

## NEO-EXPRESSIVISM

In Speaking My Mind, Dorit Bar-On develops a Neo-Expressivist approach to the epistemic status of avowals and to self-knowledge.<sup>1</sup> This paper is a critical study of Bar-On's very rich and stimulating book.

The Distinctive Security of Avowals

Sentences such as 'I am in pain', 'I am thinking about Cassis', 'I believe that Bobby Jones was the greatest golfer', and 'I want sashimi' can be used to make avowals. In making an avowal, one self-ascribes a current mental state. But not every use of a first-person, present tense sentence of the form 'I am in M' (where 'M' refers to a mental state) counts as an avowal.<sup>2</sup> First, the use of 'I am in M' must be sincere in order to count as an avowal. Second, the use of 'I am in M' must not be based on ordinary evidence. Suppose that I say, 'I believe that my sister hates me' after undergoing lengthy psychoanalysis which reveals that heretofore unacknowledged attitude towards my sister. This does not count as an avowal, due to its grounding in evidence of testimony and behavioral observation.

'I am taller than Kurt', 'My arm is bent', and 'My leg is broken' cannot be used to make avowals. Neither can 'I was in pain', 'I wanted sashimi', 'I will want more tomorrow', 'She thinks that Paris is beautiful'. 'I see a tiger' cannot be used to make an avowal, because even though I can use it to self-ascribe a mental state, the ascription does more: it is external-world-involving.

What exactly is an avowal, then? An avowal is not a sentence, since one and the same sentence can be used to make an avowal on one occasion, while it is not so used on another occasion. 'I believe that my sister hates me' can be used to make an avowal when the use is not evidentially based, in contrast to the use in the psychoanalytic context. At first blush, it would seem that an avowal is a speech act, performed by a speaker in making a sincere, non-evidentially-based utterance of a sentence of the form 'I am in M' (where 'M' refers to a mental state). According to Bar-On, though, things are a bit more complicated. She wants to consider a silent, inner tokening of 'I am in M' to be a candidate for being an avowal. Though the pertinent

mental act is not a speech act, it still counts as an avowal.<sup>3</sup>

We can make some distinctions within the class of avowals. First, there are phenomenal avowals, such as 'I am feeling thirsty'. These avowals ascribe a mental state without specifying an intentional content for the state (though the state might in fact possess such a content). Second, there are intentional avowals. The first kind of intentional avowal specifies an intentional object in ascribing a mental state: 'I am thinking about baseball'. The second kind of intentional avowals specifies a propositional content in ascribing a mental state: 'I hope that the Yankees will win the World Series'. Bar-On will seek to give a uniform account of all three kinds of avowal.

Bar-On characterizes what she calls the distinctive security of avowals as follows: (1) avowals are protected from ordinary epistemic assessments, such as requests for reasons, challenges to their truth, and simple, straightforward correction, (2) avowals are strongly presumed to be true, unlike all other empirical ascriptions, including third-person mentalistic ascriptions ('She is in pain') and non-mentalistic first-person ascriptions ('I am standing'). According to (2), then, there is an Epistemic Asymmetry between avowals and all other contingent, empirical ascriptions.

Bar-On does not hold that avowals are infallible. For her, some avowals are mistaken, or, as she says, false. For example, someone who is expecting a pain in his back says, 'I am feeling pain now' when a cube of ice is applied. Someone who is unaware of his pattern of racist behavior says, 'I believe that all men are created equal'. These, we can imagine, are false avowals, in that the speaker is not in pain, does not have the belief in question. The avower can reasonably be challenged, can be presented with evidence that warrants a retraction. So avowals are not incorrigible either.

One of Bar-On's chief goals is to explain the distinctive security of avowals: why do they have such a special epistemic status?

### Semantic Continuity

Bar-On says, "Semantically speaking, an avowal will typically identify an individual--the speaker--and ascribe to her a certain condition--being thirsty, feeling scared, hoping that...,etc." (p. 8-9) Thus, on Bar-On's view, an

avowal would seem to be an assertion of a proposition that is true iff the speaker is in the condition that she ascribes to herself.<sup>4</sup> Avowals are thus truth-evaluable, in the sense that the proposition asserted by an avowal has a truth-value. Hence avowals can serve as premises in inferences, such as my utterance of 'I am scared of the dog, and so are you; so that makes two of us'. Further, for any avowal, there is a corresponding non-avowing assertion of a proposition with the same truth-conditions as the proposition asserted by the avowal. For example, when I now avow 'I am in pain', the proposition I assert is true under the same conditions as the proposition you assert by now uttering 'AB is in pain', as well as the proposition you assert by later uttering 'AB was in pain'. These semantic features of avowals are given the moniker Semantic Continuity by Bar-On.<sup>5</sup>

### Simple Expressivism

One way of trying to explain the distinctive security of avowals is to adopt a view that Bar-On calls Simple Expressivism (hereafter SE). Wittgenstein famously said, "[T]he verbal expression of pain replaces crying and does not describe it".<sup>6</sup> On SE, an avowal such as an utterance of 'I am in pain' is not a report about, or a description of, one's mental state. Avowals are not assertions of propositions. Saying 'I am in pain' is like crying out in pain, or writhing, or saying 'Ouch!'. Avowals, then, are bits of expressive behavior which are not truth-evaluable and are not associated with any asserted truth-evaluable propositions. A related view is Elizabeth Anscombe's denial that the first-person pronoun has a referent in 'I am in pain'.<sup>7</sup> If that view is correct, then no proposition about AB is asserted in the course of my utterance of 'I am in pain'.

SE affords a sort of negative, or deflationist, explanation of the distinctive security of avowals. A groan is not the sort of thing that can be based on evidence, "epistemically grounded in a subject's judgement or belief about some state of affairs". (230) There is no need to justify or give reasons for a grimace. Writhing is not subject to challenge. These points all hold as well for avowals, on SE. Further, a grunt of pain cannot be mistaken in the way that an assertion of a false proposition is mistaken. The same goes for an avowal. These considerations also account for the thesis of Epistemic Asymmetry between

avowals and ordinary, non-expressive empirical ascriptions, such as 'She is in pain' and 'That's a goldfinch'.

Bar-On rejects SE for a number of reasons. First, the view contravenes Semantic Continuity. On SE, it is not the case that the proposition asserted by my utterance of 'I am in pain' is true under the same conditions as is the proposition asserted by your utterance of 'AB is in pain', since the avowal does not involve the assertion of a truth-evaluable proposition at all. Second, on SE, avowals do not "represent" or express self-knowledge, since there is no proposition in the offing to be known.

The third problem for SE arises in connection with a comparison between SE and ethical non-cognitivism. The non-cognitivist believes that utterances of sentences like 'Murder is wrong' are not assertions of truth-evaluable propositions. Instead, such an utterance serves to express the speaker's negative attitude towards murder, as does an utterance of 'Boo on murder!'. SE is a sort of non-cognitivism about avowals. P. T. Geach pointed out a prima facie problem for ethical non-cognitivism.<sup>8</sup> First, how can the non-cognitivist account for the occurrence of ethical sentences embedded within conditionals, such as 'If murder is never wrong, then John Kennedy's assassination was not wrong'? In an utterance of this sentence, the speaker is not expressing a pro-attitude towards murder in uttering the antecedent. Second, and relatedly, how can the non-cognitivist account for the role of ethical sentences in reasoning, such as:

- (A) If murder is never wrong, then John Kennedy's assassination was not wrong.
- (B) John Kennedy's assassination was wrong.
- So, (C) It is not the case that murder is never wrong.

If no truth-evaluable propositions are asserted in utterances of (A)-(C), then why do we think that a valid piece of reasoning has occurred?

Geach's problem also afflicts SE. First, how can the view make sense of conditionals like 'If I am in pain, then I will let you know'? An utterance of this sentence does not involve any expression of pain via the utterance of the antecedent. Second, it is hard to see how SE can account for the apparent validity of the reasoning:

- (D) If I am in pain, then someone is in pain.
- (E) I am in pain.
- So, (F) Someone is in pain.

### The Epistemic Approach

So Bar-On rejects SE, and she considers an alternative sort of approach to avowals: the Epistemic Approach (hereafter EA), according to which avowals "involve self-judgements made on an especially secure epistemic basis". (11) One version of EA is Cartesianism, according to which one has an infallible faculty of inner sense which provides access to one's non-physical mental states.<sup>9</sup> Another version of EA is Materialist Introspectionism, according to which "the human brain is equipped with a special mechanism—a scanner—designed to deliver reliable higher-order judgements about our first-order mental states", such as pain. (96) We can formulate a more general version of Introspectionism as follows: one is equipped with a mechanism of some sort or other which reliably produces second-order beliefs about one's first-order mental states, such as thoughts about baseball, desires for pizza, beliefs about numbers, and so on. The introspective mechanism is thought to involve a causal link between such first-order states and second-order beliefs about them.

One advantage the EA has over SE is that Semantic Continuity is not contravened. On both versions of EA, avowals are assertions of truth-evaluable propositions concerning the avower's mental states (either specially accessed immaterial states, or the states which reliably cause second-order beliefs about them). Avowals are reports about the avower's mental states, "made on an especially secure epistemic basis". (11)

Bar-On rejects both versions of EA. We will not go into her reasons in any detail, beyond the following sketch. She does not wish to explain the distinctive security of avowals in the Cartesian way because she does not wish to posit entities whose special immaterial nature allegedly gives rise to a special accessibility to a special inner sense. She thinks, further, that post-Cartesian Introspectionism is saddled with a number of difficulties. If we do possess reliable introspective mechanisms, then it is a contingent fact that they are highly reliable. But this distinctive status of avowals does not seem to be a contingent matter.

Further, the contingency of the mechanism's reliability would give rise to the following two possibilities. First, in ordinary perception, brute error is possible.<sup>10</sup> This is error in which the perceptual system is functioning properly but the world fails to cooperate

(e.g., an infant fooled by a bent-looking stick in the water). Given the similarity between introspection and perception, we should expect the introspective mechanism to be sometimes caught in brute error about first-order mental states, leading to false avowals.<sup>11</sup> But according to Bar-On, false avowals do not involve brute error. False avowals do not involve "mistaking one state for another". (201) In a false avowal, the fault lies with the subject, not with an uncooperative, tricky inner world.

Second, the contingency of the reliability of our introspective mechanisms gives rise to the possibility of global error in a subject's avowals, on analogy with systematic perceptual error in a defective perceptual system. But this possibility seems deeply counterintuitive to Bar-On and others.<sup>12</sup> One might well wonder why we do not in fact find the possibility of global error realized, if it really is a genuine possibility.

Still further, if I could somehow scan your first-order mental states and form highly reliable second-order beliefs about them, then my pronouncements about your mental states would, on the Introspectionist view, have the same distinctive security as your avowals. But this consequence of introspectionism seems absurd: the distinctive security of avowals is "non-transferable". (20)

### Self-verification

When I say, 'I have a broken leg', there is the possibility of error through misidentification. For example, suppose that I have just come to consciousness in a field hospital and see a leg in a cast. I mistakenly believe that I am the man whose leg is in the cast, due to my disorientation (the man on the cot next to me is the man whose leg is in the cast). Thus, I err in saying that I have a broken leg, due to a mistaken identificatory belief. By contrast, when I say, 'I am in pain', it makes no sense for me to step back and wonder whether it is really I, and not someone else, who is in pain. This is because avowals in general are not grounded in identificatory beliefs such as that I am the man whose leg is in the cast. Thus avowals are immune to error through misidentification.<sup>13</sup> In the hospital case, it does make sense for me to wonder whether it is really I, and not someone else, who is the man whose leg is in the cast.

Consider an intentional avowal which concerns a mental state whose propositional content is specified by the avowal: 'I believe that it is raining'. In making this

avowal, I do not identify myself as the subject of the avowal. I do not ground the avowal in a recognitional judgement about myself, and there is no corresponding identificatory belief. In making such an avowal, I self-ascribe a state with a certain content. Bar-On thinks that such avowals have an ascriptive immunity to error that is like their immunity to error through misidentification. When I say, 'I believe that it is raining', it would be "out of place" for me to wonder whether it is really that it is raining, rather than some other content, that I believe.

Bar-On wants to illuminate this content-ascriptive immunity to error by considering the self-verifying character of avowals of entertained thoughts, such as an utterance of 'I am thinking that water is wet'.<sup>14</sup> Let us understand thinking that P as the occurrent entertaining of the proposition P. So one can think that P without believing that P or judging that P. According to Bar-On's exposition of the Burgean notion of self-verification, when one utters 'I am thinking that water is wet', "one's success in making a true self-ascription is guaranteed simply by the act of avowing". (209) This is because "the very act of making the thought ascription summons up the ascribed thought". (209) There is thus no need to identify or recognize the self-ascribed content. According to Bar-On, this point generalizes: "When avowing an occurrent belief, hope, preference, etc., I typically single out the content of the ascribed state directly, by articulating it in the very act of avowing". (212)

However, in avowing 'I hope that the Yankees win', it does not simply follow that I have the avowed hope, in the way that in avowing 'I am thinking that water is wet', it does simply follow that I have the thought. But exactly why does it follow? According to Bar-On, the act of avowing guarantees that I have the thought in question, thereby guaranteeing the truth of the proposition that I am thinking that water is wet. But this is not quite the way that Burge sees the matter. For him, when I sincerely avow 'I am thinking that water is wet', I make a second-order judgement about my state of mind: I judge that I am thinking that water is wet. This is a judgement about a first-order mental state: my thought that water is wet. The second-order judgement is guaranteed to be correct when made, according to Burge, because in judging that I am thinking that water is wet, I must therein entertain the proposition that water is wet. But entertaining the proposition that water is wet just is thinking that water

is wet. Thus the mental act of judging that I am thinking that water is wet ensures the truth of the proposition that I judge to be true (viz. the proposition that I am thinking that water is wet). Thus the mental act of judging that I am thinking that water is wet ensures its own correctness (where the correctness of a judgement consists in the truth of the judged proposition). So it is not the speech act of avowing that guarantees that I have the thought in question. It is rather the second-order mental act of judging that I am thinking that water is wet that does the guaranteeing in question. We will return later to Bar-On's ambivalent attitude towards second-order mental acts and states.

Let us return to the my avowal of 'I hope that the Yankees win'. If I judge that I hope that the Yankees win, then, according to Burge, it follows that I entertain the proposition that the Yankees win. So it follows that I am thinking that the Yankees win, since entertaining the relevant proposition just is thinking the thought in question. But from that fact that I make the judgement in question, it does not follow that I hope that the Yankees win. So the my avowal of 'I hope that the Yankees will win' is not self-verifying.

### Expression

On Bar-On's understanding of self-verifying avowals, they share a common feature with other intentional avowals: in all intentional avowals, "when avowing an occurrent belief, hope, preference, etc., I typically single out the content of the ascribed state directly, by articulating it in the very act of avowing"; I single out the content "simply by spelling it out". (212) One way of making my critical point from the preceding section is this: the self-verifying character of avowals of occurrent thoughts has nothing to do with the articulative nature of the pertinent speech acts.

Still, this critical point is compatible with the idea that articulating, or spelling out, an intentional state's content is relevant to the distinctive status of avowals generally. But what exactly is the articulating, or spelling out, of content that is achieved by an intentional avowal? Here we arrive at a key notion in Bar-On's explanation of the distinctive security of avowals: "articulating the content serves directly to give voice to my present state; it constitutes expressing the very thought I am ascribing to myself". (215) Bar-On aggress

with the Simple Expressivist that in general, avowals serve to express the avower's mental states, rather than to report them.

Following Wilfrid Sellars, Bar-On distinguishes three kinds of expression.<sup>15</sup>

EXP1 is the action sense: a person expresses a state of hers by intentionally doing something.

EXP2 is the causal sense: an utterance or piece of behavior expresses an underlying state by being caused by that state.

EXP3 is the semantic sense: a sentence expresses an abstract proposition by virtue of the sentence's meaning and the intentions of the speaker who utters the sentence.<sup>16</sup>

EXP1 requires the performance of an intentional act, unlike EXP2. An involuntary twitch can express discomfort in the causal sense, and a blush can express embarrassment in the causal sense. It would be best to sharply distinguish EXP1 and EXP2, on the one hand, from EXP3, on the other. The former two notions concern the expression of mental states, while EXP3 does not.

When I say 'Cassis is in France', my use of the sentence semantically expresses the proposition that Cassis is in France. When I say, 'I am thinking that Cassis is in France', on Bar-On's view, my use of the sentence semantically expresses the proposition that I am thinking that Cassis is in France, which has the same truth-conditions as the proposition semantically expressed by a use of 'AB is thinking that Cassis is in France'.<sup>17</sup> So unlike SE, Bar-On's Neo-Expressivism (hereafter NE) is in conformity to Semantic Continuity. Further, on NE, when I sincerely say, 'I am thinking that Cassis is in France, my utterance expresses, in both the action and causal senses, my thought that I am thinking that Cassis is in France'.<sup>18</sup>

Suppose that I sincerely say, 'I am in pain' when I am not in pain (as in the above example concerning an expectation effect). Then obviously I do not express pain in the causal sense, since no pain causes my linguistic behavior. Do I express pain in the action sense? Bar-On says that I express pain though I do not express my pain, since I have none. Similarly, according to Bar-On, an actor expresses pain without expressing his pain (since he is feeling fine).<sup>19</sup> I think that Bar-On's usage here is not

true to our ordinary concepts of expression. It seems that (non-semantic) expression is at bottom a causal notion. The actor is feigning at expressing pain, and in the expectation-effect case, I am, unbeknownst to myself, failing to express pain. So it seems to me that all expression in the action sense is expression in the causal sense as well, though the converse does not hold. Expression in the action sense, then, is a factive notion: if S expresses M in the action sense, then S is in M. From here on out, unless I explicitly speak of semantic expression, I will focus exclusively upon expression in the action sense.<sup>20</sup>

Bar-On holds that I can express my belief that it is raining in a ground-level manner by saying, 'It is raining'. I can express the same belief self-ascriptively by avowing, 'I believe that it is raining'. This affords us a novel take on Moore's Paradox. An assertion of 'It is not raining, and I believe that it is raining' can very well semantically express a true proposition but is nevertheless exceedingly odd. Though there is no contradiction between the two conjuncts, there is an expressive conflict, according to NE. My utterance of the first conjunct normally expresses my belief that it is not raining, while my utterance of the second conjunct normally expresses my belief that it is raining. So NE explains the oddity of the Moore-paradoxical assertion.

#### More on NE

Bar-On holds that the most vivid illustration of the expressive character of avowals is found in the case of what she calls avowals proper. These are avowals that are wrung from the speaker and not backed by any communicative intention: 'I love that painting!', 'I hate this mess!', 'I wonder what he's up to!'. Avowals proper may or may not have an audience. If they do, then these avowals will not be elicited by the audience, and they will not be directed towards the audience with an intention to produce a belief.<sup>21</sup> In avowals proper, it is quite clear, for Bar-On, that the speaker is expressing his mental state. What type of expression, though? In such cases, it is not clear whether the speaker's utterance counts as an intentional act. For example, an avowal proper is not backed by a Davidsonian belief-desire pair that concerns communication with an audience. So maybe avowals proper are closer to expressions in the purely causal sense than to expressions in the action sense.

Following Wittgenstein's lead, Bar-On holds that avowals generally are replacements for natural expressions of mental states. A natural expression of a child's desire for her stuffed bear is her action of reaching for it. A replacement avowal is 'I want Teddy'. Pain-behavior such as crying and wincing expresses the afflicted one's pain. This behavior is replaced by the avowal 'I am in pain'. Wherever it is plausible to suppose that a kind of avowal replaces a natural expression of a mental state, it is plausible to hold that the avowal also expresses the pertinent mental state. A prima facie problem arises, however. There are a great many avowals that cannot be seen as replacements for natural expressions of mental states. For example, there is no natural expression of an olfactory sensation of smelling vanilla. So we cannot say that the avowal 'I smell vanilla' expresses an olfactory sensation in virtue of replacing a natural expression of the sensation.<sup>22</sup>

Bar-On's solution to this problem is based on the idea that assertions such as 'That's a cat' express the beliefs of the speaker. These are the ground-level, non-avowing expressions of belief. Avowals can replace such ground-level linguistic expressions of mental states: 'I believe that that's a cat' replaces 'That's a cat'. But there is a problem, which Bar-On does not note, in applying this idea to the case of smelling vanilla. Suppose it is said that 'I smell vanilla' replaces 'There's vanilla nearby'. The latter assertion expresses the speaker's belief about the presence of vanilla, whereas the corresponding avowal expresses the speaker's olfactory sensation. So the avowal cannot be seen as an expressive replacement of the ground-level expression of belief.

### NE and the Distinctive Status of Avowals

Bar-On takes over SE's resources in explaining some aspects of the distinctive status of avowals. Insofar as an avowal is simply an expression of a mental state, rather than a report that is yielded by Cartesian inner sense or introspection, it "should indeed seem inappropriate to ask after the reasons...[the avower] has for the different aspects of the self-ascription she produces when avowing". (263) Asking for reasons is an inappropriate response to writhing and other naturally expressive pain-behavior. The same holds for avowals, insofar as they express pain and

other mental states. Just as wincing is protected from "the kinds of epistemic assessment and criticism that are appropriate to ordinary perceptual reports", so it goes for avowals. (263)

However, there is a disanalogy between NE and SE that gets in the way of the foregoing explanations. On NE, avowals semantically express truth-evaluable, contingent, empirical propositions, whereas on SE, this is not so. So on NE, avowals are in the "space of reasons". Thus, the Neo-Expressivist, unlike his Simple counterpart, needs to explain why demands for reasons and epistemic assessments are inappropriate, given that avowals make claims to truth.

Let us now turn to an argument that Bar-On gives for the truth of NE. It turns upon the assumption that the only mentalistic ascriptions that have distinctive, special security are avowals. Avowals, in turn, concern only one's own occurrent mental states. So these are the  favored states, i.e., the states represented by the mentalistic ascriptions that have distinctive security. It just so happens that the favored states are exactly those states which are expressible by speakers! Past and future mental states are not expressible, and neither are others' mental states. So we can explain the distinctive security of avowals by assuming the truth of NE.<sup>23</sup> But explain in what way? The fact that the favored, avowable states are the expressible states does not explain why being expressible is a feature that is relevant to the epistemic status--the distinctive security--of avowals of the favored states.

Let us continue to explore NE's approach to the distinctive security of avowals. One aspect of this security is the presumption that avowals are true. Bar-On holds that we can explain this presumption in the following way. You say, 'I am in pain'. I take your avowal to express your mental condition. Which condition? I consult my linguistic understanding. The proposition that is semantically expressed by your avowal concerns the condition of being in pain. So I take it that you are expressing your pain. By the factivity of the notion of expression (in the action sense), if you express your pain, then you are in pain. So I take it you are in pain. But your condition of being in pain is the truth-maker for the proposition that is semantically expressed by your avowal. So I take it that your avowal is true (in that it semantically expresses a proposition that I take to be true for the foregoing reasons).<sup>24</sup>

Thus, an avowal expresses the very condition that makes it true. By contrast, suppose that my utterance of

the non-avowal 'This is boring' expresses my boredom. It does not follow that I have said something true, because the condition of my being bored is not the truth-maker for the semantically expressed proposition. The truth-maker for the non-avowal is instead a condition in the world outside my mind, which may or may not be present even though I am bored.

The foregoing account of the presumption of the truth of avowals rests on the assumption that avowals express avowers' mental states. In a way, it seems to be a truism that avowals express avowers' mental states. But what exactly does this mean? It is a truism to say that avowals can express avowers' mental states. But is it a truism to say that avowals almost always express avowers' mental states, almost never fail to do so? Maybe this is a truism, but if it is, then it is nevertheless very close to our explanandum. We are trying to explain why avowals are presumed to be true. A crucial part of the explanation turns out to be that almost always, when a speaker makes an avowal, he is in the avowed state, which is a cause of his intentional act of avowing. That is what it means to say that almost always, avowals express avowers' mental states. One might well ask, why is there this almost perfect expressive success? Why are expressive failures involving false avowals so rare? NE seems to provide no answer to these questions.

### Self-knowledge

Bar-On distinguishes between these two questions: (1) Do avowals in fact possess the distinctive status we have been discussing? (2) Do avowals "represent", as she puts it, privileged self-knowledge? Bar-On holds that an affirmative answer to (1) does not force an affirmative answer to (2). In the end, though, she does answer (2) in the affirmative. A negative answer would be driven by the deflationist challenge:

To the extent that avowals as expressive performances are protected from doubt and epistemic correction because made on no epistemic basis, they cannot represent any kind of knowledge, let alone an especially secure form of knowledge. (342)

Bar-On seeks to meet this challenge.

Now if an avowal "represents" self-knowledge, then there must be a justified, true belief associated with the

avowal. This takes us back to the question of the relation between avowals and higher-order mental states, which arose in connection with the phenomenon of self-verification. If my avowal of 'I am in pain' "represents" self-knowledge, then the pertinent belief amounting to knowledge would be a second-order belief that I am in pain.

Many philosophers subscribe to the following Belief Expression Principle:

- (BE) If (a) a speaker sincerely utters sentence S, (b) the utterance is backed by competent linguistic understanding, and (c) the utterance of S semantically expresses P, then the utterance expresses the speaker's belief that P.

If BE is true, then an avowal expresses a second-order belief about the mental state that is self-ascribed by the avowal. One reason to think that there is this kind of expression is that avowals, as we are understanding them, are sincere utterances. The obvious way to understand that aspect of avowals is this: the avower believes the proposition that is semantically expressed by his avowal. Such a belief will be a second-order belief about the self-ascribed mental state.

Bar-On is not quick to accept the claim that avowals express second-order beliefs. Since she thinks that an avowal expresses the mental state that it self-ascribes, BE would yield the result that my avowal of 'I am in pain' expresses both my pain and my belief that I am in pain. Bar-On calls this the Dual Expression Thesis, and she is reluctant to accept it. But there are at least two strong reasons to accept it (beyond the aforementioned reason regarding the sincerity of avowals). First, the view that avowals "represent" self-knowledge requires that the avower at least possesses the second-order belief that he is in the mental state he self-ascribes in avowing. So why not simply add that the avowal expresses the second-order belief that the avower possesses? Second, if the Dual Expression Thesis is denied and BE is restricted to non-avowals, then this would constitute something close to a violation of Semantic Continuity. Whereas my utterance of 'I am standing' would express my belief that I am standing, my utterance of 'I am in pain' would not express my belief that I am in pain. Every non-avowing 'I'-ascription would express a belief of the semantically expressed proposition concerning the speaker, whereas no avowing 'I'-ascription would do so. This asymmetry seems very odd indeed.

Bar-On in the end accepts the Dual Expression Thesis. She says that in avowing, "a subject may be said to...[express in the action sense] her self-belief in virtue of actively making the self-ascription". She continues, "All there is to having a self-belief in the case at hand, it might be held, is the (intentional) issuing of a self-ascription". (366)

So avowals express higher-order "self-beliefs" which can amount to knowledge, on Bar-On's view. For Bar-On, a necessary condition for such a belief to constitute knowledge is that the belief be justified. On her view, these self-beliefs are not formed upon any "epistemic basis". This means at least that the self-beliefs expressed by avowals are not held on the basis of other beliefs of evidential propositions.<sup>25</sup> Bar-On worries that, on the face of it, this lack of an epistemic basis might preclude the pertinent self-beliefs from being justified. That is the worry arising from the deflationist challenge. As we will see, Bar-On assuages her worry. But a foundationalist would not see even a *prima facie* worry here. He could maintain that self-beliefs are like basic perceptual beliefs and like some beliefs of a priori propositions, in respect of a lack of evidential basis (at least, if evidence is thought to consist in believed evidential propositions).<sup>26</sup>

One way for the Neo-Expressivist proponent of privileged self-knowledge to proceed is to adopt a reliabilist theory of justification. Bar-On calls this the Low Road to Self-Knowledge. On this view, a self-belief is justified if it issues from a sufficiently reliable belief-forming process. This takes us back to the Introspectionist version of the Epistemic Approach to the distinctive security of avowals, which Bar-On rejected. She accordingly rejects the reliabilist approach to the justification of self-beliefs, revisiting several objections that we considered above. Reliabilism would make it a contingent matter that we have self-knowledge, since it is contingent matter that the mechanisms that generate second-order belief about first-order mental states are reliable. The possibility of a global failure of self-knowledge would arise. Further, the privileged status of self-knowledge would be compromised, since the reliabilist would construe the epistemic status of self-knowledge as being on a par with perceptual knowledge. The justification of both self-beliefs and perceptual beliefs would be seen as flowing from the reliability of the pertinent processes.

Let us now turn to Bar-On's favored account of the justification of the second-order beliefs expressed by

avowals.<sup>27</sup> When discussing the ascriptive immunity to error possessed by avowals, Bar-On offered the following cryptic formulation:

...when one avows, one has no reason or grounds for thinking that someone is in some mental state and that the state has some content other than whatever reason or grounds one has for making the judgement that is semantically expressed by the avowal: namely, that the avower is in state M (with content c). (223)

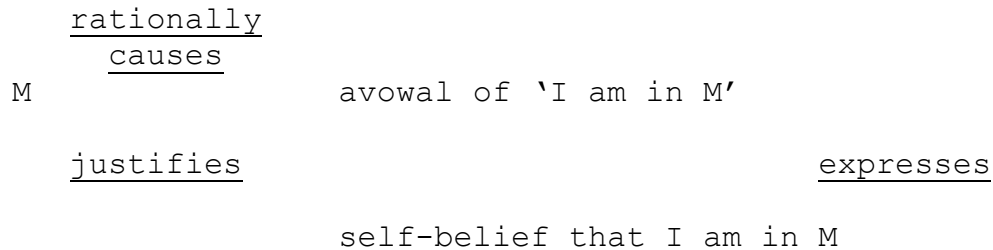
One might well wonder why an opponent of the Epistemic Approach to avowals would speak of "the reason or grounds" one has for making an avowal. For Bar-On, we have repeatedly seen, avowals have no epistemic basis. What Bar-On has in mind here becomes clear, though, when we come to understand her account of the justificatory aspect of self-knowledge.

Let us recall that the notion of expression that is crucial for NE is expression in the action sense. In this sense, an avowal expresses the avower's mental condition in virtue of being an intentional action. In general, some causes are rational causes: they are causes of intentional actions that serve to rationalize the action, such as a Davidsonian belief-desire pair. According to Bar-On, when I avow 'I am thirsty', my feeling thirsty is "my reason for the act" of avowing. (390) She allows that I may have other reasons as well. Though Bar-On does not say so, I take it that I may have communicative reasons as well, if the avowal is not an avowal proper. I (1) intend to get my audience to believe that P, and (2) intend that he get that belief as a result of recognizing my intention (1). Or so the Gricean story goes.

So for Bar-On, my feeling thirsty is the, or at least a, rational cause of my act of avowing 'I am thirsty'. Bar-On goes on to say that my condition of feeling thirsty is also "the epistemic reason for my self-ascription, or what warrants me in avowing". (390) The self-ascription that is semantically expressed by my avowal is "epistemically grounded in the avowed state", viz. my feeling thirsty.

Even if one finds this picture to be congenial, it needs to be slightly tweaked. In discussing the justificatory aspect of self-knowledge, the issue to be confronted is the epistemic status of second-order, self-beliefs, for it is such beliefs that amount to, or constitute, self-knowledge under the right conditions (such as the satisfaction of the justification condition). Bar-On

has said that when I avow 'I am in M', the avowal--the self-ascription--is grounded in, epistemically warranted by, the avowed state M itself, which is an epistemic reason. But we are inquiring into the justified, warranted, grounded-in-reasons status of the second-order self-belief expressed by the avowal. So it would seem better for Bar-On to say that that self-belief is justified by, warranted by, epistemically grounded by M itself, which is an epistemic reason for belief. So the following picture emerges:



M rationally causes my avowal; the avowal expresses both M and my belief that I am in M; M justifies my belief that I am in M.

A peculiarity of this conception of the justification of second-order self-beliefs is that the truth-maker for such belief turns out to be identical to the justifier. More precisely, the truth-maker for the believed proposition is the justifier for the belief of the proposition.<sup>28</sup> This is in contrast to all other sorts of epistemic justification. For example, the perceptual belief that I express by saying 'That cat is black' has as its truth-maker the state of affairs of the pertinent cat's being black. Again, more precisely, that state of affairs is the truth-maker for the believed proposition. The external-world state of affairs involving the cat is obviously not identical to the justifier of my belief about the cat. That justifier is my visual experience of the cat. The same point holds for non-empirical beliefs such as that which I express by saying '2+3=5'. The mathematical state of affairs that is the truth-maker here is obviously not identical to the justifier (which is presumably my grasping of the state of affairs, or my rational intuition of its obtaining).

Bar-On notes an objection to her picture of self-knowledge which arises from the structural peculiarity just noted. On a fallibilist conception of epistemic justification, it is possible for a thinker to have a justified yet false belief, i.e., a justified belief of a

false proposition. As we have seen, on Bar-On's view, false avowals are possible. But on her view, the self-beliefs that are expressed by false avowals cannot be justified beliefs. For if the proposition that is semantically expressed by a false avowal is indeed false, then this will be so in virtue of the absence of the avowed mental state. This means that the proposition's truth-maker will be absent, and hence the potential justifier of the self-belief will be absent. The self-belief expressed by the false avowal will thus be unjustified.<sup>29</sup>

Bar-On accepts this consequence of her view, likening it to disjunctivism about perceptual experience.<sup>30</sup> On that view, it is denied that subjects in the good case and in the bad case possess tokens of a single psychological-state-type: the good case is that in which I believe that I have hands in virtue of seeing my hands, and the bad case is that in which I am, say, a brain in a vat and mistakenly believe that I have hands on the basis of computer-generated illusions.<sup>31</sup> According to the perceptual disjunctivist, there is no common perceptual factor between these two cases. Similarly, on Bar-On's view, there is no common epistemic factor between a case of a true avowal and a case of a false avowal (let the cases be exactly similar but for the absence of the avowed mental state in the bad avowing case). Even if one accepts disjunctivism about perceptual experience (say, because one holds that perceptual experiences in good cases have object-involving propositional contents), Bar-On's disjunctivism has the further counter-intuitive consequence noted above: the structure of justification second-order self-beliefs is completely different from that of all other beliefs, in that the truth-maker is the same as the justifier. This has the consequence that the self-beliefs expressed by avowals are the only beliefs that contravene the fallibilist conception of justification.<sup>32</sup>

Let us consider another problem for Bar-On's view of the justification involved in self-knowledge. I do not wish to suggest that Bar-On is somehow committed to a reliabilist account of the justification of the self-beliefs expressed by avowals. However, there are similarities between her own account and that of the reliabilist. On her picture, the expressed self-beliefs are almost always justified--and almost always amount to instances of self-knowledge--in virtue of these beliefs' being caused by the mental states they are about, these mental states themselves constituting the justifying epistemic reasons for the beliefs.<sup>33</sup> It seems fair to say that it is a contingent matter that there

exists this widespread, nearly perfect correlation between intentional acts of avowing, which express self-beliefs, and the presence of avowed states, which render true the self-beliefs (i.e., by rendering true the believed propositions, such as that I am thirsty). Thus, it is a contingent matter that avowals almost always express self-knowledge: it is a contingent matter that the justifying truth-makers (the avowed states) of self-beliefs are almost always present. It seems, then, that nothing rules out the possibility of a global failure of self-knowledge. These are precisely the same objections that Bar-On laid at the door of the reliabilist approach to self-knowledge, and they mirror her objections to Introspectionist account of the distinctive security of avowals.<sup>34</sup>

### Concluding Difficulties

Let us conclude by considering two final worries about NE. The first concerns the distinctive security of negative avowals, such as 'I do not want coffee'. NE obviously cannot account for the security of this avowal by holding that it expresses an absent mental state. Bar-On's strategy for dealing with this problem is to find a mental state that is, so to speak, the "complement" of the absent one. In the case at hand, we could view the avowal as expressing the mental state of wanting not to have coffee. But this strategy will not generalize to all negative avowals. For example, when I avow, 'I do not believe that there are exactly ninety-nine wombats in the Ukraine', I am not thereby expressing a belief that there are not ninety-nine wombats in the Ukraine. For I suspend judgement on the Ukrainian wombat population. Bar-On's response to this problem is to construe my utterance not as an avowal at all but rather as a "self-report" of my "negative finding" regarding my present state of mind, a report that is "open to doubt, questioning, requests for reasons/explanation". (335) For my own part, however, I find no phenomenological difference between my negative avowal about wombat-belief and my positive avowals of belief about other subject matters.

The final worry about NE concerns the possibility of unarticulated self-knowledge. Can there be self-knowledge that is not expressed by an avowal, neither by an outward speech act nor by an inner episode? If so, then NE does not

explain the existence of such self-knowledge, for such self-knowledge would not involve the expression of the known mental state. Suppose that I know that I am in pain and yet fail to make a corresponding avowal in an outward speech act. Does it follow that I must be making an inner, silent avowal that expresses my pain? The suggestion that this must be so seems to run counter to phenomenology. Bar-On, though, would respond as follows: "thinking that one is in M is just self-ascribing M in thought". (408) By the same token, she could maintain, believing that one is in pain (as one silently suffers) is just self-ascribing pain in thought. So if I do believe that I am in pain in the case at hand (as is required by our assumption that I know that I am in pain), then I am self-ascribing pain in thought--a silent, inner avowal that expresses my pain.

This leads to a question about the explanatory utility of NE. According to Bar-On's account of the justification of a self-belief that is expressed by an avowal, the justifier is the avowed mental state that is the rational cause of the avowal. This state constitutes the justifying, epistemic reason for the self-belief. It is not clear in the end what role is being played here by the idea that avowals express mental states. One who is dubious about the epistemic role of the notion of expression could simply take over Bar-On's account of the justification of self-beliefs and remain silent on the question of expression.<sup>35</sup>

### Conclusion

In sum, the main worry that I have about Bar-On's Neo-Expressivism concerns the epistemological force of the notion of expression (in the action sense). Let us grant that when a speaker makes a true avowal (an avowal which semantically expresses a true proposition), he performs an intentional act which expresses the state that he therein self-ascribes. But why is it that avowals almost always succeed in expressing the self-ascribed states? Why is it, then, that avowals are almost always true? Let us set aside the truthmaker-as-justifier account of self-knowledge for a moment (since it does not hinge on the notion of expression) and ask: What role does the idea that avowals are expressions play in accounting for self-knowledge? Granting that a given avowal is an expression of a mental state, how does this explain the fact that a belief of the semantically expressed, truth-evaluable, contingent proposition requires no ordinary evidence and is not open to epistemic assessment? Finally, it is a contingent matter

that avowals almost always express self-ascribed mental states. Thus, if the notion of expression plays a central role in explaining the distinctive epistemic status of avowals and the existence of privileged self-knowledge, these epistemic phenomena would seem to be merely contingent as well.

Anthony Brueckner  
University of California, Santa Barbara

---

<sup>1</sup> Speaking My Mind: Expression and Self-Knowledge (Oxford: Clarendon Press, 2004).

<sup>2</sup> Sentences lacking the first-person can also be used to make avowals: ‘My pain is intense’.

<sup>3</sup> In what follows, I will sometimes refer to avowals simply by specifying a sentence, rather than by specifying a particular use of, or utterance of, a sentence.

<sup>4</sup> As we will see, it is not clear that Bar-On wishes to think of avowal as assertions.

<sup>5</sup> Bar-On characterizes Semantic Continuity by saying that avowals are “interchangeable salva veritate (in context) with ordinary, unproblematic [non-avowing] statements”. (9) This seems to be a bit misleading if we are considering sentential contexts, since the avowals we are most interested in do not occur embedded within a larger sentential context.

<sup>6</sup> Philosophical Investigations, section 244 (Oxford: Blackwell, 1953). The pronoun ‘it’ presumably should be taken to refer to pain, not crying.

<sup>7</sup> See “The First Person”, in Samuel Guttenplan (ed.), Mind and Language: Wolfson College Lectures, 1974 (Oxford: Oxford University Press, 1975).

<sup>8</sup> “Assertion”, in Philosophical Review 84, 449-65 (1965).

<sup>9</sup> One of the reasons why Anscombe thinks that the first person fails to refer in avowals is that she thinks that if ‘I’ did have a referent in avowals, a Cartesian ego would be the likely candidate.

<sup>10</sup> See Tyler Burge’s “Our Entitlement to Self-Knowledge”, in Proceedings of the Aristotelian Society 96, 1-26 (1996).

<sup>11</sup> One disanalogy between introspection and perception is that in introspection, there is nothing analogous to the perceptual states upon which beliefs about the objects of perception are based.

<sup>12</sup> See, e.g., Crispin Wright’s “Self-Knowledge: the Wittgensteinian Legacy”, in Crispin Wright, Barry Smith, and Cynthia MacDonald (eds.), Knowing Our Own Minds (Oxford: Oxford University Press, 1998), 13-46.

<sup>13</sup> See Sydney Shoemaker’s “Self-Reference and Self-Awareness”, in Journal of Philosophy 65, 555-67 (1968).

<sup>14</sup> The locus classicus for discussion of the notion of self-verification is Tyler Burge’s “Individualism and Self-Knowledge”, in Journal of Philosophy 85, 649-63 (1988).

<sup>15</sup> See “Language as Thought and as Communication”, in Philosophy and Phenomenological Research 29, 506-27 (1969).

<sup>16</sup> Bar-On says that a sentence semantically expresses thoughts and judgements. We must be careful to think of thoughts and judgements in this context as being propositions rather than mental states.

<sup>17</sup> Bar-On distinguishes avowals as acts (presumably speech acts) from avowals as products. An avowal as product is “a linguistic, (or language-like) token, an item with certain semantic properties”, which is the product of an act of avowing. (251) In this paper, I have so far been focusing upon avowals as acts. We can now say that an avowal as product is that which semantically expresses a proposition.

<sup>18</sup> Though she does not directly address the question, the expressive, non-reportive character of avowals might conflict with my earlier claim that avowals are assertions of propositions. Bar-On thinks that avowals semantically express truth-evaluable propositions, and she speaks of the truth and falsity of avowals (presumably, a true avowal is an avowal which semantically expresses a true proposition). Thus I do not see

why we cannot say that avowals both (a) express mental states and (b) are assertions of the propositions they semantically express.

<sup>19</sup> The actor's utterance of 'My pain is excruciating!' cannot be considered to be an avowal, since the utterance is not sincere.

<sup>20</sup> Bar-On says that in the action sense of expression, a person expresses his mental state. However, I will often simply say that a person's avowal (his utterance of a sentence) expresses his mental state.

<sup>21</sup> There is no discussion of Paul Grice's work on the connection between meaning and speakers' intentions in Bar-On's book. See Studies in the Way of Words (Cambridge: Harvard University Press, 1989).

<sup>22</sup> See Wright's "Self-Knowledge: the Wittgensteinian Legacy".

<sup>23</sup> See p. 284 for this argument, which I have reconstructed a bit.

<sup>24</sup> See pp. 316-7.

<sup>25</sup> Some self-beliefs are evidentially based, as in the psychoanalytic case discussed above, which does not involve an avowal.

<sup>26</sup> A disanalogy: whereas perceptual beliefs may be thought to have an epistemic basis in experience, and my simple arithmetical beliefs may be thought to have an epistemic basis in rational intuition, there is no corresponding epistemic basis for self-beliefs.

<sup>27</sup> Bar-On also discusses what she calls the High Road to Self-Knowledge, along which various philosophers argue that having justified self-beliefs is somehow a condition for the possibility of rationality. I discuss some examples of the High Road in the following papers: "Two Recent Approaches to Self-Knowledge", in Philosophical Perspectives 13, 251-71 (1999), and "Two Transcendental Arguments Concerning Self-Knowledge" in New Essays on Semantic Externalism and Self-Knowledge, Susana Nuccetelli (ed.), (Cambridge: MIT Press, 2003), 185-200.

<sup>28</sup> For similar views, see the following papers: Brian McLaughlin and Michael Tye, "Is Content-Externalism Consistent with Privileged Access?", in Philosophical Review 107, 349-80 (1998); Aaron Zachary Zimmerman, "Basic Self-Knowledge: Answering Peacocke's Criticisms of Constitutivism", in Philosophical Studies 28, 337-79 (2006).

<sup>29</sup> For discussion of this problem, see my "Problems for a Recent Account of Introspective Knowledge", in Facta Philosophica 3, 69-75 (2001). See also Brian McLaughlin's "Introspecting Thoughts" in the same journal, 77-83.

<sup>30</sup> See pp. 391-6 for Bar-On's attempt to take some of the sting out of her form of disjunctivism.

<sup>31</sup> The terminology derives from Timothy Williamson's Knowledge and its Limits (Oxford: Oxford University Press, 2000). A locus classicus for disjunctivism is Paul Snowdon's "Perception, Vision, and Causation", in Proceedings of the Aristotelian Society 81, 175-92 (1980-1).

<sup>32</sup> I am assuming that fallibilism holds for all other kinds of beliefs.

<sup>33</sup> Reliabilists do not highlight epistemic reasons for self-beliefs, as does Bar-On. Further, a reliabilist will reject Bar-On's disjunctivism: a belief expressed by a false avowal will be justified if the belief-generating process is sufficiently reliable. See my "Problems for a recent Account of Introspective Knowledge".

<sup>34</sup> Thanks here for a suggestion by Carmen McDonald.

<sup>35</sup> See, e.g., the papers in fn. 28, in which a view of self-belief-justification close to Bar-On's is presented independently of any treatment of the notion of expression.