

Externalism and Skepticism: Recognition, Expression, and Self-Knowledge

Dorit Bar-On, UNC-Chapel Hill

According to *external-world skepticism*, were I now to think: There's water in this glass, I wouldn't *know* that there is water in the glass, even if there was water in the glass, my eyes were wide-open, etc. This is because, for all I know, my way of telling what is in front of me does not allow me to rule out the possibility that I am only under some kind of illusion about what is in front of me. Surely, though, I know that I'm now *thinking* that there's water in the glass? According to *content skepticism*, I do not. This is because, for all I know, my way of telling what I am thinking does not allow me to rule out the possibility that I am only under some kind of illusion about what I am thinking. Yet, as commonsense would have it, my ordinary belief about what I am presently thinking is remarkably secure – at least much more secure than my ordinary beliefs about my extra-mental world. My first aim in this paper is to examine whether the commonsense confidence about ordinary knowledge of content can be sustained in the face of a skepticism that proceeds by analogy to external-world skepticism. It might seem that we could get the right analogue, by enlisting the doctrine known as *content externalism* (externalism, for short). But I will argue that, for the analogy to work, the external-content skeptic must rely on a certain *recognitional conception* of our ordinary knowledge of content in particular and of ordinary self-knowledge more generally. I think there are good reasons to reject this conception. And my second aim will be to sketch an alternative to it, which would allow us to sustain the idea that ordinary self-knowledge is not threatened by a skepticism that is analogous to external-world skepticism.

1. External-Content Skepticism Meets External-World Skepticism

Suppose I now say, or think:

EW: There is water in the glass.

For purposes of raising her doubt, the external-world skeptic will allow me, at least initially, to suppose that I do know that I now have a thought with the content of EW. Still, she invites me to question whether EW

represents something I actually know.¹ Full symmetry would require the content skeptic to invite me to doubt that I now have a thought with content p while at the same time holding fixed, or at least not questioning, my belief that p . But it is not clear how I am to proceed. How can I confidently affirm that there is water in the glass, even as I doubt that I am now thinking that there's water in the glass? Am I to think: "There's definitely water in the glass. But, for all I know, I may not now be *thinking* that there's water in the glass"? This seems like a version of Moore's paradox (viz., " p , but I am not thinking that p "). Moreover, just as I cannot doubt p without thinking that p , so I cannot affirm p without thinking that p ; and this too is something I can easily recognize. Explicitly entertaining a doubt about whether some external-world proposition p , which I believe is true, as well as expressing confidence that p is true, are both inconsistent with simultaneously entertaining a doubt about whether I am thinking that p . But this suggests that content skepticism cannot be entertained in a fully explicit form that parallels the way (Cartesian) external-world skepticism is entertained. Each of us can internalize the external-world skeptic's challenge, and pose it in the form: as I am thinking that p is the case, for all I know, p may not be the case. But it seems we cannot rationally pose to ourselves the content skeptic's challenge in a parallel form, viz.: as I am thinking that I am thinking that p , for all I know, it is not the case that I'm thinking that p .

It is also worth noting that the commonsense conception of the content of a mental state does not readily furnish us with skeptical alternatives. It takes some fairly sophisticated philosophical stage-setting to bring us to appreciate the kinds of alternatives that would support a skeptical doubt concerning what we are currently thinking. This is precisely what the view called *externalism* (or anti-individualism) about content is supposed to have provided. The thesis of content externalism says, roughly:

EXT: The contents of our thoughts (and other mental states) depend for their individuation on the nature of our physical or social environment. Whether an individual is in a mental state with one

¹ An extreme skeptic could of course later turn around and question my knowledge of the content of my thought, as well.

content rather than another depends in part on relationships between that individual and her extra-mental environment. ²

Now, the commonsense view of our knowledge of the contents of our mental states can be summarized as follows:

CSK: Normally, when we ascribe to ourselves present contentful mental states (in speech or in thought), our self-ascriptions do not rely on observation, evidence, or inference. Nonetheless, our self-ascriptions are especially secure and privileged, and they represent genuine knowledge we have of the content of the relevant states.

Though CSK does not imply absolute infallibility, incorrigibility, or self-omniscience regarding content, it does imply that our pronouncements regarding what we think, want, hope, etc. carry special weight and represent some kind of privileged knowledge.

EXT is often claimed to be incompatible with CSK. Suppose I make the following Self-Ascription of a Contentful state

SAC: I am thinking that there's water in the glass.

For SAC to constitute something I *know* requires not only that it be *true* that I am thinking there's water in the glass, but also that I be *warranted* in holding it true. If EXT is true, then it seems that, to ascertain the truth of SAC, it is necessary to verify that I stand in the right relation to my extra-mental environment. Yet according to CSK, my own self-ascription of the thought doesn't rely on *any* kind of study or observation. So how could I possibly be warranted in my self-ascription? ³

2. Is Externalism Compatible with Self-Knowledge?

However, assuming that the alleged problem for CSK that is based on EXT is *not* supposed simply to follow from a more general skeptical problem about extra-mental knowledge, it is not as yet clear what the

² The *loci classica* are Putnam (1975) and Tyler Burge (1979) (which applies the externalist thesis to mental contents).

³ Following Davies in Wright *et.al.* (1998), who distinguishes between the "achievement problem" for content externalism which he says "does not pose any insuperable achievement problem for first-person authority" (p. 342), and "the consequence problem" which he thinks is more difficult. I address the latter problem elsewhere (in 2004a) Although I agree with Davies that what he calls "the achievement problem" is not insuperable, I think solving it takes some doing.

problem is. Granted, if externalism is true, contents have essential features – a kind of nature, if you will – that are not transparent to us purely in virtue of our being thinking subjects; they admit of theoretical understanding. But if I can normally know I am drinking water, even in ignorance of chemistry, why is it that I cannot know that I am *thinking* that I am drinking water, even in ignorance of the correct theory of thought content?

In general, for any range of truths, there may be different ways of knowing those truths, not all of which involve a theoretical probing into the underlying nature of the known items. Still, even in the absence of such theoretical understanding, if someone is said to know some truth *p*, it is because she has *some* positive way of telling that *p* is the case. And that *seems* to imply that she must go by some features of the relevant situation that are somehow correlated – and reliably so – with the obtaining of *p*. Thus, if I can be said to know that the man I see is George, even though I have conducted no DNA tests, this is because visual recognition represents a reliable method for correctly identifying people. Similarly for knowing that there's water in the glass, or that a shirt in front of me is red. But if externalism is true, how can there be *any* telltale signs by which we could immediately recognize what our own thoughts are about?⁴

However, even if it's true that, to know some truth *p*, we must have some reliable epistemic access to the worldly conditions that render *p* true, it is not clear that this access must afford us *counterfactual discrimination*. If there is no twater, but only water in my environment, can I not know that there is water in the glass, even though *had* I been confronted with a glass of twater, I *might* have been fooled? If so, then why could I not know that the thought I am having is that there's water in the glass, even though, had circumstances been different, I might have been fooled into thinking that that's what I was thinking?

Note that skepticism about the external world can proceed without trading on an implicit counterfactual discrimination requirement. The Cartesian external-world skeptic invites us to consider alternatives that may *in fact* obtain for all we know, ones that (so the challenge goes) we are unable to rule out.

⁴ The present suggestion has obvious affinities to what is known as epistemological reliabilism, which is usually presented as an externalist view of knowledge. It has often been observed that, to avoid the charge of incompatibility with self-knowledge, content externalism can (perhaps must) be coupled with epistemological externalism. I have avoided the label 'reliabilism' here, since I am not certain that the proposal I am working with here qualifies as a purely externalist proposal. See below, Section 3.

Such an inability to rule out alternatives is supposed by the skeptic to undermine our claims to know external-world propositions, insofar as it shows a diminished capacity for discrimination in actual, and not merely counterfactual, situations. And such diminished capacity does seem to bear on epistemic warrant. If I am to have perceptual knowledge that this man is George, or that the liquid in my glass is water, for example, I must be able to tell George apart from other individuals I know, and water from other liquids, respectively, through perception. Such discriminative abilities are necessary if I am to *know which* individual is George, or which liquid is water, which in turn seems to be a precondition of my being warranted in thinking that this man here is George, or that the liquid in the glass is water.⁵ (Of course, how good my ability to discriminate must be in order for me to be warranted is open for debate. The safe answer for our present purposes would be: good enough so that it is not a sheer coincidence that I hold the relevant claim true, but not so good that it would allow me to rule out counterfeits in all imaginable situations.) As it happens, I am quite good at discerning George by his looks, voice, manner, etc., and am quite adept at telling water apart from other liquids by its color, smell and taste (or lack thereof). But now suppose that George in fact has an identical twin brother; or suppose twin-earth is discovered, and we begin to import twin-earth's twater (XYZ) to earth. My world in such cases encompasses actual candidate counterfeits. If I still cannot perceptually tell George apart from his twin, or water apart from twater, could I nevertheless be warranted at times in thinking that I am looking at George, or that I am drinking water? Arguably not. If so, then the external-world skeptic could legitimately mount a challenge to our knowledge of the external world by trying to offer actual alternatives that we cannot rule out.⁶

To preserve the analogy with external-world skepticism, then, the external-content skeptic would have to challenge content self-knowledge as described by CSK without relying on the counterfactual discrimination requirement. The external-content skeptic needs to persuade us of the following: As I ascribe a thought with a certain content to myself, there are in fact, though perhaps unbeknownst to me, actual

⁵ Russell proposed that this "knowing which" requirement must be satisfied by anyone who is able even to *think* about an individual. For a construal of the requirement in terms of an ability to discriminate the object of judgment from other things, see Evans (1982, Chapter 4).

⁶ Whether the dream or the evil demon/brain-in-the-vat hypotheses qualify as legitimate such alternatives can of course also be debated. For relevant discussion, see Goldman (1976).

counterfeits of that thought; i.e., thoughts I may now be having which differ in content, and which I could not tell apart from the thought I take myself to be having. This would be a situation in which, given how things are in the world I live in, I could be fooled into thinking, or could be merely under the impression, that I'm having a particular thought. But this is not as easy to accomplish as in the case of external-world skepticism (where one can trade on familiar phenomena such as dreaming, or hallucination, to generate the threatening possibility). Specifically, the external-content skeptic needs to describe a situation in which I do ascribe to myself a *water*-thought, but where it is possible, for all I know, that I am instead having a *twater*-thought. The trouble is that, if my present situation is one in which there is (and has always been) only water around, then I can only have *water*-thoughts, and so I cannot go wrong by ascribing to myself a *twater*-thought. And any situation in which I fail to meet the conditions for having genuine *water*-thoughts, will be a situation in which I cannot so much as *think* that I am having a *water*-thought. This is just the familiar point that, if my first-order intentional states are all about *twater*, since I have only ever been in contact with XYZ, then none of my higher-order self-ascriptions can involve erroneously assigning *water* contents to my intentional states. By contrast, a situation involving counterfeit glasses of water is *not* one in which I cannot think (or believe, or be under the impression) that there is a glass of water.

At this point things get a bit complicated. Briefly, to envisage the right situation, it is not enough to imagine, e.g., that *twater* is brought to earth, since such a change will not automatically render false my belief that I am thinking there's water in my glass. I do not cease to have *water*-thoughts, or begin to have *twater*-thoughts upon the introduction of *twater* into my environment, or even upon first encounters with *twater*. Similarly, if I undergo a "fast switch" – I suddenly get transported to twin-earth – I will still be thinking that there is water in my glass. Since there is no water on twin-earth, my thought is bound to be false. But my self-ascription to the effect that that is what I am thinking will still be true. So now suppose that I undergo a "slow switch": I am unwittingly transported to twin-earth and stay there long enough to have acquired the concept *twater*. (And maybe I even start travelling between earth and twin-earth with some frequency.) This description of the scenario is not sufficient by itself either; we must add that I stay long enough on twin-earth to acquire the concept *twater*, but I do not thereby lose my concept *water*. For only if I have *both* concepts

simultaneously in my cognitive repertoire, would it be possible for the contents of my first-order thought and my self-ascriptive thought to pull apart in the requisite way.

This suggests that the content skeptic is less well-placed than the external world skeptic to mount a challenge to our ordinary knowledge. However, recall that the commonsense expectation is that content self-knowledge should be *more* resistant to skepticism than knowledge of the external world. In particular, we might expect that our ability to know what we are thinking should *not* be threatened *even* in a slow switching situations. And indeed, some externalists - Burge, for example - assures us that even in slow switching cases "there is no possibility of counterfeits," because basic self-knowledge is "self-referential in a way that ensures that the object of reference just is the thought being thought".⁷

But we may wonder: What is it about basic self-knowledge that makes it "self-referential" in Burge's sense? That is, what guarantees that the content of the ascribed thought and the ascribing thought will still interlock even in cases where a thinker has both twin-contents in her actual, active cognitive repertoire? I will offer an answer to this question later on. But before doing so, I want to bring out something that is easily obscured by focusing too narrowly on the compatibility of externalism and content self-knowledge. In fashioning a skeptical challenge using content externalism, the content skeptic is relying on a certain conception of content self-knowledge in particular, and of self-knowledge more generally. If we adhere to this conception, which I will call the *recognitional conception*, it will turn out that skeptic's challenge has nothing to do especially with content, nor is it exclusively threatening to externalists. But then perhaps it is the recognitional conception itself that should give way.

3. Content Skepticism and the Recognitional Conception of Self-Knowledge

The content skeptic's challenge can be put as follows:

How could a thinker know what content her present thought has, if it is possible, for all she knows, that the content her thought *really* has is different from the content it *appears* to her to have?

Put this way, however, it should be clear that the challenge potentially faces not only externalist views of content, but *any* objectivist view of content, whether externalist *or* internalist – any view that allowed for a

⁷ Burge (1988: 659).

possible gap between what objectively makes it the case that a thought *has* a certain content and what content the thought *appears* to anyone, including the individual herself, to have. To meet the skeptical challenge it is not enough simply to insist that contents depend for their identity only on facts internal to the individual. We have to be assured that the relevant internal facts will always somehow be immediately available to the subject who has contentful mental states. But mere acceptance of content internalism does not provide this assurance. As long as what makes it the case that someone has states with certain contents includes objective facts that are not directly available to her, then, regardless of whether they concern things "inside" or "outside" her head, we will have the makings of an appearance/reality gap with respect to content. Where there is such a gap, there is a possibility of imposters: mental states that appear to us to have contents that are different from the contents they actually have. It is this possibility that is relevant for generating content skepticism on the model of external-world skepticism.

The analogy with external world skepticism, however, requires an additional idea: that, in the normal case, telling what content a present state of ours has is a matter of *reliance on some 'appearances'*. The alleged difficulty for externalism arises because the view allows that there can be circumstances, however extraordinary, in which a thinker would be unable to discriminate among alternative thought contents available to her, given how things *appear* to her 'from the inside.' But this is problematic *only if* having ordinary knowledge of content *depends* on a thinker's ability to make such discrimination.

According to the commonsense doctrine of content self-knowledge (CSK), recall, self-ascriptions of present contentful states are special, in that they do not rely on the use of ordinary epistemic means such as ordinary observation, evidence, or inference. Still, on what we may describe as *the recognitional conception*, if such self-ascriptions are to constitute some kind of knowledge, we still must have at our disposal *some* epistemic means or method – a way of telling – that we deploy in obtaining the information they convey. Moreover, *to be able to tell* that a state of mine has a certain content *c*, *I must be able to tell apart* states with content *c* from states with other contents that are available to me, using some characteristic 'signs' that allow me to recognize content *c* as the content my state has (as opposed to any other content that is available to me). And, in keeping with CSK, my way of telling must be especially secure and unique to me. So, even if I

do not need to ascertain directly that I satisfy the conditions necessary for my entertaining content c , still, my ability to know that a present state of mine involves c , as opposed to some other content, depends on my being able somehow to recognize distinctive characteristics of my state or its content.

Thus, in the sorts of cases we looked at before – knowing that this is George, even without checking his DNA, or knowing that there is water in the glass, even without running chemical tests – I clearly have certain recognitional abilities with respect to the relevant items, which allow me to tell them apart from other items. Similarly, it is supposed that in the case of content self-knowledge, there must be some recognitional ability that allows us to tell that a present mental state of ours has one particular content, as opposed to some other candidate contents. But, according to externalism, there can be situations in which an individual will inevitably fail to meet this recognitional requirement, so she will fail to know what she is thinking. And this threatens the idea that ordinary knowledge of content can be *more secure* than ordinary knowledge of the external world.

By now, I think we should be in a position to appreciate that the dialectical situation will remain essentially the same if we broaden our focus and consider basic self-knowledge in general, and not just knowledge of content. Content skepticism is but a special case of a more general skepticism that has nothing specifically to do with content. The true analogue of external-world skepticism is then *internal world skepticism*, which poses the following generalized challenge:

How could a subject know what mental state she is in (as well as what content it has), if it is possible, for all she knows, that the state she is actually in is different from the state that it appears to her she is in?

(Let me note in passing again that this internal-world skepticism is indifferent to the specific view one holds regarding the nature of mental states, be it "externalist" or "internalist". Skepticism thrives wherever we combine an appearance/reality distinction with a conception of non-theoretical knowledge as epistemically mediated by some recognitional ability.)

As regards ordinary self-knowledge, what this means is that, if we want to adhere to the idea that we possess an especially secure and privileged kind of knowledge of our present states of mind, we must *either*

reject the appearance/reality distinction as it applies to one's own mind *or* reject the recognitional conception of ordinary self-knowledge. I myself do not see how we can have an objectivist view of mentality while denying that it is possible for things in the mental realm to be different from the way they appear on any given occasion to *anyone*, including oneself. But we have another option: to deny that ordinary self-knowledge is a matter of recognitional judgments that represent how things appear to one.⁸ This seems to me, in any event, a plausible route to take, since I don't think the recognitional model is an appropriate model on which to understand ordinary, basic self-knowledge.

4. The Security of Self-Knowledge: Immunity to Error

Ordinary present-tense mental self-ascriptions that are not made on the basis of therapy, careful reflection, self-interpretation, or scientific findings – what are often called "avowals" – seem to enjoy a special security. As I ascribe to myself in the normal way, say, a desire for a glass of water, I seem to *know* that I want water. I do not seem vulnerable to certain skeptical doubts about the content of my state. Could it be that I really want a glass of *twater*, rather than water? Somehow, the question seems to have no grip. My self-ascription seems secure also as regards its being water *rather than gin* that I want. Furthermore, I can tell securely not only that it is water that I want, but also that I'd *like* some water right now, rather than merely speculate that there's water in the glass, or that I have some other water-involving attitude. Indeed, I would argue that the security of content-assignment is but a special case of a more general security we enjoy when making mental self-ascriptions in the normal way.

Now, I have suggested that, if we accept a recognitional conception of ordinary self-knowledge, then skeptical alternatives abound, and the apparent security of avowals may be illusory. However, since this conception may seem to be our best model for understanding non-theoretical knowledge, then, if we answer

⁸ I see a prima facie difficulty in reconciling the following three claims:

- (i) There is an objective difference between, e.g., my being in M and my being in M'.
- (ii) We normally tell which state we are in by recognizing it introspectively.
- (iii) When ascribing a state to ourselves introspectively we cannot be fooled.

Holding (i) and (ii) has the effect of portraying what is in one's mind as, in a crucial sense, like what is in one's extra-mental world, in being epistemically external to one, whereas (iii) attempts to recover some kind of distinction between the mental and the extra-mental world. I am inclined to think, but cannot argue here, that the distinction cannot be preserved without rejecting the recognitional conception. In any event, it should be clear that accommodating

the skeptic by *rejecting* the recognitional conception, we may be saving self-knowledge from skepticism by embracing *deflationism* about self-knowledge:

If avowals are not open to doubt in view of alternatives, this is not at all because they represent an especially secure form of knowledge, but because they represent *no* knowledge whatsoever.

In what follows, I want to sketch a *non*-deflationist view of what allows ordinary mental self-ascriptions to be both especially secure and represent knowledge, even though they do not rest on recognitional determination.⁹

I propose that we understand the knowledge I have of what content a present mental state of mine has on the model of the knowledge I have that it is *I myself* who is instantiating a certain property. In "Self-Reference and Self-Awareness," Sydney Shoemaker makes an observation regarding certain self-ascriptions that use a first-person device such as 'I':

The statement "I feel pain" is not subject to error through misidentification: it cannot happen that I am mistaken in saying "I feel pain" because, although I do know of someone that he feels pain, I am mistaken in thinking that person to be myself. ...

If I say "I feel pain" or "I see a canary," I may be identifying for someone else the person of whom I am saying that he feels pain or sees a canary. But there is also a sense in which my reference does not involve an identification. My use of the word "I" as the subject of my statement is not due to my having identified as myself something of which I know, or believe, or wish to say, that the predicate of my statement applies to it.¹⁰

Shoemaker offers the notion of *immunity to error through misidentification* to characterize the special epistemic status of self-ascriptions of the kind he describes in these passages (a notion also discussed by

all three claims would require more than adopting an internalist conception according to which mental states supervene on an individual's internal organization.

⁹ For presentation of the challenge in terms of self-knowledge involving "no cognitive achievement" see Wright, Boghossian, and Fricker in Wright *et. al.* eds. 1998. The explanation I offer below is but a brief sketch of a view a develop at length elsewhere. (See Bar-On 2004, esp. Ch. IX.)

¹⁰ Shoemaker (1968: 557-8).

Gareth Evans and others).¹¹ The following features of judgments that are immune to error through misidentification (IETM, for short) are important for my present purposes. First, such judgments are not in general incorrigible. "I see a canary," "I am sitting down," "My legs are crossed," can all be plausibly said to be IETM, yet no one would be tempted to take them as incorrigible. Further, the class of self-ascriptions that are IETM includes *nonmental* self-ascriptions. Second, when an ascription is IETM, then, even if it can be mistaken, it is not open to a certain kind of error: an error of mistaking **me** for another because of a mistaken identification. Third, whether or not a self-ascription is IETM depends not on its semantic content, but rather on its "epistemic pedigree".¹² One and the same (semantically individuated) self-ascription can be IETM or not, depending on the *basis* on which it is made. If I tell that I am sitting down by seeing in a mirror, or on a video screen, someone who is sitting down, and whom I take to be me (perhaps mistakenly), my self-ascription will *not* be IETM, even though it is IETM when it is made in the normal way. Finally, if I issue a [true](#) self-ascription "I am F" that is IETM, I can be legitimately said to *know that it is I myself* who has the property F. However, my knowledge is not based on some recognitional identification of myself as the "right" subject of my ascription.¹³

This last point bears elaboration. Suppose I'm asked how I know it's a canary that I see at the bird-feeder. I may cite my ability to recognize canaries. This ability provides my epistemic grounds for the judgment that it is a canary I see. By contrast, when I say or think in the usual way: "I see a canary," it is not as though I think of someone in particular that she sees a canary, and take it to be me, or recognize that there is someone who sees a canary, and take it to be me who sees it. Indeed, I have *no* grounds or basis for thinking it is I who is the correct subject of my ascriptions, over and above, or separately from whatever

¹¹ Shoemaker in Cassam (1994: 82-3); Evans (1982: *passim*). For a recent discussion that distinguishes two different kinds of immunity to error through misidentification, see Pryor (1999). Since, as far as I can see, nothing hangs on this distinction for my purposes here, I will largely ignore it in what follows.

¹² Also, being IETM is not unique to first-person ascriptions. Both Shoemaker (*op. cit.*) and Evans (*op. cit.*) think that demonstrative thoughts such as "This is moving fast" are IETM (though clearly not infallible or incorrigible).

¹³ There is a "thin" sense in which I do identify myself as the subject of the ascription, for I do manage to *refer* to myself as opposed to someone else. But my ascription does not rest on a recognitional judgment of the form "I = r" (where "r" stands for some identifying description or demonstrative representation of an individual), or of the form "I am the individual who is F" (where I have an independent reason to think that someone is F). It may be useful to distinguish between the referential notion of identifying (a semantic notion) and the recognitional notion of identifying (an epistemic notion). Compare Evans (1982: 218).

grounds I have for thinking that *I* see a canary. And whatever knowledge I have that it is me who is F, for the relevant F's, does not rest on any independent my knowledge that *someone* is F; for in the normal case, my only grounds for thinking that someone is F is whatever grounds I have for thinking that *I* am F.

But what about counterfeits? Suppose I say or think: "I am sitting down." Couldn't it be, for all I know, that it is my identical twin who is sitting down? Well, that possibility would seem entirely irrelevant in this case, given that I do not tell that it is me who is sitting down by telling myself apart from other candidates. Since in the normal case, my self-ascription does not rest on any recognitional judgment identifying me as the individual who is sitting down, there is no room for me to doubt whether it is me, rather than someone else, who is sitting down.

Evans suggests (see quotation) that the *reason* why in certain cases it will make no sense to wonder: "Someone may be F, but is it I who am F?" is that, in these cases, there is no way to doubt that *I* am the one who is F without doubting that *someone* around is F. To put it somewhat awkwardly, in these cases, there is no stable judgment of appearances concerning who is F that I can hold onto so as to raise a doubt about how things really are. This is because the only reason it *appears* to me that someone is F is that it appears to me that I am F. I can, of course, be wrong that I am F. But Evans denies that this is sufficient to show that my knowledge that I am F must rest epistemically on my (correct) recognitional identification of myself as the individual who is F. What, then, allows a self-ascription of the form "I am F" to represent something I *know* about myself? Evans notes in this connection that we possess two general and distinctive capacities for gaining information about some of *our own* states and properties. First, we possess "a general capacity to perceive our own bodies" (which includes "our proprioceptive sense, our sense of balance, of heat and cold, and pressure"), and second we also have a capacity for determining our own "position, orientation, and relation to other objects in the world ... upon the basis of our perceptions of the world."¹⁴ It is the exercise of these capacities that gives rise to self-judgments that are IETM and which represent a certain kind of bodily self-knowledge..

¹⁴ Evans (1982: 220, 222).

If we accept Evans' analysis, then we have an interesting and powerful model that is useful for understanding the epistemic status of content self-ascriptions (and mental self-ascriptions more generally).

On this model, there are epistemically secure self-ascriptions that have the following features:

- (i) Though not absolutely infallible, they are *immune to a certain kind of error and doubt* (concerning the identity of the subject of ascription)
- (ii) They represent *knowledge* that something is the case (viz., that I am F)
- (iii) The knowledge is *non-recognitional*, in the sense that it does not depend on the subject's having direct recognition of that which is the case (i.e., the judgment that I am F does not rest on a recognitional identification of myself)

I think we can fruitfully enlist this model to understand what I call the distinctive *ascriptive security* of the mental self-ascriptions earlier referred to as avowals. Specifically, I maintain that

When making a self-ascription of a contentful state in the ordinary way, my self-ascription enjoys a special *immunity to error through misascription* and it can constitute knowledge I have about my state, even though the intentional content of my state is not for me a *recognitional target* (and neither is my state).

It is often tacitly assumed that I could only be said to know the content of my state if I had a separate epistemic reason for taking the content to be this rather than that. The idea of recognitional identification is invoked to supply the relevant reason: I know *c* rather than *c'* to be the content of my state because I can *recognize* *c* as the right content; I have a good way of at telling when it's *c* rather than *c'*. That's what epistemically grounds that part of my self-ascription. However, consideration of judgments that are immune to error through misidentification opens up the following possibility. When I simply avow being in state M with content *c*, I have no epistemic grounds for the ascriptive component of self-ascription that is not at the same time grounds for the self-ascription as a whole. But I see no reason why my avowal could not in principle represent knowledge that I am in state M with content *c*, even in the absence of any separate grounds for thinking that *it is c (rather than c') that is the content of my state (or that it is M, rather than M', that I am in)*.

When I make content ascriptions to others, my ascriptions do rest on such separate grounds. I see by your behavior that you are scared of *something*, and I look around to figure out what it is you are scared of. I conjecture that it is the dog you are scared of by noticing that you're looking at the dog. And I can sensibly wonder whether it is the cat, not the dog, that you are scared of, while not questioning that you are scared of *something*. This can be true even of content *self*-ascriptions. If I were to determine that *it is my neighbor* that I am annoyed at using indirect methods such as reviewing of the circumstances as well as my behavior, consulting with friends, or costly therapy, then I would be relying on specific grounds for taking *my neighbor* to be the intentional object of my annoyance that are not at the same time grounds for simply thinking I'm annoyed at my neighbor. But this is clearly not the *normal* case.

My proposal is that we should regard avowals of contentful states not only as immune to error regarding *who* it is that is in the relevant contentful state (that is, as IETM), but also as *immune to error through content misascription*. When assigning content *c* to a present state of mine in the ordinary way, I have no reason for taking my state to have content *c* other than whatever reason I have for taking myself to be in contentful state *M*. If my self-ascription of the contentful state *M* is false, which it *can* be, this will not be due to my mistaking the state's content for another, say, because *appears* to me a certain way. Whereas I can go wrong in my proprioceptive report "My legs are crossed" precisely because my legs may merely appear to me to be crossed when they're not (just as I can wrongly judge a colored patch in front of me to be blue, simply because it merely appear blue to me), when avowing a contentful state I cannot go wrong through mistaking one content for another due to its merely appearing to me to have that content.

5. The Security of Avowals: A Neo-Expressivist Account

We saw that, on Evans' analysis, although self-ascriptions that are IETM are immune to a certain kind of epistemic error, they can nonetheless represent knowledge that we have about ourselves obtained through the exercise of distinctive capacities for gaining information about some of our own properties and states. If we are to insist, as against the deflationary view, that avowals can represent knowledge we have about our mental states and their contents, we would do well to identify what is it that allows avowals to

represent knowledge that we are in a state with a particular content. My proposed answer is that, here too, we can appeal to a distinctive capacity that we deploy when issuing avowals. The proposal is as follows:

When ascribing to myself a state with content c in the normal way, I exercise an *expressive* capacity: the capacity to *use* content c (rather than content c^{\wedge}) to articulate, or give voice to my present state. When avowing, say, a present thought that is crossing my mind, I tell what content the thought has by telling *it*, i.e., that content. If I "identify" the content c , it is not by recognizing c to be the content I take my thought to have, but rather by using *that* content in the course of giving voice to my thought.

I shall first explain this expressivist proposal in connection with the case of content self-ascription with which we began, and then explain how we can actually see this case as but a special case of a more general phenomenon.

So let us return to

SAC: I am thinking that there is water in the glass.

Understood as a self-ascription of a presently entertained thought, rather than something one affirms, or judges to be true, SAC is maximally secure. For, when making it, I cannot misidentify the subject of my ascription, but I *also* cannot misascribe to myself the specific thought that there is water in the glass. Indeed, on a natural reading of self-ascriptions of this special kind, i.e., self-ascriptions of presently entertained thoughts, they are *self-verifying* – they make themselves true.¹⁵ Now contrast SAC with: "I am thinking something boring," which clearly does not seem self-verifying. What makes "I am thinking that there is water in the glass" self-verifying is the fact that the very act of self-ascribing the thought summons up the ascribed thought, as it were. Here what renders my self-ascription true is the fact that, in ascribing the entertained thought, I directly spell it out.

When the state ascribed is simply one of entertaining a certain thought, all I need to do in order to succeed in *truly* ascribing to myself the entertaining of a thought with content c is to entertain a thought with that content, which I am bound to do if I ascribe the thought explicitly. As I spell out the content, I cannot but entertain the very thought I ascribe to myself, which is why my avowal is self-verifying. This is not true

of all intentional avowals. However, although intentional avowals are not in general self-verifying, they do typically articulate their content. When avowing, say, the hope that dinner will be served soon, I do not obliquely describe my hope as when I report: "I am hoping for the same thing you are hoping." Rather, I explicitly articulate the content *that dinner will be served soon*, in the course of giving voice to my hope. And, I maintain, when I *avow* a contentful mental state, as oppose to ascribing it to someone else, or even ascribing it to myself indirectly, on the basis of some specific epistemic basis – when I self-ascribe a state in the 'avowing mode,' as we might put it – the explicit articulation of the state's content does not rest on an independent assessment of a feature of my first-order state.¹⁶ Rather, it is the upshot of putting to use the very same resources I would put to use in saying, or thinking "Let dinner be served soon" *hopefully*.¹⁷ (And as regard the state ascribed, if I am avowing, then I issue the self-ascription by way of expressing my present hope itself.)

Semantically speaking, a self-ascription that is IETM, such as "I am sitting down," *specifies* a particular individual who is said to be sitting down. However, what allows such a self-ascription to qualify as a piece of *knowledge* that I am sitting down is not the fact that I have correctly recognized myself as the one who is sitting down. Rather, it is the fact that in issuing the self-ascription I put to use certain capacities for gaining information, which I have only with respect to certain of my own states. I am making an analogous claim regarding self-ascriptions that are, as I put it, immune to content misascription. Semantically speaking, an avowal such as "I am hoping that dinner will be served soon" specifies a state I am said to be in as well as assigning it a specific content. However, what allows it to qualify as a piece of knowledge I have about my present state is the fact that in issuing the self-ascription, I put to use a distinctive capacity that I have only

¹⁵ For an analysis of self-verifying thoughts that is designed to be consistent with content externalism, see Burge "Our Entitlement..." in Ludlow and Martin eds. (1998).

¹⁶ When I assign content to another's state it by explicitly articulating it, I still rely on a recognitional judgment; the content is for me still a recognitional target. Whereas in my own case, as I explain below, the articulation of the content is made in the course of giving voice to the contentful state. For more discussion, see Bar-On (2004, Chapter VI).

¹⁷ Burge (1979) remarks that "[w]hen one knows that one is thinking that *p*, one is not taking one's thought (or thinking) *that p* merely as an object. One is thinking that *p* in the very event of thinking knowledgeably that one is thinking it. It is thought and thought about in the same mental act." (p. 70) Peacocke in Ludlow and Martin eds. (1998) also appeals to the idea of a self-referential or reflexive element in self-verifying self-ascriptions. The expressive take I offer on knowledge of content can help explain how this idea can be generalized.

with respect to my own mental states: the expressive capacity to give articulate voice to my present state of mind, or to *speak from* it.

6. Neo-Expressivism Expounded

The account I am invoking here – my *Neo-Expressivism* – is in the first instance an account of the distinctive security of avowals – the *epistemic asymmetries* between the way we treat certain present-tense self-ascriptions of mental states and the way we treat all other ascriptions. Unlike traditional epistemic accounts of this security, my account does not appeal to the security of a special epistemic method, or basis, or access on which I rely when avowing. Like more familiar expressivist accounts, it appeals instead to the notion that avowals, unlike ordinary reports, serve directly to express the avower's present mental states. However, unlike "Simple Expressivist" views, my Neo-Expressivism is motivated in a non-traditional way, without appealing to similarities between, e.g., saying "I have a headache" and wincing, or saying "I feel so happy!" and giving a big grin. Rather, I approached the expressivist idea by considering a sort of case that is a far cry from paradigm cases of naturally expressive behavior (yelping, wincing, gasping). I used the case of self-ascribing an articulate presently entertained thought (our SAC). The advantage of doing so is that it can help forestall a standard complaint against "Simple Expressivism", namely, that it compromises the *semantic continuity* between avowals with ordinary reports. My Neo-Expressivist account, unlike its predecessors, is designed from the start to accommodate the idea that, when making an avowal I issue a genuine self-ascription that can be true or false, and can be interchanged in context with reports about the mental state I am in, in virtue of its semantic content. Indeed, I believe the expressivist insight is best approached by focusing on the epistemology of avowals and their use, rather than their semantics.

Although I have appealed to the expressivist insight in connection with the security of self-verifying self-ascriptions, I think it is applicable to non-self-verifying cases of avowals, and more generally, to all avowals, whether intentional or phenomenal. The general expressivist claim I endorse is that avowals derive their special security not from being based on some special epistemic method or access, but rather from being expressive acts in which subjects directly give voice to—share, air, vent—a self-ascribed mental state. A subject who issues an avowal, like one who says: "This is great!" or "How gross!" or "God, please no!" *speaks*

from her state, instead of giving a nonverbal, natural expression to her state. Although the avowal proper tells us that the self-ascriber is in some state, just like a report of that state (by someone else or by the avower herself), in the circumstances, it may still play an expressive role similar to that played by a natural expression of the state.

To bring out the various aspects of similarity and dissimilarity between avowals and natural expressions, we need to make some distinctions. Specifically:

(I) Sellars' three senses of "expression":

a-expression: in the *action* sense a *person* expresses a state of hers by intentionally doing something; For example, when I intentionally give you a hug, or say: "It's so great to see you," I express in the action sense my joy at seeing you. One may also express one's feeling of sadness in the action sense by *letting* tears roll down her cheeks, instead of wiping them out and collecting oneself. (Note that a-expression requires doing *something* intentionally, not necessarily *expressing* intentionally.)

c-expression: in the *causal* sense an *utterance* or piece of behavior expresses an underlying state by being the culmination of a causal process beginning with that state; For example, one's unintentional grimace or trembling hands may express in the causal sense one's pain or nervousness, respectively.¹⁸

s-expression: in the *semantic* sense e.g., a *sentence* expresses an abstract proposition, thought or judgment by being a (conventional) representation of it.

For example, the sentence "It's raining outside" expresses in the semantic sense the proposition that it is raining at time [t] outside place [p].

The similarity I find between avowals and natural expressions concerns a- and c- expression, *not* s-expression. Natural expressions do not express anything in the semantic sense. And avowals s-express a proposition/judgment that I am M. But it's still possible that in the causal and action senses an avowal, just like a natural expression, expresses the state it ascribes. Indeed, to give full purchase to the expressivist idea, we should focus more narrowly on the comparison between avowals and natural expressions that fall under a-

expression — that is, cases where one produces a facial expression or a gesture intentionally, and where we can think of the expressed mental state not only as the 'brute cause' but also as the rational cause of, or reason for, one's expressive behavior. And more narrowly still, the crucial comparison that really matters is between avowals and *non-self-ascriptive linguistic utterances* through which speakers are said to express their present feelings, emotions, and other occurrent mental states. (We'll come back to this shortly.)

(II) Avowals as acts and avowals as products:

As we consider cases of intentionally producing expressive behavior, we should distinguish between an *act* of expressing and its *product*. The important similarities between avowals and other intentionally produced expressions require thinking of avowals as acts, whereas the salient dissimilarities come to the fore when we think of avowals as products.

The product of an act of avowing, unlike a smile or a wince, or even a verbal cry such as "Ouch!", is a semantically articulate self-ascription, an item with semantic structure and truth-conditions. It is a product whose properties allow it to serve, and be caught up, in other kinds of distinctively linguistic (and mental) acts. Importantly, avowals, understood as products, s-express (in Sellars' semantic sense) self-ascriptive propositions, to the effect that the avower is in some state. And the mentalistic terms that constitute parts of avowals semantically represent the relevant states. Natural expressions understood as products—a facial expression, a gesture, bodily movements or demeanor, an inarticulate sound, however produced—do not s-express anything. There are no semantic conventions in virtue of which laughter represents amusement, no linguistic rules in virtue of which a hug signifies, or refers to, joy at seeing the person hugged. Moreover, it could perhaps be argued that natural expressions *show* the expressed mental states, whereas avowals only *tell* of the states. This too should be understood as a claim about the respective products. But we should keep in mind that we often take linguistic acts – saying "What a mess!" or "How gross!" or "This is great!" – to be a-expressive of speakers' mental states, even though they have semantically articulate products, and even though these products do not mention the mental states they a-express. So we can agree that there are

¹⁸ As I understand it, expressing 'in the causal sense' is not a purely causal notion. For discussion, see Bar-On (2004, Chapter VII and VIII).

notable differences between the products of avowals and of acts of natural expression while retaining the expressivist insight.

Let's consider another example. A linguistic utterance such as "It's so good to see you!" which typically serves to a-express the speaker's joy, does so with an expressive vehicle that s-expresses the proposition or thought that it is very good to see one's hearer, in virtue of the rules of English. If, as I maintain, one can also a-express her joy by avowing "I am so glad to see you!" this would involve using a sentence which, in virtue of the rules of English, s-expresses the *self-ascriptive* proposition or thought that I am very glad to see her hearer. The minimal expressivist claim is that these are two acts that equally serve to a-express the agent's joy, though they use different expressive vehicles (both linguistic, but only the latter self-ascriptive). I think that once we see how to understand the similarity claim, we should be more open to seeing that these two linguistic acts can be in turn similar, in point of the act performed, to an act of giving your friend a cheerful hug. The product of latter act is not governed by any syntactic or semantic rules of English, and it does not s-express anything. Still, in all three cases we can see the agent as performing the same kind of act—i.e., giving expression to (a-expressing) her joy at seeing someone.

(III) Expressing M vs. Expressing One's M:

The third distinction I want to call on is between expressing *a* mental state and expressing *my* mental state. An actor on stage may express anger making use of a variety of expressive vehicles, from growling to throwing her hands in the air, to uttering: "Darn it!" to saying: "I am so angry". But she will typically not be expressing *her* anger.¹⁹ The distinction is equally applicable to both natural and linguistic expressions. Intentionally produced natural expressions exhibit a certain measure of 'expressive autonomy' (a feature regularly exploited in the arts). A disgusted face, a scared shriek, a yawn of boredom, can all be produced on occasion in the absence of the relevant feelings or emotions.

¹⁹ Perhaps the *character portrayed* by the actor is expressing her anger (if we think *Romeo* can die in the play, even though the actor portraying him doesn't, there's no obvious reason why *Romeo* can't express 'his' anguish, even if the actor portraying him doesn't). We may, however, think that the actor only *pretends* to be a-expressing his anger. (And, while no one can c-express a state she's not in, one can presumably *pretend* to c-express such a state – an actor pretending to stub her toe, when portraying a character who is stubbing her toe, can pretend to c-express feeling a sharp pain). For discussion, see Bar-On (2004, Ch. VIII).

Natural expressions can not only be simulated on purpose, but they can also be 'pressed' from someone even when the person is not in the relevant mental state, due to unusual circumstances. A person on a dentist's chair with a long history of painful dental work may emit a gasp of pain as the drill approaches her mouth, *as well as* say: "Ow!" or "My tooth hurts!" even though in reality she feels no pain. Circumstances conspired to squeeze out of her an inappropriate expression. We need not suppose her to be insincere, nor do we need to suppose that her verbal or non-verbal emissions were due to a recognitional mistake she's made regarding the state she is in. A *psychological* explanation of what led to the inappropriate expression is ready at hand. It would allow us to make sense of the possibility of false avowals without resorting even to a non-Cartesian *recognitional* story and yet without giving up the expressivist explanation of avowals' security. (Standard cases of so-called wishful thinking and self-deception can be treated along the same lines. See Bar-On 2004, Chapter VIII.)

A strand of the Neo-Expressivist account that has so far remained in the background can now come to the fore. We saw that the account invokes the expressivist idea to explain the distinctive *ascriptive* immunity to error of avowals. But the expressivist idea, properly understood, can also help explain why avowals are governed by a strong presumption of truth (even though, as indicated above, this presumption *can* be overridden; for we can make sense of cases of false avowals). If we take someone's behavior to be an act of genuine – as opposed to pretend – (a- or c-) expression of a mental state, we take it that she is *in* the mental state that we perceive in her behavior. Except under non-standard circumstances, if we think that someone is engaged in behavior expressive of, say, fear, we take it that not only does the behavior show us fear, but also that it shows us the *agent's* fear. The verb "express" shares in the facticity, or 'success' aspect, of the verb "show". In the special case of avowals, though, the product of the act actually *tells* us that the agent is in a certain mental state. Thus, to take the avowal to be expressive of the agent's mental state *just is* to take the avowal to be *true*. Avowals are unique among expressive acts in that their products semantically express a proposition that ascribes to the agent the very same mental state that we take her to be expressing in the action sense in the act of avowing. If someone says: "This is so boring!" we can take her to be a-expressing her boredom *without* presuming her utterance to be true. (We can perfectly consistently suppose her to feel

bored by something that isn't boring at all.) But to take the utterance: "I feel so bored" to express the utterer's boredom *is* to take the utterance to be true. (As hinted above, though, there is a certain slack between a-expressing c-expressing, a slack characteristic of all intentional acts; and herein lies the possibility of false avowals.)

To sum up: On the Neo-Expressivist account, when avowing, a person issues a self-ascription of a mental state in the course of speaking directly from her present condition, rather than reporting her finding regarding her present state of mind. Unlike the person who grunts, or smiles, she is *speaking* her mind. And unlike the person who says "Darn it!" or "How awful!" or "This is great!" she speaks self-ascriptively. On the present account, to regard someone's self-ascription as an avowal is to take her to be a-expressing her self-ascribed state. This is why it would make little sense to contradict or correct her self-ascription. For, to take it that she has (successfully) a-expressed her state just is to take it that she has given voice to a state she is in. Recognizing an expressive element in all avowals thus contributes to a non-recognitional explanation of avowals' distinctive security. Insofar as we think of all avowals as enjoying a special security that goes beyond the epistemic security of well-grounded or highly reliable self-reports, it is *because, or to the extent that*, we regard them as acts in which the speaker successfully a-expresses the very conditions ascribed by the proposition the avowal as product s-expresses.

7. Self-Knowledge

Back to self-knowledge. On the Neo-Expressivist account, the *security* of the self-ascription does not come from the epistemic security of the avower's recognition of the presence and character of her state, but it is rather a consequence of the fact that the avowal is a successful act of a-expressing the self-ascribed state. But if this is so, one may worry that the deflationist challenge has still not been met. For, even if avowals are not vulnerable to "internal-world" skepticism, they are altogether poor candidates for genuine knowledge that the avower has about herself. This worry can be perspicuously couched in terms of the familiar notion that to know *that p* one must have a *warranted belief that p*. Now, if an avowal (a- or c-) expresses the mental state it ascribes, doesn't that mean it does *not* (a- or c-) express the avower's *belief* or judgment that she is in the mental state? But then how could it possibly represent something she knows? This is a good question to

ask, and I have addressed it at length elsewhere.²⁰ Here I can only offer a few brief remarks. Notice, first, that the question presupposes that, for example, if an avowal (a- or c-) expresses my hope that *p*, it cannot at the same time *also* (a- or c-) express my belief, or judgment that *p*. This exclusivity claim can be denied. My expressivist account denies that my avowal represents a recognitional self-judgment that is based on the way my present state or its content appear to me. And it maintains that my avowal a-expresses my first-order state. But it can allow that my avowal *also* a-expresses something I *hold true* about myself, and not only in the weak dispositional sense that *if* I were to consider what the avowal says I would affirm it. For I can also be said to make the relevant occurrent judgment, at least in the sense that I intentionally produce the relevant self-ascription. Granted, since my self-ascription is not based on any recognitional judgment, and it is certainly not based on any observation, evidence, or inference, it doesn't seem that I could have *justification* for holding-true the self-ascription, at least not in any traditional sense. However, this does not mean that I cannot be warranted in my belief, in the sense of being *entitled* to hold-true the self-ascription. Perhaps, as the subject of the relevant mental state, who is speaking from the mental state—who is giving it articulate voice—I am simply warranted by default, as it were, since no defeating alternatives to the ascription I make are in the running. Or perhaps my self-ascription, though not based on any recognitional judgment, is grounded (à la Peacocke)²¹ in the very same state that gives reason for the act of avowing, namely, the mental state that I self-ascribe. If any of these options works, then we could think of avowing subjects not only as expressive beings capable of speaking their mind, but also as *knowing* selves, capable of giving articulate voice to warranted beliefs about their present states of mind.

If the neo-expressivist account is right, we can agree with Descartes that our ordinary mental self-ascriptions are not open to the same skeptical threat as our ordinary claims about extra-mental reality. Should this tempt us to Cartesian dualism? Wittgenstein describes the temptation to postulate Cartesian Egos as follows:

²⁰ In Bar-On (2004, Chapter IX).

²¹ See Peacocke (in Ludlow and Martin eds. 1998).

We feel then that in the cases in which "I" is used as subject, we don't use it because we recognize a particular person by his bodily characteristics; and this creates the illusion that we use this word to refer to something bodiless, which, however, has its seat in our body. In fact *this* seems to be the real ego, the one of which it was said, "Cogito, ergo sum".²²

Wittgenstein's "uses of 'I' as subject" are precisely what Shoemaker and Evans try to capture with the notion of immunity to error through misidentification: uses of "I" in judgments that rest on no recognition of "a particular person by his bodily characteristics." But, Wittgenstein warns us against concluding that they rest on recognition of someone by *non-bodily* characteristics. We should equally resist an analogous temptation in the case of the ascriptive part of mental self-ascriptions. We often issue such self-ascriptions without any reliance on recognition of physical (or even functional) features of our mental states. Though our avowals enjoy immunity to error through misascription, we should not conclude that avowable mental states have *nonphysical* qualitative features that only we can securely recognize in our own case and on which privileged self-knowledge is based. And replacing the recognitional conception of self-knowledge with the neo-expressivist conception allows us to see that.²³

²² Wittgenstein (1960: 69)

²³ My thanks to Ram Neta, Dylan Sabo, and Ted Parent for comments on earlier drafts of this paper. Thanks to audiences at MIT, Davidson College, University of Maryland, Auburn University, and to the participants at the Self-Knowledge and the Self Workshop in Bigorio, Switzerland (2003) for helpful questions and comments.

References

- Bar-On, Dorit. *Speaking my mind: Expression and self-knowledge*. (Oxford: Oxford University Press), 2004.
- . 'Externalism and Self-Knowledge: Content, Use, and Expression,' *Nous*, **38**, 430-455, 2004a.
- Boghossian, Paul A. "What the Externalist Can Know A Priori." Pp. 271-84 in Wright *et. al.* 1998. (originally published in *Proceedings of the Aristotelian Society* 97 (1997): 161-75).
- . "Content and Self-Knowledge." Pp. 149-74 in Ludlow and Martin. eds. 1998. (originally published in *Philosophical Topics* 17 (1989): 5-26).
- Burge, Tyler. "Individualism and Self-Knowledge." Pp. 65-79 in Cassam ed. 1994. (originally published in the *Journal of Philosophy* 85, no. 11 (1988): 649-63).
- . "Individualism and the Mental." Pp. 21-83 in Ludlow and Martin eds. 1998. (originally published in *Midwest Studies in Philosophy* 4 (1979): 73-122).
- . "Our Entitlement to Self-Knowledge." Pp. 239-64 in Ludlow and Martin eds. 1998. (originally published in *Proceedings of the Aristotelian Society* 96 (1996): 91-116).
- Cassam, Quassam ed. *Self-Knowledge*, Oxford: Oxford University Press,
- Davies, Martin. "Externalism, Architecturalism, and Epistemic Warrant." Pp. 321-61 in Wright *et. al.* eds. 1998.
- Evans, Gareth. *The Varieties of Reference*. Oxford.: Clarendon Press, 1982.
- Fricker, Elizabeth. "Self-Knowledge: Special Access Versus Artefact of Grammar – A Dichotomy Rejected." Pp. 155-206 in Wright *et. al.* eds. 1998.
- Goldman, Alvin. "Discrimination and Perceptual Knowledge." *The Journal of Philosophy* 73, no. 20 (November 1976): 771-91.
- Peacocke, Christopher. "Our Entitlement to Self-Knowledge: Entitlement, Self-Knowledge and Conceptual Redeployment." Pp. 265-303 in Ludlow and Martin eds. 1998. (originally published in *Proceedings of the Aristotelian Society* 96 (1996): 117-58).
- Pryor, James. "Immunity to Error Through Misidentification." *Philosophical Topics* 26, no. 1 & 2 (Spring & Fall 1999): 271-304.
- Putnam, Hilary. "The Meaning of 'Meaning'." Pp. 215-71 in Putnam, *Mind, Language and Reality*. Cambridge; Cambridge University Press, 1975.
- Shoemaker, Sydney. "Self-Reference and Self-Awareness." Pp. 80-93 in Cassam ed. 1994. (originally published in the *Journal of Philosophy* 65, no. 19 (1968): 555-67).
- . "Self-Knowledge and 'Inner Sense'." *Philosophy and Phenomenological Research* 54, no. 2 (June 1994): 249-314.
- Wittgenstein, Ludwig. *The Blue and Brown Books: Preliminary Studies for the "Philosophical Investigations"*. 2nd Edition. New York, NY.: Harper and Row, 1960.
- Wright, Crispin. "Self-Knowledge: The Wittgensteinian Legacy." Pp. 15-46 in Wright *et. al.* 1998.
- Wright, Crispin, Barry C. Smith, and Cynthia Macdonald eds. *Knowing Our Own Minds*. Oxford.: Clarendon Press, 1998.
- . "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics." Pp. 233-64 in *Reflections on Chomsky*, edited by Alexander George. Oxford.: Basil Blackwell Ltd, 1989.