believe that John is innocent," or "I hope she'll get here soon," or "I wish that noise would stop," and so on. Here, the mere act of ascribing to myself the relevant intentional state does not secure the truth of the ascription. Still, such self-ascriptions enjoy at least one kind of security regarding their intentional content: they seem protected against skepticism generated by content externalism.¹⁸

Briefly, externalist theories of content claim that the contents of our thoughts are determined by our relations to various objects and events in our environment. Whether one has *water* thoughts, as opposed to *twater* thoughts (i.e., thoughts about stuff that is superficially similar to but chemically different from water, which is to be found on twin earth) will depend on whether one stands in the appropriate causal relation to the liquid water. If this is so, however, it would seem that we can be no more privileged in telling the contents of our thoughts (and other intentional states) than we are in telling what determines those contents. And we are often in error in our beliefs regarding the nature and presence of objects, events, and substances in our environment.

Externalists have responded to this form of skepticism by arguing that, whatever is involved in (so-called) authoritative knowledge of the contents of our thoughts, it does not require knowledge of the causal determiners of our thoughts. Knowing that you are having a *thought about* water cannot require being able to distinguish whether you are on earth or twin earth; nor does it require that you know the chemical constitution of the stuff you refer

to by the term *water*. This would be asking too much of knowledge of content. Thus, suppose, as externalism claims, a necessary condition of my having intentional states involving the concept of water is that I have had some causal contact with water (H₂O). Still, as I self-ascribe: "I believe there is water in that glass," it cannot be an open possibility for me that perhaps my belief is about *twater* (XYZ), rather than plain old water. Notice that, unlike in the case of an ascription to someone *else* (or to myself at other times), my assignment of intentional content to my own present states is bound to draw ingredients from the same conceptual pool as the ascribed intentional states themselves. If my first-order intentional states cannot involve *twater*, since I've never been in contact with XYZ, then my self-ascriptions of intentional states cannot involve *twater*, either. We can sum this up by saying that one way in which a subject cannot go wrong with an intentional self-ascription is by ascribing to herself an intentional content which is not in her cognitive repertoire. To this extent, the contents that all articulated intentional self-ascriptions assign to one's intentional states are "locked onto" the contents of the first-order states they ascribe.¹⁹

On a plausible understanding of the way the pronoun "I" functions to pick out its referent, successful reference to oneself does not depend on one's ability to identify oneself in any "thick" epistemic sense—e.g., to ascertain that one fits certain identifying descriptions. Relaxing the epistemic requirements on successful reference can go a long way toward explaining the special guarantee of reference enjoyed by uses of the pronoun "I" and its equivalents in speech or in thought. The idea is that one can succeed in referring to oneself even when one is in complete ignorance of, or error about who one really is (or what one's real nature is). Given this feature of "I," as we saw earlier, "I"-ascriptions can be genuinely *about* the self-ascriber, even when they involve no identificatory judgment of the subject of ascription, and are thus immune to error through misidentifying the subject of the ascription.

The above account of how content assignment in self-ascriptions is protected from a certain kind of skepticism can be understood along similar lines. The externalist's claim is that success in ascribing to oneself a present state with a certain intentional content does not depend on one's ability to establish that she meets conditions that are (metaphysically) necessary for having states with that content. The externalist can go on to claim that, since associating content with her intentional state in no way requires the self-ascriber to engage in an epistemic effort to determine the causal sources of her concepts, her intentional self-ascriptions are immune to errors as to whether they are about, say, *twater* rather than water. Furthermore, the externalist could insist that this immunity to errors in content assignment is consistent with the idea that intentional self-ascriptions genuinely ascribe to

subjects intentional states with specific contents. (This would be analogous to the claim that immunity to error through misidentification is consistent with the idea that the relevant "I"-ascriptions are genuinely about the subject of the ascription.)

However, we should realize that this does not yet afford us a full explanation of the ascriptive immunity of avowals, for that security goes beyond protection from content skepticism. In ascribing to myself the desire for a glass of water, say, I seem secure not only as regards its being water that I want, but also as regards it being a *desire* that I have, rather than a belief, or some other water-involving attitude. So it is not only my assignment of content that seems secure, it is also my ascription of the particular state. This can also be appreciated when considering that so-called phenomenal avowals—i.e., avowals of thirst, hunger, pain, etc.—seem equally, if not more secure than intentional avowals.²⁰

Even as regards the assignment of intentional content, the security I enjoy in avowing an occurrent intentional state goes beyond the above-described protection from "twin-alternatives." When I avow a desire for a glass of water, it is not just that I seem able to tell securely that it is water, *rather than twater* that I want; I can also securely tell it is water I want *rather than gin*. This kind of security pertains to my ability to tell which of various intentional contents that *are* in my cognitive repertoire is the content of my avowed intentional state. And it seems specifically characteristic of one's intentional avowals (as opposed to more "alienated" intentional ascriptions). When I avow an intentional state, it does not seem that the correct intentional content of the self-ascribed state could remain up for grabs, as it might when I interpret someone else's utterances. It is not as though, having somehow recognized that I am in a condition of believing, or being angry, or wanting, I independently need to determine what intentional content to assign to the ascribed state. And it seems wrong to think of my position as one of having to pick out the intentional content of my state by somehow recognizing that it is the right intentional content to associate with it. Thus, although intentional avowals in general are not self-verifying, they seem to share with self-verifying ascriptions a certain epistemic directness in the assignment of intentional content.

I think (though I will not argue this here) that this epistemic directness tells against taking the special status of avowals to stem from the fact that they represent highly secure judgments about one's mind—judgments that are arrived at on an especially secure epistemic basis or through the operation of a very reliable tracking mechanism. By contrast, on the view defended here, the distinctive security of avowals is negative: it is a matter of immunity to certain kinds of epistemic errors, which is due to the fact that the self-ascriber does *not* do certain things in arriving at the ascription. If there

is no room for mistaking what intentional content to assign to the ascribed state, this is because the assignment of intentional content does not involve deploying a procedure for ascertaining that this is the right content to assign to one's intentional state in the first place. If, more generally, it is not up for grabs what mental condition the avowing subject is in, this is because (or to the extent that) the ascription is not made on the basis of recognizing the presence or character of the condition in her.

IV. ASSIGNING CONTENT: ARTICULATION AND EXPRESSION

How is it, then, that we do tell what content our intentional states have? We should note that, as regards their intentional content, intentional avowals share the following feature with self-verifying ascriptions of entertained thoughts: in avowing an occurrent belief, hope, preference, etc. I typically "latch onto" the content of the ascribed state directly, by *articulating* it in the very act of avowing. A typical avowal of a hope will involve "going through" the relevant state's propositional content (in speech or in thought): "I hope *she gets here on time*." A typical avowal of a desire will involve mention of its intentional object: "I want *a cup of tea*." This, in contrast with oblique self-ascriptions such as: "I am hoping for the same thing I did yesterday," or "I want what you want," where the relevant intentional content is not spelled out in the self-ascription.

Explicit articulation should be recognized as a legitimate way of assigning intentional contents. Its use is not restricted to avowals. We commonly assign intentional content to others' intentional states by displaying or spelling out the assigned content ("She thinks that it's raining outside." "He's annoyed by that noise. "). When doing so, we put to explicit use elements from our own cognitive repertoire so as to latch onto the content of their states. We latch onto the content not by identifying it in some indirect way (as in "He is thinking about what I was just thinking") but rather by *expressing* it directly ("He is thinking about the dinner he's about to miss. "). Of course, when we ascribe an intentional state to another, observers may straightforwardly challenge the correctness of our ascription—they may question on the basis of their own assessment whether the person is in the state ascribed, or whether the content assigned is the right content to assign to her state. But in the case of avowals our self-ascriptions are normally taken at face value; one is typically presumed to be in a privileged position to tell whether one's state has this rather than that content, as well as whether it is a state of belief, hope, desire, fear, that one has, and so on. This is just one aspect of Epistemic Asymmetry which the present account seeks to capture.

Putting together the fact that avowals in general are not self-verifying with the fact that they are nonetheless more secure than other ascriptions may seem to force on us an account that appeals to a distinctive epistemic basis on which avowals are made. On the one hand, as we have noted, when I avow: "I hope Emma will come soon," the mere articulation of the hope's content does not provide apriori guarantee of the truth of my self-ascription. This may suggest that, in the general case of non-self-verifying avowals, more epistemic work is required on the part of the self-ascriber in order to succeed in making a true self-ascription. One who avows a particular hope, it may be thought, does need to have some way of determining her current condition to be that of hoping, rather than, say, believing; she may also need to have some epistemic method for ascertaining that the state has the content that she spells out in the self-ascription, rather than some other content available to her. She can at least in principle go wrong, and this explains why intentional avowals in general are not self-verifying. On the other hand, since intentional avowals still seem relatively more secure than ascriptions of similar states to others, as well as more secure than bodily self-reports such as: "I'm sitting on a chair," it may seem that we must invoke an especially reliable epistemic mechanism (though perhaps not an introspective one) for finding out about our own intentional states.

This is a line of thought that I wish to resist. I think we can explain the distinctive security of non-self-verifying avowals without appeal to an especially secure epistemic basis or especially reliable epistemic mechanism, by extending the idea of immunity to error to aspects of self-ascriptions that go beyond content assignment. Though avowals in general are not self-verifying, when they go wrong, it is *not* because of the subject's failing to recognize correctly the ascribed state. To reiterate, if there is no room for my mistaking the content *or* character of the state I am in when I avow being in the state, this is so to the extent that I do not engage in any epistemic effort to determine what state I am in, or what its content is, in the first place. Avowing as such does not rely on the deployment of any ("thick") epistemic procedure for ascribing an occurrent state to oneself. To that extent, there is a crucial similarity between the relative security of all avowals, on the one hand, and the absolute security of self-verifying avowals.

Consider again the self-verifying "I am thinking that I may be deceived by an evil demon" (where "thinking" is used in the sense of a thought crossing one's mind, not in the sense of judging, or affirming). The explanation of the self-verifying character of this self-ascription of thought appealed to a certain "reflexive" element—the fact that it involves an explicit articulation, in speech or in thought, of the very content assigned to the self-ascribed state. When the state ascribed is simply one of entertaining a certain thought, explicit articulation of the thought suffices for making the

self-ascription true; for it requires the subject of ascription to undergo the very mental process she ascribes to herself, namely, the entertaining of a thought with that content. All I need to do, epistemically speaking, in order to succeed in truly ascribing to myself the entertaining of the thought *that p* is to entertain the thought *that p*, which I am bound to do if I ascribe the thought explicitly, by displaying its content. No substantive procedure of recognizing the presence of the state in me, or of identifying its content, is necessary. Nor is there any need to use some other kind of epistemic basis for determining what state one is in. When I avow a presently entertained thought, there is no need for me to have any epistemic access to the ascribed state or its content, and no need to make it an epistemic target.

By contrast, when I ascribe an entertained thought to someone else, the state the other is in is an epistemic target for me. This is so even if I do not make the ascription on the basis of hypothesizing on the internal causes of the other's behavior, but rather do it more directly, by projecting onto him an explicitly articulated thought of my own. For, in general, projection is still an epistemic method designed to enable us to determine what mental state the other is in, and what content it has (if any). Whereas when I ascribe an entertained thought to myself, the articulation of the thought obviates the need to use *any* epistemic method for determining what state I am in. In this case, the articulation does not merely serve to demonstrate the content one is prepared to assign to the ascribed state by displaying it, as it does in the case of ascription to a subject other than oneself. Rather, it serves *directly to express—in the sense of giving voice to—the self-ascriber's condition of entertaining the thought.*

In his "Language as Thought and Communication,"²¹ Sellars distinguishes three senses of expressing: the *action* sense, in which a *person* intentionally expresses a condition of hers by "venting" it, or "giving it voice" (in my terminology), the *causal* sense, in which an *utterance* or piece of behavior expresses an underlying condition by being the culmination of a causal process beginning with that condition, and the *semantic* sense, in which e.g., a *sentence* expresses an abstract thought or judgment by being a (conventional) representation of it. Expressing regret, say, in the first two senses requires expressing *one's* regret, which is factive: I (or an utterance of mine) cannot be said to express regret without my feeling regret. Expressing regret in the semantic sense does not require expressing *one's* regret. In this same sense, the linguistic locution: "Regretfully *p*" expresses regret, though on a particular occasion, it may not be used for—or may not succeed in—expressing the *utterer's* regret.

Simple declarative sentences of English can be seen as conventional vehicles for the (semantic) expression of thoughts. It is for this reason that, in the special case of self-ascribing an entertained thought that *p*, giving

semantic expression to *the thought that p* suffices for giving expression to *one's thought that p*. This is because one cannot articulate an entertained thought without the articulation being causally connected to the entertaining of that very thought. The articulation will be the culmination of a process (however brief) that begins with the entertaining of the thought, or it will at least involve the entertaining of the thought, in the course of articulating it. One's self-ascription, then, will inevitably be an expression of the entertained thought in Sellars's causal sense. And inasmuch as one's articulation is intentional, one will also be assured success in expressing *his* condition of entertaining that thought in Sellars's action sense.

The above case of self-verifying thoughts illustrates the following idea. Articulating, or spelling out, the content of a self-ascribed state, which constitutes expressing it in the *semantic* sense, can amount to giving voice to an intentional state one is in. That is, it can constitute giving it expression in the *causal* and *action* senses. I now want to suggest that this idea may be generalized beyond the special case of self-verifying avowals, to all intentional avowals. Consider first avowals of beliefs. Saying (or thinking) "I believe John is angry with me" is not self-verifying. Still, if this self-ascription is an avowal, the point of making it seems to be at least in part to "vent" my (first-order) belief—that John is angry with me. This can be attested to by the anomaly of so-called "Moore-propositions"—pronouncements such as "I believe that John is angry with me, but John is not angry with me." There is no overt or formal contradiction in a pronouncement of this sort. But seeing the first part of the pronouncement as an avowal allows us to identify a certain kind of conflict. Being an avowal, the first part does not merely ascribe to me the belief that John *is* angry with me, but serves directly to express my belief that this is so. Since sincerely asserting the second part also expresses my belief—in this case that John *is not* angry with me—we get what we may describe as expressive dissonance or conflict.²²

More generally still, I think we should recognize that all intentional avowals, whether of beliefs, or of hopes, desires, angers, fears, and so on, play (in part) an expressive role as described above. A typical self-ascription such as "I hope John will come," or: "I'd really like some water!" or "I'm mad at you, Mom" will play a role played by more direct expressions of one's own intentional states, such as saying, or thinking: "John will come, won't he?" *hopefully*, or "Water, please!" *eagerly*, or "Mom!!" *angrily*, and so on. The point of articulating the state's intentional content in avowing is directly to express one's self-ascribed intentional state in the very act of avowing. (Speaking of point here should not be taken to suggest that there is a deliberate plan on the part of the self-ascriber to use the avowal for a certain purpose.)

We can now see why intentional avowals are more secure than intentional ascriptions to others as well as being more secure than alienated self-

ascriptions of intentional states (e.g., ascribing to myself hatred of my brother on the basis of therapy). On the present expressive approach, the point of avowing an intentional state is not so much to provide a descriptive report of it as to “vent” it. What is distinctive of acts of avowing an intentional state, as opposed to reporting its presence in oneself or in others, is not that the self-ascription is made on a distinctive epistemic basis. To use earlier locutions, such acts do not require the self-ascriber to traverse any epistemic distance, in order to arrive at the self-ascription. When avowing an intentional state, the ascribed state is in no way an epistemic target for the self-ascriber. It is for this reason that epistemic errors of intentional ascription are crowded out. That is to say, however an intentional avowal can go wrong, its failure will not be due to mistaking *either* the subject of the ascription *or* the ascribed state. In this, intentional avowals are different both from intentional ascriptions to others and from nonmental self-reports that are immune to error through misidentification.

We can also see why not all intentional avowals are self-verifying. In the special case of avowing an entertained thought, being in the relevant state, viz., the state of entertaining the thought is guaranteed by the spelling out of the thought whether to others or to oneself, whether in language or in some other medium or representation. So in that case, giving (semantic) expression to *p* suffices for success in expressing *one's* thinking that *p*, and thus in making a true self-ascription. But there is more to being in a state of hoping *that p*, wishing *for x*, being afraid *of y*, etc., than the spelling out of the relevant propositional content or intentional object. So, while the point of articulating the content may still be to express one's first-order intentional state, the articulation does not guarantee that one will succeed in expressing one's hope *that p*/wish *for x*/fear *of y*, etc. The truth of the self-ascription is not automatically secured; it is not self-verifying.

V. AVOWALS AS EXPRESSIVE ACTS

So far I have extended the expressive account from self-verifying self-ascriptions to intentional avowals more generally. We are now in a position to generalize the account even further. As we saw earlier, not all avowals involve the articulation of content, e.g., “I’m so uncomfortable!”, “I’m really thirsty.”; “I feel very achy.” Nonetheless, identifying an expressive aspect of all avowals will allow us to understand why all avowals—whether intentional or phenomenal—are treated as more secure than ordinary empirical reports, including mental ascriptions to others, and explanatory (“alienated”) ascriptions to oneself, on the one hand, but also all bodily self-ascriptions, on the other hand. My proposal is that *avowals in general*,

and not just intentional avowals, *are expressive acts in which the subject gives vent to—or directly speaks from—a mental condition.*²³

My strategy in leading up to this proposal so far has followed a very different path from that taken by the more traditional expressive approach. On the approach often attributed to Wittgenstein, we begin with the idea of a close similarity between, e.g., holding one's cheek and wincing, on the one hand, and saying: "I have a toothache." (Emitting: "Ow!" can serve as a nice intermediate case.) We are then invited to think of avowals of sensations and emotions such as pain, hunger, fear, anger, as being on a par with *natural expressions* of such states—moaning and groaning, trembling, bearing teeth. An avowal of pain, just like a moan, is not a vehicle for expressing one's *judgment* that one is in pain. Instead, it serves to replace the natural expression—it is a way of moaning.²⁴ On a common understanding of this comparison, an utterance such as "I am in pain" turns out to be incorrigible for the same reason a moan is incorrigible. If we take it to be protected from error, this is so only in a degenerate sense. For, like a moan, such an utterance is simply in the wrong "grammatical" category for being subject to *any* kind of epistemic or semantic assessment.

Granting this kind of "immunity by grammar" to avowals would come at a high price. Crucially, it would require denying that avowals are genuine, truth-evaluable ascriptions of mental conditions to particular individuals. This means they cannot have truth-conditional equivalents with which they can be legitimately interchanged in certain contexts, and they cannot serve as legitimate premises in logical inferences. In short, the traditional idea of protecting avowals from ascriptive error through grammar requires buying Epistemic Asymmetry at the price of Semantic Continuity. As stated at the outset, I think we must adhere to Semantic Continuity. Bearing this in mind, my own attempt to motivate an expressive account of avowals began with a case as far from moans as groans as seems possible—the case of a fully articulate self-ascription of an entertained thought. I now want to offer an understanding of the expressive idea which is applicable to intentional and non-intentional avowals alike and which does not compromise the logico-semantic continuity between avowals and other ascriptions.

I see a mean-looking large dog. I am overcome by fear. My fear can be expressed through my agitated behavior, by my exclamation "That dog is scary!" or by my utterance: "I am really scared of that dog!" *Semantically speaking*, these are three very different performances. The first involves a natural, non-linguistic expression of the fear. It is linguistically inarticulate; it makes no reference to dogs, or being scared, or anything. The other two involve linguistic utterances that, however, say different things. The second ascribes a certain quality, being scary, to the dog. The third one says that a certain individual, me, is in a certain condition, being scared of the dog.

However, in a given context, the three performances can be on a par, *epistemically speaking*. In particular, though the self-ascription ascribes to me a certain present mental state, it need not be epistemically underwritten—it need not require my employment of any epistemic means for detecting the presence of that state in me, or otherwise determining that I am in the relevant state. The self-ascription can simply be my way of giving direct expression to that state. It is this insight of the traditional expressive view that I think is worth preserving.

On the “Neo Expressive” approach I advocate, however, the comparison to natural expressions is appealed to in a rather limited way. We can summarize the comparison by saying that avowals are like natural expressions in certain epistemic respects, but quite unlike them semantically. Avowals may be seen as epistemically unmediated: they are typically not offered or received as an upshot of the subject’s epistemically-based opinion that she is in the relevant condition. Rather, they resemble natural expressions in being direct upshots of the subject’s first-order mental condition. Thus, my spontaneous pronouncement (in speech or in thought), “I’m so happy to see you” can be regarded as epistemically on a par with my exclamation: “It’s so good to see you!” or with my cheerfully hugging you. For, these acts can all be seen as ways of directly giving voice to my joy at seeing you, rather than being upshots of my forming a higher-order judgment that I feel joy on this or that epistemic basis.²⁵ To the extent that, when I issue my self-ascription, I can be seen as speaking from my first-order condition, rather than as reporting my findings about it, my expressive self-ascription will share in the epistemic directness of the other, non-self-ascriptive expressions. In the case of avowals, we can recognize a certain lack of epistemic distance between avowing subjects and the conditions they avow. Hence, the epistemic similarity to natural expressions, and the epistemic asymmetries with other ascriptions (including mental “self-reports”).

However, avowals are crucially different from natural expressions, semantically speaking. Natural expressions express the subject’s condition by displaying it. Whereas the “ascriptive” expressions do not *as such* show us the relevant condition. An ascriptive expression—an avowal—of joy does not show the joy as a cheerful voice or hug might. Rather, it *tells* us something. What the avowal tells us, specifically, is that someone—the self-ascriber—is feeling joy. For all that, sometimes a telling can replace, or play the role of, a showing. To use previous terminology, the articulate self-ascription expresses *the subject’s* first-order condition—in the *causal* and *action* sense—by giving *semantic* expression to the relevant claim, namely, *that* the subject is in such and such condition. The semantically articulate expression ascribes to the subject a mental state. This is what allows it to be interchangeable (in context, for some purposes) with other ascriptions.

This can perhaps be made plausible, if we consider how linguistic self-ascriptions could come to play this kind of role. Think of a small child, Jenny, eagerly reaching for a teddy bear. Her behavior is a natural expression of her desire for the toy. It is not related to the desire simply causally, the way a reliable indicator is related to its cause, or a natural sign—say, smoke—is related to the fire it signals. Rather, the naturally expressive behavior is something that is *read* a certain way. As we see Jenny's reaching behavior, we take it to reveal to us directly both the presence and the character of her condition. It is as though we could see the condition through the behavior, in a way *somewhat* similar to the way we see the iceberg by seeing its tip. (This is, of course, not to say that expressive behavior cannot be faked or simulated, or that we do not need to acquire it or learn how to read it.) Having thus read the behavior, we are in a position to begin offering Jenny alternative, more socially tamed, *linguistically articulate* ways of expressing her desire. We might first say "That's teddy!" allowing Jenny verbally to embellish her reaching behavior by uttering: "Teddy!", thus articulating the intentional object of her desire. And later, we might say: "You really want that teddy, don't you, Jenny. Can you say: 'I want the teddy'?", whereupon she becomes able also to give verbal expression to the kind of condition to be expressed, not just to its intentional object. At least some of Jenny's subsequent self-ascriptive utterances "I want the teddy" can be seen as resembling her original reaching behavior in that they serve to "give vent" to her very desire. The verbal self-ascription can come to supplement or supplant the natural expression.

Not all acts of mental self-ascriptions are expressive in this way. We have seen that one can come to ascribe a desire to oneself on a purely theoretical basis—say, as a result of what one's therapist has told her, or by applying to oneself the findings of cognitive science experiments. Such "alienated" self-reports belong in a different category from avowals, and do not share in their special security. Furthermore, even within the category of avowals, we may recognize avowals that play a dual role. In response to your question: "How are you feeling?" I may say, wanting to provide you with the information you are requesting: "I am feeling very nervous." Here it does not seem that my avowal *only* serves directly to express my nervousness. It also serves to report my present condition, in response to your request. Yet we would still consider the avowal ascriptively secure.²⁶ The present proposal, however, is that even in such a case, what provides the avowal with protection from epistemic challenge or criticism is the fact that it plays an expressive role, rather than the fact that it constitutes a well founded or highly reliable self-report. If we do not question a present tense mental self-ascription, if we do not expect it to be supported by any evidence or reasons whatsoever, if we strongly presume it to be true provided

only that it is sincere, this is so *to the extent that* we regard it as taking us—as would a natural expression—directly to the subject's mental condition, without any epistemic mediation.

To sum up, when avowing, a subject gives voice to her condition; she speaks from it. When speaking from her condition, a subject is immune to error through mistaking the ascribed condition, as well through the misidentification of herself as the right subject of the ascription. The subject no more uses epistemic means to determine the presence of a mental condition in her than she uses such means to pick herself out as the subject who is in that condition. In this way avowals resemble natural expressions. In both cases, unlike in the case of ordinary ascriptions, many self-ascriptions included, there is in a sense no epistemic distance between the ascriber and the condition ascribe. Hence Epistemic Asymmetry. However, the ascriptive immunity of avowals is not the purely grammatical immunity of grunts and winces, which prevents them from being true *or* false. Rather, it is due to the fact that, when a subject avows, no space opens up for *epistemic* error in ascribing a mental condition (or in assigning it intentional content). Semantically speaking, avowals, unlike natural expressions, are true (or false) sayings about contingent matters of fact. They involve referring to particular individuals and ascribing to them present mental conditions. This, despite the fact that the ascription may not be arrived at through any kind of ascertaining of the presence or character or content of the relevant condition in the right subject. For, as we have seen, substantive recognition or detection of the presence of a mental state (or of its correct content) is no more required for ascriptive success than substantive recognition of oneself as the appropriate subject of the ascription is required for successful reference to oneself. Thus Semantic Continuity is preserved.

Earlier I suggested that we can see the privileged status of avowals as a joint product of the immunity to error through misidentification in the subject part and the ascriptive immunity. It should now be appreciated, however, that these immunities cannot be ultimately pulled apart in the case of avowals. It is not the case that there is, on the one hand, the security of speaking (or thinking) about oneself, and, on the other hand, the security of pronouncing on the condition one is in. On the expressive account developed here, when I issue an avowal, I am telling you things *about* my mind by simply *speaking my mind*. If this is so, then there is an important sense in which the immunity to error through misidentification of the *subject* of the avowal is parasitic on the immunity to error through mistaking the *subject's condition* (and its intentional content, when it has one). It is because, when avowing, you are speaking your mind—i.e., giving voice to a condition of yours—that you have no need to identify, for yourself or for others, the subject of your self-ascription. It is because you have no need to employ

epistemic means for detecting the presence or character of the condition you (semantically) ascribe to yourself that you have no need to employ identificatory means for latching onto yourself as the right individual about whom you want to assert or affirm something.

VI. AVOWALS AS TRUTH ASSESSABLE

It is often objected that, if avowals are seen as expressive of the mental conditions mentioned in them, rather than as being assertoric reports about those conditions, then avowals *cannot* be regarded as true or false sayings. *Ipsa facto* they cannot be things subjects can be said to *know*, let alone know with any special authority. I would like to address this objection more directly now. On the Neo Expressive account presented here, avowals play the following expressive role: they serve to express a subject's particular first-order mental state, rather than²⁷ her higher-order judgment about the presence and character of a first-order mental state. That avowals play this expressive role is not a fact about the grammar or semantic content of the self-ascriptions involved in avowals, but rather has to do with their use in the given situation. An utterance or thought can have the content of a self-ascription in that it can be true if and only if a particular subject (the avower) is in the state referred to by the relevant mental term, and at the same time not serve to express the subject's own *judgment* that she is in that state. In other words, the grouping of utterances or thought-episodes into the category of avowals is not a grammatical or semantic grouping; it is a grouping in terms of the type of act performed (in speech or in thought).

A familiar comparison may help here. Suppose I utter: "I promise to take you to dinner tonight." I am using a sentence that is true if and only if DB promises to take [so-and-so] to dinner [on a certain evening]. In uttering the sentence, I will be *saying that* I promise to take the addressee to dinner—for I intentionally utter a sentence that has that content. But, in the typical case, I will not be reporting a promise; I will be making it. Or, if I say: "I intend to kill you," my utterance will be true if and only if at the given time DB has the intention of killing the person addressed. But the utterance may very well constitute a declaration of my evil intention, rather than a report of the presence in my mind of that intention. Similarly, suppose I utter: "I am really thirsty," as I desperately reach for water. I will be using a sentence that has the grammar and meaning of a self-ascription. On the Neo Expressive view, the sentence will be true under the same conditions as the comment: "You are really thirsty," which you may make upon watching me gulp down a whole cup of ice water. Still, in uttering the sentence I may not be reporting my thirst condition but rather expressing it.

On a now standard view, the semantic content of an utterance of an indicative sentence—what the utterance says—is at least in part a matter of its truth condition. The utterance's having a particular content, then, entails that it is *truth assessable* or truth evaluable. Truth assessability is a matter of grammar and meaning. But it should be clear that an utterance can be truth assessable without it being thereby an assertoric report. In the case of avowals, as in the familiar cases of so-called other moods, grammar and meaning do not line up with force or use.²⁸

The charge that any expressive account of avowals must fall afoul of their truth assessability may be due to a false comparison to ethical expressivism. For example, the emotivist view in ethics claims that, when someone says: "This is good," she is merely commending it or expressing approval or some other "pro-attitude" toward it.²⁹ The relevant contrast here is between expressing one's subjective preference or pro-attitude, and objectively describing an individual or object, an act, or a state of affairs. The expressivist claim is in aid of an ontological denial that there *are* such properties as goodness, justice, and the rest, which properties can be thought objectively to inhere in things, independently of our approvals. By contrast, it is by no means part of the Neo Expressive account of avowals to deny that there are such things as being thirsty, being in pain, thinking of Paris, etc. Of course there are such conditions; otherwise we could not express them! In virtue of their semantic content, avowals ascribe these conditions to the subject they refer to (namely, the avower). What is essential to an expressive view of avowals, however, is the insistence that the special security of avowals comes from the fact that they serve directly to express a first-order condition, rather than a higher-order judgment about it.

Going back to Jenny, as she utters: "I want the teddy!" we may plausibly regard her as saying that she, Jenny, wants the teddy. For, assuming she has become a competent user of all relevant terms, she will presumably be intentionally using these words, which have that content as their literal meaning. This means that Jenny's utterance can be seen as expressing the judgment that Jenny wants the teddy, but only in the *semantic* sense. *Qua* an avowal, I have urged, we should not see Jenny's utterance as expressing *her* judgment that she, Jenny, has the desire for the teddy, in either the causal or the action sense. Instead, we should see it as directly "giving vent" to her desire through the use of an articulate expressive vehicle, a self-ascription.

Now, I have emphasized that, except in the special limit case of self-verifying self-ascriptions, the ascriptive immunity to error of avowals affords no absolute guarantee of truth. On the present view, though an avowing subject may not have her condition as an epistemic target—though she may simply be giving it verbal expression—it is still possible for her to say something false. Let us return to an earlier example. I sit in the dentist's

chair. Having a long history with dental work, I dread what is to come. The dentist puts a sharp-looking instrument in my mouth, and I wince, or grunt. If only I could speak, I might say (something which I may in fact *think* to myself): "Ow, that tooth!" or, more explicitly: "My front tooth hurts so much!" Here, I am speaking from my condition, in the sense that I am venting, rather than arriving at the judgment that my tooth hurts through introspective reflection. But for all that, I may not be suffering a pain in my front tooth; so my explicit self-ascription would be false. *Something* has gone wrong in that case, though it does not seem to be a case of mistaking one state for another through some kind of failure of recognition.³⁰

What has gone wrong, I suggest, is that, assuming I am not in pain, my expressive act has failed: I have expressed *pain* without succeeding in expressing *my pain*. I have used an expressive tool that is referentially associated with the condition of having a toothache (viz., the verbal expression: "tooth hurts") to give vent to a condition which is *not* one of having a toothache. This separation between expressing a (type of) condition without expressing *one's* condition should be familiar from cases of dissimulation. It is not unique to ascriptive self-ascriptions, either. The liar uses a semantic expression of the belief *that p* (uttering: "It's going to rain later," or "I think it's going to rain later") as though it expressed *her* belief *that p*. A child may pretend to feel hurt to gain sympathy or attention, and an actor on stage can effectively fake natural expressions of pain—grimaces, grunts, etc.—to (falsely) convey her being in pain. We are also familiar with simple cases of misspeaking oneself: saying "Jack is here!" to express one's belief that John has arrived. To this familiar array of mismatched expressions we can now add cases where one speaks from a condition that is not of the kind associated (naturally, or semantically, or contextually) with the expressive means she has used. Avowing as such involves speaking from a condition, with no epistemic targeting of it; but the expressive means used to express one's condition may fail to be associated (through meaning or context) with the condition one is in. When this is so, the self-ascription made will be false.

It is important to acknowledge that avowals can be false, since we do sometimes find occasion to reject an avowal, as the above examples show. But it is also important to make room for false avowals of the kind described above, if we are to preserve the distinction between self-verifying self-ascriptions (e.g., "I am thinking about being deceived by an evil demon," as an avowal of an entertained thought passing through my mind) and other avowals. Put in terms of expression, the difference between the self-verifying and the non-self-verifying cases is as follows. Whereas an avowal of an entertained thought *that p* is necessarily an expression of *one's thought that p*, an avowal of a belief or hope or desire *that p* is *not* neces-

sarily an expression of one's belief/hope/desire *that p* (and similarly for non-intentional avowals).³¹ For, in the general case, it is possible to make use of a tool that (semantically) expresses a mental condition without being in that condition, and thus without expressing *one's* condition, where the "expressive mismatch" is not due to an intention to deceive, or dissimulation, or act-playing.³²

Avowals, then, *can* be false. However, a central mark of their unique status is that they are governed by a strong presumption of truth. An avowal will ordinarily be presumed to tell us correctly what condition the subject is in (although the presumption can, as we saw, be overridden). On the view defended here, the distinctive strength and character of this presumption can be explained by the fact that avowals involve speaking from one's mental conditions, rather than offering descriptive reports of them. Unlike other ascriptions, including "theoretical" mental self-ascriptions and relatively secure bodily self-ascriptions, avowals share in the epistemic directness of first-order expressions of mental states, whether verbal or natural.

VII. EXPRESSIVE AVOWALS AS POTENTIAL ITEMS OF SELF-KNOWLEDGE³³

I have portrayed avowals as expressions of the avower's first-order mental states. But avowals as such do not use expressive means that display or exhibit the mental states; instead, they employ as expressive vehicles genuine, truth-assessable self-ascriptions, in speech or in thought. When avowing, I express my mental condition not by showing (a glimpse of) it, but rather by calling its name, as it were, and articulating its intentional content when it has one. Avowals, then, are *ascriptive expressions* of one's present first-order mental states; they give *articulate* voice to one's sensations, thoughts, occurrent beliefs and wants, and so on. Seeing avowals this way, I maintain, allows us to explain Epistemic Asymmetry without sacrificing Semantic Continuity. It allows us to answer the question with which we began:

- (i) How can avowals, seen as true or false self-ascriptions of contingent states, enjoy an unparalleled security?

As I indicated in the introduction, however, answering (i) does not by itself suffice for answering

- (ii) What allows avowals to represent articles of *privileged self-knowledge*?

Question (ii) concerns the special claim to *knowledge* subjects have when avowing their states of mind. Any non-deflationary answer to this question must supply an account of what makes avowals especially apt to represent judgments that are both true and epistemically warranted. And it may be thought that my non-epistemic answer to (i) precludes my accepting any substantive, non-deflationary answers to (ii). In this final section, I want to explain why I do not think that this is so.

The Neo Expressive account presented here can straightforwardly accommodate the *truth* aspect of a special claim to self-knowledge. When avowing, I have argued, subjects issue genuine self-ascriptions that are governed by an especially strong, though overridable, presumption of truth. An avowal of, e.g., a belief, a fear, or a sensation, can be strongly presumed to be true insofar as it can be presumed to take us to the subject's relevant first-order condition (believing *that p*, being scared *of x*, feeling a sharp pain) as directly and as securely as would e.g. the subject's sincere statement *that p*, his pointing *at x* while trembling, or his wincing. No other type of ascriptions is governed by such a presumption. Furthermore, *non-self-ascriptive* expressions of one's mental states (whether natural or verbal) are not covered by this same presumption of truth. Such expressions are not *self-ascriptions*; *ipso facto* they are not apt to be *true* self-ascriptions, and so are in no way candidates for self-knowledge.

However, it may seem that the Neo Expressive account would not fare as well when it comes to epistemic warrant. If we took avowals to represent judgments subjects arrive at on a distinctive and especially secure epistemic basis, then they could easily be seen as possessing a special kind of justification. I have resisted this kind of epistemic approach to the special security of avowals. Instead, I have presented their security in terms of the negative notion of immunity to epistemic errors, both in the identification of the avowal's subject and in the ascription. How could it then be maintained that there is a special epistemic warrant associated with avowals?

Suppose it is held that a knowledge claim can only be justified if the person making the claim is engaged in some epistemic doing or effort, such as drawing an inference or observing or judging on some basis or even just attending to the relevant phenomenon. Then it would follow that the Neo Expressive view excludes seeing avowals as justified. However, it seems unreasonable to expect that every knowledge claim should be backed up by justification in this sense. Indeed, by that standard, it is not clear that many of our beliefs could constitute knowledge; consider basic perceptual beliefs, typical memory claims, and so-called *apriori* judgments. Possession of knowledge requires having epistemic warrant. But justification in the above sense is only one kind of epistemic warrant. There may be others.

For example, a person who makes pronouncements in the absence of

justification in the above sense may still have a special *entitlement* to those pronouncements, in the sense of having a special epistemic right to make those pronouncements.³⁴ As suggested by the other cases of knowledgeable-but-not-justified beliefs just mentioned, it is fairly clear that we ordinarily make room for epistemic warrant beyond justification. The idea behind the notion of entitlement would be to take up this slack, by identifying a form of epistemic warrant that is not bound up with the idea of a cognitive effort, or an epistemic doing on the part of the entitled subject.³⁵ An account of self-knowledge couched in terms of special entitlement could avoid commitment to there being a special epistemic route leading subjects to their avowals. Instead, it would seek to ground the special entitlement in other ways. Outlining these other ways, and assessing their adequacy in providing for privileged self-knowledge has not been my task here. But my account of avowals does not close off these avenues.³⁶

It might be thought that since, on my view, an avowal does *not* express a subject's higher-order judgment that she is in some mental state, there can be no room for regarding the subject as entitled to the judgment. This would be too hasty. Granted, I have denied that the avowal will represent a "cognitive achievement" on the part of the subject, at least inasmuch as there is no epistemic targeting of the truth she conveys about her present condition through the avowal. The avowal as such is *not* "the upshot of something which the person can do, or try to do, an effort of some kind which she may make in order to achieve [self-knowledge]."³⁷ However, I have also maintained that avowals do have the semantic content of self-ascriptions. An avowal conveys a truth about the subject's condition by *telling* it like it is. It says that *S* (the self-ascriber) is in *M* (a particular mental state). The avowal thus expresses that judgment *semantically*, though it does not express the *avower's* judgment to that effect (in either the causal or the action sense of "express"). We can, then, sensibly raise a question regarding the epistemic relation between the subject *S* and the judgment that her avowal (semantically) expresses. In particular, we can ask: Does *S* have a special entitlement to that judgment?

As far as I can see, the Neo Expressive view of the special status of avowals does not rule out a positive answer to this question. The view precludes seeing the special status of avowals as due to their expressing judgments that are arrived through a special way of knowing about one's own present mental states. Instead, it requires seeing it as due to their immunity to epistemic error, where the immunity has its source in their expressive character with its attendant lack of epistemic distance. Now, a *deflationist* about self-knowledge could maintain that *all there is* to self-knowledge is the ability to produce self-ascriptions that are immune to error in the ways outlined here. But it is important to recognize that the view I have defended

does not *require* us to take this deflationary route. For, it does not rule out the possibility that we do possess a special entitlement to the judgments semantically expressed by our avowals.

The account advocated in this paper maintains that avowals typically play a role similar to that of non-self-ascriptive expressions in that their primary contextual function is to express first-order conditions rather than higher-order judgments about those conditions. On this view of avowals, when avowing, we tell things about our minds by speaking our mind. To the extent that, when avowing, we can be seen as speaking directly from the mental conditions, rather than issuing reports about them, our avowals are immune to epistemic errors. And they can be seen as especially apt to be true, in the sense that they reveal what conditions we are in as directly (epistemically speaking) as would non-self-ascriptive expressions of those conditions. This should explain why avowals are treated so differently from other ascriptions.

We can certainly find room in this account for a special first-person privilege. Expressing my first-order mental conditions is what as the *subject* of these conditions I am uniquely placed to do. I can only inform you about someone else's state of mind by conveying my belief about it. But in my own case, I can do it also by speaking my mind. Speaking my mind is something I am in a unique position to do. Only I can express or give voice to *my own* current states of mind by speaking my mind—i.e., by issuing semantically articulate self-ascriptions. And it is only states of my *mind* that I can express (in the relevant, non-semantic sense). For bodily conditions, such as being of a certain height, or having a bent arm, or a cold, or a stomach ulcer, are not conditions one can speak from. I can speak my mind, but I cannot speak my body.

First-person privilege understood in this way is tied to being a subject who *has* mental states and who is a competent user of ascriptive vehicles for the expression of those mental states. It is not tied to possession of a special way of knowing (though, of course, mastery of ascriptive expressions may itself involve various epistemic achievements). However, it may well be that having first-person privilege in the above sense—i.e., being in the privileged position to express my own mental states using self-ascriptions—affords me a special entitlement to the judgments that are semantically expressed by those self-ascriptions. Perhaps it could be argued that this entitlement is different from any entitlement I may have to any other ascriptions I make. The special entitlement would derive from the fact that I am in a unique position

to proclaim the very states—the thoughts, hopes, wants, pains, etc.—that move me in thinking and in acting, at the same time as I ascribe those states to myself. If this is so, then the account of avowals advocated here could make room for genuinely privileged self-knowledge. For it would allow us to maintain that, when I speak my mind, not only do I speak from my conditions, but I also know what I am talking about.³⁸

NOTES

My thanks to Simon Blackburn, Doug Long, Bill Lycan, Keith Simmons and especially to Bijan Parsia, for comments on drafts of this paper. Earlier versions of this paper were presented at the College of William and Mary, the University of South Carolina, and the Canadian Philosophical Association (May 2000).

1. Unless otherwise specified, my discussion is intended to apply equally to self-ascriptions made in speech and in thought. I shall therefore use as much as possible the relatively neutral notion of *ascription* (as opposed to *utterance*, on the one hand, or *thought/judgment*, on the other). In keeping with common practice, I shall allow myself to speak of the “semantics,” as well as “truth conditions” of ascriptions, as well as of the “I” component of an ascription and its “reference.” I shall also use quotation marks to cite the content of the relevant ascription, without presupposing that self-ascriptions in thought must consist in one’s *talking* to oneself. Where the differences between ascribing in language and ascribing in thought matter, I shall indicate it.
2. For discussion of how leading accounts currently on offer fail to respect one or the other claim, see Dorit Bar-On and Douglas C. Long, “Avowals and First-Person Privilege,” in *Philosophy and Phenomenological Research* 62 (March 2001): 311–35; and Dorit Bar-On, *Speaking My Mind: On the “Grammar” and Epistemology of “I”-Talk* (in progress). Focusing on the difficulties in accounting for what is here called Epistemic Asymmetry, Crispin Wright points out (in “Self-Knowledge: The Wittgensteinian Legacy,” in *Knowing Our Own Minds*, ed. Wright, Smith, and Macdonald [Oxford: Clarendon Press, 1998], 13–46), the importance of preserving features grouped here under Semantic Continuity. In “On Knowing One’s Own Language” (op. cit. 391–428), Barry Smith argues that the main difficulty in reaching a satisfactory philosophical account of so-called first-person authority is to reconcile (what is here called) Epistemic Asymmetry with Semantic Continuity.
3. The epistemic-basis approach so described encompasses a wide range of views, and is in no way restricted to traditional Cartesian introspectionism. For discussion, see Bar-On and Long, “Avowals and First-Person Privilege,” sec. 1 and 2. See also Elizabeth Fricker, “Self-Knowledge: Special Access versus Artefact of Grammar—A Dichotomy Rejected” (in Wright, Smith, and Macdonald, eds., *Knowing Our Own Minds*, 155–206).
4. There is an affinity between my two questions and the two sub-tasks Fricker identifies for a philosophical account of self-knowledge in section 2 of “Self-Knowledge: Special Access versus Artefact of Grammar—A Dichotomy Rejected.”
5. Bar-on and Long’s “Knowing Selves: Expression, Truth, and Knowledge” (in *Privileged Access: Philosophical Accounts of Self-Knowledge*, Brie Gertler, ed., Ashgoode Epistemology and Mind Series, forthcoming, July 2002) attempts to answer question (ii).
6. Sydney Shoemaker, “Self-Reference and Self Awareness,” *Journal of Philosophy* 65

(1968): 555–67; Gareth Evans, *The Varieties of Reference*, ed. John McDowell (Oxford: Clarendon Press, 1982).

7. Both authors refer to the following, much-cited paragraph from Wittgenstein's *The Blue and Brown Books* (Oxford: Blackwell, 1958) in which he draws the relevant distinction:

There are two different cases in the use of the word "I" (or "my") which I might call "the use as object" and "the use as subject." Examples of the first kind of use are these: "My arm is broken," "I have grown six inches," "I have a bump on my forehead," "the wind blows my hair about." Examples of the second kind are: "I see so and so," "I hear so and so," "I try to lift my arm," "I think it will rain," "I have a toothache. One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these cases the possibility of an error, or as I should rather put it: the possibility of an error has been provided for. . . . It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbour's. And I could, looking into a mirror, mistake a bump on his forehead for one on mine. On the other hand, there is no question of recognizing a person when I say I have tooth-ache. To ask "are you sure it is *you* who have pains?" would be nonsensical. (1964: 66f.)

8. Most notably, Anscombe's, in her celebrated "The First Person" (in *Mind and Language: Wolfson College Lectures, 1974*, ed. Samuel Guttenplan [Oxford: Oxford University Press, 1975]). Reprinted in Quassim Cassam, ed., *Self-Knowledge* (Oxford: Oxford University Press, 1994).
9. For the distinction between two senses of identifying an object, see Evans, *The Varieties of Reference*, 7.2. Both Evans and Shoemaker offer demonstratives such as "this" and "that" as examples of genuinely referential devices that do not require "thick" identification of an object. A person can be thinking about a particular object, as in: "*That* thing is moving very fast," without having to think of the object as a such-and-such, and certainly without possessing uniquely identifying information about it. If we wonder how the thought can succeed in "latching onto" a particular object in such a case, we should bear in mind that in such a case there are typically various *non-epistemic relations* between the person and the relevant target of reference. For example, she is able to locate it in space, or track it as it moves, and so on.
10. Note that when this is so, my judgment that *someone* is walking fast can still survive; for it does not depend on my judgment that that someone is John. Not so with judgments that are IETM.
11. For a recent source, see Wright, "Self-Knowledge: The Wittgensteinian Legacy," 19f.
12. See his "Incorrigibility as the Mark of the Mental," *Journal of Philosophy* 67:12 (1970): 399–424.
13. By "brute epistemic error" I mean error that is simply due to the world failing to cooperate, rather than being due to some kind of failure on the subject's part. The contrastive model here is of perceptual errors, in judging, e.g., that there is a cup in front of me, which can occur even when all my faculties are working perfectly and my use of language is in perfect order. For some discussion, see Burge, "Our Entitlement to Self-Knowledge," *Proceedings of the Aristotelian Society* 96 (1996):1–26 (see especially pages 13ff).
14. The use of the label "first-person authority" is very prevalent, as can be evidenced by a cursory look at the articles in three recent collections, *Self-Knowledge* (Cassam, ed.), *Knowing Our Own Minds* (Wright, Smith, and Macdonald, eds.), and *Externalism and Self-Knowledge* (Peter Ludlow and Norah Martin, eds. [Stanford, Calif.: CSLI Publications, 1998]). In "Avowals and First-Person Privilege," Bar-On and Long introduce the notion of "first-person privilege," which I shall use below. The notion of author-

ity is typically understood in epistemic terms, whereas it seems desirable to leave room for a non-epistemic account of the relevant privilege.

15. I am here in part glossing the following remarks by Wright in "Self-Knowledge: The Wittgensteinian Legacy": "the authority which attaches to [avowals] is, in a certain sense, *inalienable*. There is no such thing as showing oneself chronically unreliable in relation to . . . avowals" (17); and later: "the authority I have over the avowable aspects of my mental life is not transferable to others: there is no contingency . . . whose suspension would put other ordinary people in a position to avow away on my behalf, as it were" (24). See also Bar-On and Long, "Avowals and First-Person Privilege," sec. 2.
16. On the self-verifying character of such thoughts, see Burge, "Individualism and Self-Knowledge," in Cassam, ed., *Self-Knowledge*, 65–79.
17. See Brian Loar, "Subjective Intentionality," *Philosophical Topics* (spring 1987): 89–124.
18. For much discussion, see the recent volume *Externalism and Self-Knowledge* (Ludlow and Martin, eds.). I discuss content externalism at length in "Externalism and Self-Knowledge: Content, Use and Expression" in S. Nuccetelli and U. Yekini, eds. *New Essays on Semantic Externalism, Skepticism, and Self-Knowledge*, Rowman and Littlefield (forthcoming).
19. This line of defense of externalism against content-skepticism draws on Burge, "Individualism and Self-Knowledge" and "Our Entitlement to Self-Knowledge"; Davidson, "First-Person Authority," *Dialectica* 38:2–3 (1984): 101–11; and "Knowing One's Own Mind" (in Cassam, ed., *Self-Knowledge*, 43–64), and John Heil, "Privileged Access," *Mind* 97 (1988): 238–52.

Burge sums up this line of defense as follows:

[I]n [these] cases there is no possibility of counterfeits. No abnormal background condition could substitute some other object in such a way as to create a gap between what we think and what we think about. Basic self-knowledge is self-referential in a way that ensures that the object of reference just is the thought being thought. If background conditions are different enough so that there is another object of reference in one's self-referential thinking, they are also different enough in that there is another thought. The person would remain in the same reflexive position with respect to this thought, and would again know, in the authoritative way, what he is thinking. ("Individualism and Self-Knowledge," 75)

20. On some views, all mental states, phenomenal states included, are intentional in the sense that they have representational content. I am not certain, however, that such views cannot recognize a significant difference between intentional and non-intentional *avowals*. Even if phenomenal states have representational content, it still seems true that avowals of phenomenal states are different from intentional avowals in that the former do not involve any *specification* of the state's intentional content.
21. Reprinted as chapter 5, 93–117, of Wilfrid Sellars, *Essays in Philosophy and Its History* (Dordrecht, Holland: D. Reidel Publishing Co., 1974). (The distinction is presented on page 109.)
22. For a congenial discussion of Moore's paradox, see Moran, "Making Up Your Mind: Self-Interpretation and Self-Constitution," *Ratio* NS 1 (1994): 137–51. See also his "Self-Knowledge: Discovery, Resolution, and Undoing," *European Journal of Philosophy* 5:2 (1997): 141–61.
23. For a development of this proposal which partially overlaps with the remainder of this section, see Bar-On and Long, "Avowals and First-Person Privilege," sec. 4. Here I am offering my present take on the proposal (which is the topic of chapters 4–7 of Bar-On, *Speaking My Mind: On the "Grammar" and Epistemology of "I"-Talk*).
24. For exposition and references, see P. M. S. Hacker, *Wittgenstein: Meaning and Mind*, vol. 3, pt. 1, (Oxford: Blackwell, 1993). See also Wright, "Self-Knowledge: The Wittgensteinian Legacy," sec. 9.

25. The expressive character of such non-doxastic avowals can also be attested to by the anomaly of combinations such as: "I like this painting a lot, but it's horrible," or "I am excited about this meeting, but it's very boring," or "I feel so hot" uttered (or thought) as one is shivering and rubbing one's arms. Like Moore-sentences such combinations, exhibit expressive dissonance or conflict.
26. In Bar-On and Long, "Avowals and First-Person Privilege," in 4.1–4.2, we distinguish in this connection between "avowals proper," which are purely expressive, and "reportive avowals," which also serve non-expressive purposes.
27. Or maybe sometimes in addition to (as suggested above).
28. For discussion, see Rockney Jacobsen, "Wittgenstein on Self-Knowledge and Self-Expression," *Philosophical Quarterly* 46:182 (1996): 12–30.
29. The *locus classicus* for the emotivist view is Alfred Ayer, *Language, Truth and Logic* (Gollancz, London, 1936), chap. 6. For some introductory discussion, see Simon Blackburn, *Spreading the Word* (Oxford: Clarendon Press, 1984), 167ff.
30. A similar case is discussed in the final section of Bar-On and Long, "Avowals and First-Person Privilege." The diagnosis offered there, however, is somewhat different from the one I offer here.
31. I am here setting aside the possibility of *uttering* a false self-ascription of a presently entertained thought due to a slip of the tongue.
32. And even when no slip of the tongue, or its equivalent in thought (if there is such) is involved.
33. For development and modification of ideas sketched in this section, see Bar-on and Long, "Knowing Selves: Expression, Truth, and Knowledge."
34. Compare Burge: "I take the notion of epistemic warrant to be broader than the ordinary notion of justification. An individual's epistemic warrant may . . . also be an *entitlement* that consists in a status of operating in an appropriate way in accord with norms of reason, even when these norms cannot be articulated by the individual who has this status" ("Our Entitlement to Self-Knowledge," 3).
35. Arguably, the familiar reliabilist approach in epistemology is one attempt to do just that. The problem with invoking this approach in the present context, however, is that it is not clear how a reliabilist could ground distinctively *privileged* self-knowledge. In other words, I suspect that a reliabilist answer to (ii) would end up having to adopt a fairly deflationary view of self-knowledge, for much the same reason that the reliabilist answer to (i) would end up compromising Epistemic Asymmetry. Even so, it is useful to recognize that one could adopt a reliabilist epistemology of self-knowledge while accepting a non-reliabilist account of the privileged status of avowals (like the one offered here).
36. In recent years, Tyler Burge has offered a Kantian derivation of a special entitlement we have to self-knowledge. (See, e.g., "Our Entitlement to Self-Knowledge," and "Reason and the First Person," in *Knowing Our Own Minds*, ed. Wright, Smith, and Macdonald.) Burge's subtle and intricate discussion is mainly devoted to self-ascriptions of states that play a role in critical reasoning. But I believe parallel considerations to ones he offers could be used in connection with self-ascriptions of other mental states. (I should mention, though, that Burge thinks the case of sensations requires separate treatment—see "Our Entitlement to Self-Knowledge," footnote 13.)
37. See Fricker, "Self-Knowledge: Special Access versus Artefact of Grammar—A Dichotomy Rejected," 173, for this characterization of the denial that self-knowledge involves any cognitive achievement. (Fricker offers a second characterization as well.)
38. This paper is based on a manuscript I am currently completing, entitled *Speaking My Mind: on the "Grammar" and Epistemology of "I"-Talk*. The manuscript motivates and develops in detail a non-epistemic account of avowals, certain aspects of which are presented in Bar-On and Long, "Avowals and First-Person Privilege."