

## 6

# Expression, Truth, and Reality: Some Variations on Themes from Wright

*Dorit Bar-On*

## 1. Introduction: Expressivism as an 'Anti-Realist Paradigm'

Expressivism, broadly construed, is the view that the function of utterances in a given area of discourse is to give expression to our sentiments or other (non-cognitive) mental states or attitudes, rather than report or describe some range of facts. This view naturally seems an attractive option wherever it is suspected that there may not be a domain of facts for the given discourse to be describing. Familiarly, to avoid commitment to ethical facts, the *ethical expressivist* suggests that ethical utterances (e.g. “Gratuitous torture is wrong,” “What John did was morally good”) do not serve to ascribe ethical properties to objects, actions, persons, or states of affairs. Instead, they simply function to give voice to certain of our sentiments (or ‘pro/con’ attitudes). Along similar lines, philosophers have entertained versions of expressivism about the aesthetic, the modal, the mental, what is funny, even about theoretical science and knowledge ascriptions.<sup>1</sup>

In several of his writings, Crispin Wright lists expressivism among the chief ‘anti-realist paradigms’ and not a successful one at that. Insofar as expressivism purports to allow us to avoid unwanted ontological commitments, its success depends on denying—implausibly—that the relevant discourses “really deal in truth-evaluable contents.” Yet Wright thinks that these discourses *do* possess truth-evaluability, since they exhibit all the relevant ‘assertoric trappings,’ which, on the ‘minimalist’ conception Wright develops in *Truth and Objectivity*, suffice for truth-aptness and truth.<sup>2</sup> If it were indeed true that expressivism, in every case, required denying truth-evaluability, then

<sup>1</sup> See Wright 1988, pp. 46–47, endnote 12.

<sup>2</sup> See Wright 1992, Chapter 1, and Wright 1996 for a summary.

Wright would be correct to declare the expressivist move ultimately a ‘faux pas’ (Wright 1988: 34). For, in that case, expressivism would require compromising the palpable *semantic continuity* between putatively expressive and ordinary descriptive discourses. Ethical sentences, for example, behave just like ordinary descriptive sentences syntactically as well as logico-semantically. They can be involved in logically valid inferences, they admit negation, embed in conditionals and in propositional attitude contexts, and so on.<sup>3</sup> To maintain expressivism as a thesis that directly concerns the *semantics* of ethical, aesthetic, psychological, modal, etc. sentences—a thesis that treats them as effectively *synonymous* with verbal expressions such as “Good for you!,” “Boo!,” “Yuk!,” “Ouch!” and so on—is to do serious violence to a wide array of linguistic appearances and intuitions.

I think that the alliance standardly assumed to exist between expressivism, on the one hand, and the denial of truth-evaluability as well as anti-realism, on the other, is unfortunate, though historically understandable. With the proper distinctions in place, I believe we can decouple the core expressivist idea from the denial of truth-evaluability and break its apparently necessary link to anti-realism. The result will, I hope, present us with a viable option in various areas where philosophers have been drawn to expressivism—an option which I also hope Wright himself will welcome.

I begin by discussing one kind of expressivism, concerning avowals and self-knowledge—“avowal expressivism,” for short. I present a simple version of avowal expressivism and some reasons canvassed by Wright for rejecting it (Section 2). Wright, however, maintains that avowal expressivism fares better than ethical expressivism. In Section 3, I take up the comparison between these two expressivist views and argue that, understood as a radical anti-realist view, simple avowal expressivism is in fact incoherent. In Section 4, I consider an alternative anti-realist view of first-person authority that Wright himself flirts with in several places under the title “the default view.”<sup>4</sup> I argue that it shares a defect with simple avowal expressivism. I then turn (in Section 5) to my own *neo-expressivist* view of avowals, which, I submit, is not only independently plausible but also superior to the default view. I will not here argue for the independent plausibility of this neo-expressivist view.<sup>5</sup> Instead, in Section 6, I will explain how the view avoids the incoherence of simple avowal expressivism, and indicate how it can be generalized to other cases so as to yield a new expressivist paradigm, one that allows for truth-evaluability and avoids anti-realist commitments.

<sup>3</sup> These are referred to by Wright under “Geach’s point” (Wright 1998, sec. 3, especially p. 31, see also Geach 1960).

<sup>4</sup> Wright 1998, p. 41.

<sup>5</sup> For a full treatment, see Bar-On 2004. Though I have so far spoken of avowals as *utterances*, my account is intended to apply to avowals made in thought, and not just in speech. (The distinction between avowals and other kinds of self-ascriptions is relevant to thought as well as speech. A person’s spontaneous thought: “I’m sick of this!” has a different epistemic status from thoughts she might have about how she felt yesterday, about the position of her limbs, about others’ mental states, etc.)

In all this, I take myself to be offering some variations on themes near and dear to Wright's heart. I do not expect to be doing justice to the richness of Wright's own treatment of these themes, nor will I be attempting a careful exegesis. Nonetheless, I do hope to be articulating a view that Wright himself may find congenial—or else learn why he does not.

## 2. Avowal Expressivism

Consider everyday utterances such as “I have a terrible headache,” “I feel sick,” “I'd like some tea,” “I'm scared of that dog,” “I'm wondering whether it's going to rain,” “I hope that I'm not late.” On their face, such utterances—“avowals,” as they are often called—resemble other everyday utterances in which we report various occurrences, such as “There is a red cardinal at the bird-feeder,” “We've just run out of dog food,” or “I have a mosquito bite on my leg.” Like these reports, avowals appear to inform us of certain contingent states of affairs; they tell us of states of mind that the speaker happens to be in at a given moment. Yet when compared to other sorts of utterances, avowals appear to enjoy a special status and to be especially secure. Thus, avowals contrast with third-person present-tense ascriptions of mental states (e.g. “She is very tired” said of me by someone else, or “John is very tired” said by me about someone else), or past-tense mental self-ascriptions (“I was very tired then,” said by me at a later time). Avowals also contrast with non-observational bodily self-ascriptions (“My legs are crossed,” “I'm sitting down”) and present-tense self-attributions of psychological traits or standing dispositions (“I'm a patient person,” “I like going to restaurants”). It is also interesting to consider that an ordinary avowal of fatigue would seem much more secure than some present-tense self-ascriptions of *occurrent* mental states, such as “DB is very tired,” if it were said by me about myself when I fail to recognize that *I* am DB, or “I am very tired,” if it were said by me on the basis of, say, looking in the mirror, or inference from some test results.

Unlike these other utterances, avowals appear to be, as Wright puts it, ‘groundless’: they are seemingly made on no epistemic basis, and it would normally be considered out of place to ask of a person issuing an avowal for reasons or justification.<sup>6</sup> Yet avowals enjoy a *special security*: except under special circumstances, avowals are not subject to correction and are taken at face value. Whereas it seems that other pronouncements are only as secure as their epistemic basis, avowals appear remarkably secure *even though* they are apparently made on *no* epistemic basis.<sup>7</sup>

<sup>6</sup> Wright 1998, p. 14.

<sup>7</sup> Wright distinguishes between phenomenal and attitudinal avowals in point of their security (Wright 1998, pp. 14–17). The former, he thinks, are ‘strongly authoritative,’ whereas the latter are only ‘weakly authoritative.’ By contrast, Bilgrami 1998 argues that, properly understood, attitudinal avowals are incorrigible (Bilgrami sets aside phenomenal avowals altogether in his account of first-person authority). For reasons I detail in Bar-On 2004, I don't think much weight should be placed on the phenomenal/attitudinal divide when trying to understand the distinctive security of avowals.

As the Cartesian tradition would have it, avowals enjoy *absolute* security: they are absolutely indubitable, incorrigible, and infallible. Moreover, avowable mental states are thought to be ‘self-intimating’: one who is in a mental state is guaranteed to know it. There may be good reasons to depart from the Cartesian tradition on these scores. But even so, the *epistemic asymmetry* between avowals and all other ascriptions (including, as can be seen from the above list, various kinds of *self*-ascriptions) calls for an explanation.<sup>8</sup>

If we are no longer wedded to absolute incorrigibility or infallibility as the mark of avowals, it may be tempting to play down the contrast between avowals and other self-ascriptions. Specifically, we can note that there is a class of bodily predicates, such as “crossed legs,” or “is sitting down,” whose application exhibits a certain epistemic first-person/third-person asymmetry. Such predicates are applied on the basis of external observation in the third-person case, but on the basis of proprioception or kinesthesia in the first-person case. Likewise, we can think of mental predicates “in pain,” or “is angry at  $y$ ,” and so on, as predicates that are applied on different epistemic bases in the third- and first- person cases. They apply on the basis of observation of behavior, or inference, or conjecture, in the third-person case, but, in the first-person case, though they appear to apply on *no* basis, they in fact apply on the basis of *introspection*.

Suppose, *contra* the Cartesian understanding, we think of introspection in the materialist’s preferred terms, as a faculty whose job is to deliver reliable though *not* infallible reports on certain goings-on *inside* our bodies (more specifically, in our brains), which reports are articulated by avowals. Might we not then hope to have a solution to our puzzle regarding epistemic asymmetry without paying the heavy metaphysical cost incurred by the Cartesian introspectionist?<sup>9</sup>

The *materialist introspectionist* view (“MI” for short) identifies a distinct epistemic ‘route’ that we can associate with avowals. But doing so cannot by itself allow us to understand the perceived contrast in degree of security between avowals and certain bodily self-reports, such as those arrived at through proprioception. After all, I have a different way from my observers of learning where my limbs are, or whether I am sitting down, but my word on these matters is not clearly *better* than theirs. On a given occasion, my body could be in a state that my proprioceptive mechanism cannot distinguish from the state of my legs being crossed. I could also prove to be *systematically* unreliable in reporting my limb positions or bodily movements without looking. The ability to tell my bodily conditions through proprioception or kinesthesia thus seems *alienable*. Furthermore, we can even conceive of someone whose brain was so hooked up to my limbs as to receive direct information about their position. Such a person

<sup>8</sup> Wright does not endorse absolute Cartesian security, though he does include self-intimation (which he refers to as ‘transparency’) among the key features of ‘avowals’: “in the normal run of cases, the subject’s ignorance of the truth or falsity of an avowal . . . is not, it seems, an option”—Wright 1998, p. 15.

<sup>9</sup> The discussion that follows makes several points of contact with Wright’s rejection of the ‘observational’ model. See, for instance, Wright 1987, pp. 399–402, Wright 1989, sec. IV, and Wright 1998, pp. 22–28.

could inherit my proprioceptive ability, and would be able to tell where my limbs are the same way I do. Proprioception (as well as kinesthesia) thus seems also entirely *transferable*.<sup>10</sup>

If the distinctive security of avowals were entirely due to the epistemic security of introspective mechanisms, we might expect so-called ‘first-person authority’ to be equally alienable and transferable. Yet I agree with Wright that, by commonsense lights, it isn’t. It is one thing to show that various mental self-ascriptions we make are false; it is another to show that we can make sense of the idea of a subject who, though she *has* mental states, is ‘chronically unreliable’ in her avowals.<sup>11</sup> Even more problematic is the notion that I might be able to transfer my first-person authority to someone else—someone who was able to ascribe to me reliably and correctly present mental states without relying on any observation or evidence.<sup>12</sup> For one thing, my avowals would surely constitute a crucial part of the data against which to measure my putative mindreader’s claim to be able to provide reliable non-evidential reports of my occurrent mental states. But more importantly, so long as there was no question about my linguistic competence or my sincerity, if disagreement broke out between us over what is now going on in my mind, her consistent past success in reading my mind would not be sufficient ground for taking her word over mine. It seems, rather, that we would take the disagreement as signaling the waning of her mind-reading powers.<sup>13</sup>

There is another sort of error that should be possible on the MI model: *brute local error*—an error that is not due to any defect of any of my psychological mechanisms but is simply due to the fact that my internal introspector has been ‘fooled’ into false detection. On the commonsense view, though, the possibility of brute local error seems no less problematic than the possibility of global systematic failure, and much more problematic than MI would have us expect. It’s not that commonsense does not allow for false avowals. As Wright notes, false attitudinal avowals—e.g. of beliefs, intentions, hopes, etc.—can be issued in self-deception or as a result of wishful thinking (Wright 1998, p. 17). I also think that under certain circumstances, even a sincere phenomenal avowal may be thought false. For example, as you sit on the dentist’s chair and say “My tooth hurts!” before the drill reaches your mouth, your dentist may sensibly question whether you really are in pain, even without questioning your sincerity.<sup>14</sup> But although these are cases in which

<sup>10</sup> See also the kaleidoscope analogy in Wright 1998, p. 22.

<sup>11</sup> Wright says that, even in the case of attitudinal avowals, which he thinks are only ‘weakly authoritative,’ “[w]holesale suspicion about my attitudinal avowals—where it is not a doubt about sincerity or understanding—jars with conceiving of me as an intentional subject at all”—Wright 1998, p. 17.

<sup>12</sup> Wright 1998, p. 24.

<sup>13</sup> Similar remarks would apply to the possibility entertained by Rorty of a brain-scanner set up to replace one’s own self-scanning mechanism and to serve as a mechanical mind-reader. See Rorty 1970.

<sup>14</sup> I discuss other examples of sincere but false phenomenal avowals in Bar-On 2004, pp. 329–35 and 394–6.

I thus disagree with Wright that phenomenal avowals differ from attitudinal avowals in being (treated as) ‘strongly authoritative,’ meaning that someone’s making, or being disposed to make, an avowal is “a guarantee of the truth of what they say” so that “[a] doubt about such a claim has to be a doubt about the sincerity or the understanding of the one making it”—Wright 1998, p. 14.

we are prepared to question the truth of an avowal, they do not provide examples of brute error; for these are precisely cases in which the avowal's falsity is assumed to be due to some psychological irregularity, failure, or defect on the part of the avower. We don't suppose in any of these cases that the subject might have *simply* been *fooled* into issuing a false avowal by an 'uncooperative' mental world.<sup>15</sup>

Indeed, reflection on how best to understand cases of false avowals may help point us in the direction of an alternative solution of our puzzle. Consider again the above case of falsely (though sincerely) avowing a toothache at the dentist. At first blush, this kind of case may seem grist to the materialist introspectionist mill: your tooth doesn't really hurt, but your internal introspector mistook what is in fact (say) fear of the approaching drill for pain, and it is this finding that your false avowal represents. But how plausible is it to regard your exclamation "My tooth hurts!" as a report of your (false) belief that your tooth hurts? After all, under the circumstances, you might have equally said: "Ouch!," or emitted a yelp, or winced. Should we think of the wince as equally the upshot of a *false belief* that you formed about your internal state? Isn't it much more plausible to regard the exclamation: "My tooth hurts!" as on a par with the wince—as something forced out of you, though in this unusual case not by an actual toothache, but rather by the priming effect of the fear (as well as your previous history at the dentist)?

In this somewhat roundabout way, we have arrived at a well-known counter to introspectionism often attributed to the later Wittgenstein, namely, *avowal expressivism*.<sup>16</sup> This is the view that avowals are importantly similar to natural expressions in being bits of behavior that typically serve directly to express our present mental states rather than to report their presence. The key idea of *simple* avowal expressivism is that, although avowals closely resemble in their surface form ordinary descriptive reports, logically and semantically speaking they are more like natural expressions of sensations, such as moans and groans, grimaces, giggles, etc.

Now, natural expressions of sensations are spontaneous reactions that are in some essential way characteristic of the subject's relevant conditions. A subject in pain naturally groans, or cries out, or grimaces, or clutches the painful part. A subject being tickled will typically squirm and giggle. Natural expressions are not limited to sensations. There are natural expressions characteristic of emotions and feelings such as fear, anger, excitement, joy, etc. as well as of wants.<sup>17</sup> What is crucial about natural expressions of all these kinds is

<sup>15</sup> It is for this reason perhaps that full-blown 'internal world skepticism' (the analogue of Cartesian external world skepticism) seems incoherent. What is required to fuel such skepticism is the possibility of holding fixed all of our present 'judgments of appearances' concerning our internal world while questioning their veridicality. But plausible cases of false avowals are precisely *not* cases in which it is reasonable to suppose that it *appears* to the subject that some internal state of hers is a toothache (a belief that p, a desire for y, ...) when it isn't. (I discuss the contrast between external and internal world skepticism in Bar-On, forthcoming.)

<sup>16</sup> See §244 in Wittgenstein 1953.

<sup>17</sup> Compare Hacker: "A child who wants a toy reaches for it and tries to get it. The child's anger is manifest in striking out, contorted features, and screams or rage . . . if he is frightened, he blanches, cries, and runs to Mummy"—Hacker 1993, p. 88.

that they are in no way thought to represent a subject's *judgment* or *belief* about herself. A subject's smile is not supposed to represent the subject's *judgment* that she is pleased—rather, it is supposed to give expression to the pleasure itself. Correlatively, we do not expect the person who smiles, or cries in pain to give reasons for her smile or cry; we do not query or challenge a gasp of fear, or a sigh of relief, and so on. Insofar as we regard a form of behavior as a natural expression of a subject's condition, it seems entirely inappropriate to impose on it any epistemic assessment.

Simple avowal expressivism can be seen as offering to explain the epistemic asymmetries that give rise to our puzzle by breaking away from the face-value understanding of avowals' grammar. As a first step, the account maintains that, despite surface appearances, avowals of sensations have the same 'logical grammar' as natural expressions of sensations. That is to say, '*grammatically speaking*', mental terms as they are used in avowals do not have the semantic function of picking out mental states and ascribing them to individuals.<sup>18</sup> Avowals serve *only* to express subjects' mental states, just like cries, or wincing, or joyful hugs. As a second step we can then realize that, *epistemically speaking*, it is misguided to regard the special status of avowals as a consequence of a special access subjects have to their own states of mind, and there is no need to seek any secure epistemic basis on which avowals are made. If avowals are not vulnerable to doubt, and are not open to query, or correction, this is not because of the security of their epistemic basis, but because they do not serve to tell or describe the avower's present condition in the first place. Like natural expressions, they merely serve to express subject's mental states, and so, they belong in the wrong 'grammatical' category for any epistemic assessment. A subject who avows "I am in pain" is just like one who spontaneously lets out a cry: such a one cannot be legitimately asked to give reasons, questioned, or challenged.<sup>19</sup> And this can serve to explain the asymmetries between avowals and other ascriptions. But with this explanation of the epistemic asymmetries in place, we can finally agree with materialists that *ontologically speaking*, there is no need to invoke Cartesian entities to play the role of especially accessible mental items that are referred to by sensation (or any other mentalistic) terms.

The 'no-ascription' aspect of simple avowal expressivism is invoked to explain the epistemic asymmetries between avowals and other ascriptions while avoiding the Cartesian appeal to private objects. However, this very aspect is its downfall. If avowals do not involve genuine—let alone *true*—ascription of mental states to the avowing subjects, then they cannot have truth-conditional equivalents that *do* involve such ascription and with which they can be legitimately interchanged in certain contexts; and they cannot serve as legitimate premises in relevant logical inferences. In short, simple avowal expressivism requires explaining the epistemic asymmetries by compromising notable semantic continuities between avowals and other ascriptions.

<sup>18</sup> In Bar-On 2004, Chapter 7, I suggest that this move is analogous to Anscombe's denial that 'I' refers.

<sup>19</sup> Compare Wright 1998, pp. 34–5. Notably, however, we *do* sometimes 'take issue' with non-verbal expressive behavior. "You can't really be tired," we may say to someone who is yawning and stretching.

Avowals, understood as simply alternative ways of moaning and groaning would be invulnerable to error and incorrigible—indeed, absolutely so—but only in the degenerate sense of belonging to the wrong grammatical category. And they would be protected from epistemic assessment, but only by being altogether excluded by ‘grammar’ from the realm of linguistic meaning, judgment, and knowledge.

### 3. Simple Avowal Expressivism vs. Ethical Expressivism<sup>20</sup>

In its suggestion that avowals do not involve genuine ascription of (or reference to) mental states, but only serve to express them, simple avowal expressivism is reminiscent of traditional *ethical expressivism*. According to the latter view, ethical claims (e.g. “This is good/bad,” “Racial discrimination is unjust”) are mere expressions of certain of our emotions, preferences, or attitudes. They are not genuine, truth-evaluable assertions either about objective ethical states of affairs (as ethical objectivists would have it) or about the expressed attitudes (as subjectivists would have it). As is well known, traditional ethical expressivists, like Ayer and Stevenson,<sup>21</sup> were motivated in good part by wanting to avoid all commitment to an ethical realm of facts populated by ontologically suspect ethical properties. Rather than taking ethical claims to be assertions about our subjective attitudes, they suggested that ethical terms do not function to designate *any* properties but are rather terms of dis/approbation, dis/approval, or condemnation/commendation (much like “Hooray,” “Yuk,” “Super”).

Of the various complaints made against ethical expressivism over the years, the one most pertinent to the comparison with simple avowal expressivism is this: ethical expressivism seems unable to accommodate the semantic continuity of ethical discourse with ordinary descriptive discourses. In particular, it cannot accommodate the fact that ethical terms, just like ordinary descriptive terms, can occur in all sorts of grammatical contexts. For instance, suppose we were to accept that to say “Helping the poor is good” is only ever to express a pro-attitude toward helping the poor. Still, it is difficult to accept that the sentence is used that way when embedded in the context of a conditional: “If helping the poor is good, then we should tax the rich and give the money to poor people.” As Geach has famously argued,<sup>22</sup> in such a context, it seems as though the sentence makes a semantic contribution that is not properly captured by the expressivist analysis.

However, as Wright points out, the ‘Geach point’ can perhaps be handled by the ethical expressivist.<sup>23</sup> After all, take an utterance of a sentence such as “It usually rains

<sup>20</sup> Some of the material for this section is taken from a similarly titled section of Chapter 7 of my Bar-On 2004.

<sup>21</sup> See Ayer 1946, Chapter 6 and Stevenson 1944.

<sup>22</sup> Geach 1960.

<sup>23</sup> Wright 1995a, sec. 1, Wright 1998, sec. 3.

here at this time of the year” which normally has uncontroversially assertoric force. If we utter the same sentence as part of a conditional or embed it in any other ‘force-stripping’ context (negation, propositional attitude construction, etc.), it will no longer be used to make an assertion (though the sentences in which it is embedded may). This just means that from the possibility of embedding an indicative sentence in a force-stripping grammatical context we cannot infer anything about the illocutionary force that is standardly associated with it when it is uttered on its own. Given the familiar separability of force and surface grammar, this should come as no surprise. (“I’d like to know what time it is,” for example, is an indicative sentence that is standardly used to ask a question, rather than to make an assertion about oneself.) But then the ethical expressivist can maintain that, at most, the Geach point establishes that sentences containing ethical terms must have whatever semantic features that allow sentences to function (grammatically and logically) like ordinary indicative sentences. Thus the fact that ethical sentences can be embedded in force-stripping grammatical contexts does not tell against the expressivist claim that the standard function of ethical utterances is to express attitudes or sentiments rather than to make assertions about the world.

Applied to simple avowal expressivism, the Geach point would be that it is impossible to reconcile the claim that avowals merely serve to express the self-ascriber’s mental states, and in no way constitute assertoric reports about them, with a host of grammatical facts about avowals. Like ethical sentences, and like ordinary descriptive sentences (e.g. “I am 5’5” tall), avowals embed in force-stripping constructions (“If I’d like some ice-cream, then it would be nice for you to get me some,” “Rachel thinks that I’d like some ice-cream”), they admit of tense-transformation (“Yesterday, I wanted some ice-cream”), and so on. In these contexts, they do not plausibly serve to express the mental states apparently named in them. But the above response to the Geach point offered on behalf of the ethical expressivist can serve the simple avowal expressivist just as well. The Geach point can show that “I’d like ice-cream” must possess whatever features are required for a sentence to take indicative form. From this, it does not follow that the sentence must always be used to make assertions or descriptive reports about oneself, or that it cannot serve to express rather than report one’s desire outside force-stripping contexts.

Wright agrees that the Geach point “is powerless to determine that the standard use of a locution is to assert [truth-evaluable] content.”<sup>24</sup> However, he thinks that insofar as Geach’s observation exposes classes of sentences as capable of making positive contributions to truth-conditions (in force-stripping contexts), it shows that these sentences must possess “truth-evaluable contents.” For this reason, Wright seems to think that Geach’s point ultimately causes *more* trouble to ethical expressivism than to avowal expressivism. For the ethical expressivist is committed to the thesis that “there

<sup>24</sup> Wright 1998, p. 36.

are no real moral states of affairs; so the occurrence of what are apparently truth-evaluable contents couched in distinctively moral vocabulary has to be some kind of illusion” (ibid.).<sup>25</sup> In other words, in the ethical case, as in other familiar cases, the expressivist thesis is put to the service of an *anti-realist* ontological agenda. But the case of avowal expressivism, Wright seems to think, is different:

[since] it is no part of the present, allegedly Wittgensteinian expressivist proposal that there is no such thing as a statement of ordinary psychological fact. No one is questioning that ‘*He is in pain*’ is an assertion. The expressivist thesis distinctively concerns *avowals*.<sup>26</sup>

Wright’s treatment here brings up two issues. First, Wright’s idea that “the Geach point signals . . . the presence of truth-evaluable content” presupposes that all indicative sentences that are subject to Geach’s point must possess truth-evaluable content, *and* that the only way for a sentence to have such content is by having the kind of truth-conditions possessed by ordinary descriptive sentences. But suppose one accepts that indicative form must go hand in hand with truth-evaluability. Still, it seems that the ethical expressivist could try to explain how ethical sentences can *earn* indicative grammatical form, as well make systematic contributions to truth-conditions, even if ethical terms do not denote ethical properties.<sup>27</sup> This raises the general question whether expressivism must be saddled with the denial of truth-evaluability, even if it is coupled with anti-realism. Relevant here is the *minimalist conception* of truth and truth-aptitude that Wright himself recommends, according to which, roughly, possession by a discourse of the key features that underwrite the Geach point suffices for truth-evaluability.<sup>28</sup> For, on a minimalist conception, it would seem that one can be an anti-realist regarding a discourse *without* denying that its sentences are truth-evaluable.<sup>29</sup> But then it’s hard to see why an expressivist could not embrace the minimalist understanding and preserve truth-evaluability. Secondly, it is notable that, in the case of avowals, Wright appears to be allowing an expressivist view that *isn’t* anti-realist, thereby departing from his general characterization of expressivism as

<sup>25</sup> This claim may seem an odd one for Wright to make, given his minimalism about truth and truth-evaluability; see discussion below.

<sup>26</sup> Wright 1998, p. 36.

<sup>27</sup> For the separation between grammatical form and propositional content, on the one hand, and illocutionary force, on the other, see Austin 1961. For attempts by ethical expressivists to explain how ethical claims can earn propositional form in the absence of ethical reality, see e.g. Simon Blackburn 1984, pp. 181–94, 1993, essay 10, and Allan Gibbard 1990, Chapter 5, 2003, Chapters 3 and 4. However, both Blackburn and Gibbard’s ‘anti-realist’ proposal shy away from accepting that ethical terms can contribute to truth-conditions. Ridge 2006 defends an ‘ecumenical expressivist’ view that seems closer to the line I have in mind here. Below, I suggest that Wright’s own minimalism about truth-evaluability can also be roped into this service.

<sup>28</sup> Wright 1992, p. 29.

<sup>29</sup> See Wright 1996, p. 864. An ethical anti-realist might worry that accepting (even minimal) truth-evaluability for moral claims commits you to moral *beliefs* through a platitude according to which only beliefs can have truth-evaluable content. For discussion of this issue see Blackburn 1998, pp. 158–60. For partial response see Bar-On and Chrisman 2009.

an ‘anti-realist paradigm.’<sup>30</sup> Moreover, as I shall now explain, in the case of avowals, one *cannot* be an expressivist and at the same time be an anti-realist about mental states.

Let me begin by pointing out a difficulty for simple avowal expressivism that does not beset ethical expressivism. The difficulty has to do with the fact that avowal expressivism attempts to capture a contrast that arises within a single area of discourse—i.e. mentalistic discourse. This contrast concerns the *epistemic asymmetries* between avowals, on the one hand, and, most notably, mental ascriptions to others, as well as past-tense mental and present-tense bodily *self*-ascriptions, on the other. By contrast, ethical expressivism, at least in its traditional form, aims to capture an alleged *ontological difference* between ethics and other, straightforwardly ‘factual’ areas of discourse. The ethical expressivist who is moved by the ‘queerness’ of ethical facts<sup>31</sup> invokes the expressivist view in order to fill an apparent gap that opens up once we pull the metaphysical rug from under our ethical pronouncements—once we deny that there is a realm of facts for ethical claims to describe truly or falsely. We should distinguish, then, between

- (a) the positive *expressivist claim* that ethical proclamations regularly serve to express pro- and con- attitudes, rather than to make assertions about objective states of affairs,

and

- (b) the negative *ontological claim* that there are no ethical properties for ethical terms to refer to and no ethical facts for ethical sentences to describe,

as well as

- (c) the negative *semantic claim* that ethical sentences are not truth-evaluable, since they do not express true or false propositions.<sup>32</sup>

In the ethical case, if the expressivist can show how, despite the truth of (b), ethical sentences can earn truth-evaluability (perhaps by endorsing a minimalist conception thereof), she will have succeeded in offering a systematic, alternative conception of ethical discourse on which ethical claims need never be construed as ascribing ethical properties to some individual, or act, or state of affairs, and on which such claims, when made on their own, have the distinctive point of expressing certain of our sentiments or attitudes.<sup>33</sup> By contrast, the scope of simple avowal expressivism is limited to *a certain*

<sup>30</sup> It’s possible to take Wright here to be allowing only that third-person mental ascriptions are *minimally* true, and thus not *requiring* the independent reality of mental states. I think the argument that follows suggests that he must allow something stronger. In any event, he is at least leaving room for a realist construal.

<sup>31</sup> For locus classicus see Mackie 1977, Chapter 1.

<sup>32</sup> The closely related (but still separate) ethical *non-cognitivist* claim is that ethical sentences never express cognitive, truth-evaluable beliefs or judgments.

<sup>33</sup> The expressivist claim can be brought in *not* directly to fill in a perceived gap that opens once the negative ontological claim is embraced, but rather to capture some distinctive positive features of ethical discourse. See below.

*range of mentalistic utterances*: it encompasses (at most) those utterances that involve apparent ascriptions of mental states to oneself in the present tense.<sup>34</sup> The expressivist proposal here is that such utterances—avowals, as we have been referring to them—are mere expressions of subjects' present mental states. They do not represent assertoric reports about the subject's present mental states, and they involve no reference to those states. However, this proposal is at risk of prising mentalistic self-ascriptions too far apart from other ascriptions within a single (i.e. mentalistic) discourse.

Let me explain. Consider again my avowal: "I'd like some ice-cream." You could presumably use my avowal to infer: "She (DB) would like some ice-cream." And I could later conclude, remembering my avowal: "Earlier, I wanted some ice-cream." Semantically and logically speaking, these three sentences are transformations of each other that can support valid inferences and chains of reference. However, on simple avowal expressivism as fashioned after traditional ethical expressivism, we are to combine the idea that avowals serve only to express subjects' mental states with the negative ontological claim that there are no properties (or states/processes) for the mental terms contained in avowals to refer to the way ordinary descriptive terms do. But it is very hard to see how to reconcile this negative claim with the logico-semantic continuities between avowals and other mentalistic ascriptions. In the ethical case, the negative claim implies that ethical terms do not refer to genuine properties *wherever they occur*. Thus *if* the ethical expressivist can explain how ethical utterances can earn propositional form and content, even though there are no ethical properties for ethical terms to refer to, her work is done. The proponent of simple avowal expressivism, on the other hand, needs to explain how an avowal such as "I have a toothache," which allegedly makes no genuine reference to a state of an individual, can still be truth-conditionally equivalent to non-expressive, assertoric utterances which *do* involve such reference.

I have in mind here utterances of three types:

- (i) Third-person reports (e.g. "She/DB would like some ice-cream")
- (ii) Tense transformations (e.g. "I wanted some ice-cream earlier")
- (iii) Mentalistic present-tense self-reports (e.g. the admittedly unusual but nevertheless possible "I want some ice-cream" made on the basis of evidence or inference).

(In his discussion of avowal expressivism, Wright lists the fact that avowals submit to tense transformation along with the fact that they can be embedded in conditionals/negation under the heading "the Geach point." Wright also cites two additional facts: the fact that avowals can stand in logical relations to other statements ("I am in pain" →

<sup>34</sup> I say "at most," because some present-tense mentalistic self-ascriptions, namely, ones made on the basis of evidence, inference, testimony, therapy, or reflective self-interpretation do not exhibit the distinctive features of avowals. They are plausibly to be regarded as descriptive (fallible, corrigible) self-reports. Compare Wright 1998, pp. 15–16.

“Someone is in pain”) and the fact that they can be embedded in knowledge ascriptions, as in “He knows that I am in pain.”<sup>35</sup> I think that the case of tense-transformation ((ii) above), should be grouped together with cases of truth-conditionally equivalent third-person ascriptions (i) and of present-tense mentalistic self-reports (iii), both of which are omitted by Wright. The reason is that I take cases (i)–(iii) to illustrate how, *pace* Wright, avowal expressivism as presented so far faces a greater *prima facie* difficulty than ethical expressivism.)

It may seem that, in response, simple avowal expressivism could follow the ethical expressivist’s lead and adopt a more thoroughgoing non-referential account of psychological discourse, according to which psychological/mentalistic terms are *never* taken to refer to mental states. On this view—call it “psychological expressivism”—when making psychological ascriptions, all we ever do, whether in the first-person or third-person, whether in the present or past tense, is express various subjective conditions. So, for instance, saying of another “She is in pain” would be like emitting a sympathetic cry, thereby expressing one’s pity for her, as when a parent exclaims: “Ouch!” upon seeing her child hurt. Notice that this non-referential account goes beyond the (by now familiar) claim that psychological discourse requires special treatment, because it differs from non-psychological discourse in being subject to various normative constraints.<sup>36</sup> It also goes beyond the no-ascription account we canvassed at the end of the previous section. For, the account we are considering now requires maintaining that terms such as “is feeling pain” or “is afraid of the dog” or “is thinking about Vienna” simply do not serve to pick out states of individuals in *any* of their uses, first-person *or* third-person, present or past. Thus, an utterance such as: “DB has a headache” would be understood as (only) expressing your sympathetic attitude toward me and not at all referring to a state of mine. Or, perhaps it could be taken to express some other more complex attitude on the part of the attributer, having to do with her readiness to take a special ‘stance’ toward me (*à la* Dennett).

It should be clear that an account along these radical lines is not sustainable. The difficulty most relevant to us here can be seen by considering how we can consistently maintain that psychological terms do not refer, while at the same time holding that psychological attributions serve to *express attributers’ states of mind*.<sup>37</sup> For, it seems clear that this claim itself requires quantifying over states of mind. If psychological terms do not refer in any context, and sentences involving such terms never serve to represent truth-evaluable judgments about mental states, what are we to make of the expressivist account of avowals itself? The trouble for the psychological expressivist who opts for a thoroughgoing no-ascription account is that her alternative account of psychological

<sup>35</sup> Wright 1998, p. 35.

<sup>36</sup> See, e.g., Davidson, 1973 and 1984, and Dennett 1987, *passim*.

<sup>37</sup> Could the anti-realist avowal expressivist hold that her expressivist thesis *itself* is non-factual and serves only to express some of our attitudes (as suggested to me by Matthew Chrisman)? Perhaps. But then it is unclear to me how we are to assess the apparent *disagreement* between the psychological anti-realist and the psychological realist.

discourse must appeal to items of the very same kind she takes to be questionable, namely, mental states. Claiming that, even when ascribing mental states to others, we never report or describe those mental states, but only express our own attitudes and mental states, still requires acknowledging that there are mental states to *be* expressed. (This, of course, leaves open various accounts of what the nature of those states might be.) By contrast, claiming that when we make ethical pronouncements we never report or describe moral properties of things—people, acts, states of affairs—but only express our own pro- and con- attitudes toward things, only requires acknowledging that there are pro- and con- attitudes to be expressed; it does not require acknowledging that there are moral properties.

The foregoing can serve to elaborate on a complaint Wright makes against expressivism “applied to ordinary psychology under the aegis of metaphysical anti-realism.”<sup>38</sup> Focusing on what I described earlier as the expressivist’s *positive* claim, Wright points out that it requires separating states of mind into different kinds—feelings, as opposed to beliefs, for example. However the separation proceeds,

[t]he very claim that a discourse is expressive will presuppose an underpinning in facts about aspects of the characteristic attitudinal psychology of its participants. But facts of that genre are just what metaphysical anti-realism about psychology is unwilling to countenance. So expressivism can offer it no consistent outlet.<sup>39</sup>

The problem, in a nutshell, is that the expressivist’s positive claim depends on the invocation of the very same kind of states that the metaphysical anti-realist concerning ordinary psychological discourse wants to shun.

However, even if psychological expressivism could somehow avoid ‘teetering into incoherence’ (as Wright puts it), it should be recognized that it is not an attractive option in the context of solving the puzzle with which we began—viz., how to explain the distinctive security of avowals. Such an account, even if it were deemed adequate to explain the contrast between psychological and non-psychological discourses, would undermine the expressivist’s attempt to account for the asymmetries between avowals and *other* mentalistic ascriptions. As we have construed it, psychological expressivism attempts to account for *all* psychological talk, on the assumption that psychological terms do not refer *anywhere*, and thus that sentences containing them cannot be used to describe or report psychological facts. But then what is to account for the distinctive character of avowals *as contrasted* with third-person or first-person past-tense mental reports? As presented earlier, the central idea of simple avowal expressivism was that when *I* ascribe a mental state to myself (at least when I do so in the ‘avowing mode’), I do something different in kind from what others do when they ascribe a mental state to me. But this contrast would seem to escape the expressivist account construed as an anti-realist account that applies to *all* mentalistic discourse.

<sup>38</sup> Wright 2002, p. 212.

<sup>39</sup> Wright 2002, p. 213; see also Wright 1995b, pp. 204–5.

The upshot of our discussion so far is this. There is a crucial difference between expressivism as it has traditionally been advanced in metaethics and expressivism about avowals. In the ethical case, expressivism is invoked to preserve the cogency of ethical discourse even when it's admitted that it differs from ordinary descriptive discourses in lacking a proper grounding in matters of fact. If we take expressivism about psychological discourse to have a similar agenda, we court incoherence. Moreover, the puzzle for whose solution expressivism may be needed pertains to an *epistemological* contrast that arises *within* psychological discourse. And this puzzle is *not* addressed by psychological expressivism. Indeed, the puzzle remains even if we set aside ontological qualms about psychological discourse and take it to be straightforwardly factual.

#### 4. Wright's Default View

After explaining that avowal expressivism can adequately address the 'Geach point,' Wright raises other difficulties for the view, and concludes: "[T]he expressivist proposal flies rather further than is usually thought. But it is a dead duck all the same."<sup>40</sup> In its place, Wright considers another (allegedly Wittgensteinian) anti-realist view, which he dubs "the default view." In this section, I'd like to argue that the default view suffers a defect similar to the one identified for psychological expressivism toward the end of the previous section.

As Wright presents it, the default view, like avowal expressivism, and unlike the introspectionist view, denies that an avowal can represent a genuine "cognitive achievement, based on cognitive privilege" that can, in turn, explain so-called first-person authority.<sup>41</sup> On the default view, the first-person/third-person asymmetries "belong primitively to the 'grammar' of the language-game of ordinary psychology." The fact that psychological discourse exhibits these features should simply be regarded as part of what makes it *psychological* discourse.<sup>42</sup> On this account, then, the special status enjoyed by avowals is not due to the special nature of the states they ascribe (*viz.*, mental, as opposed to non-mental), or the special epistemic relation between subjects and their own mental states. Rather, it is "a *constitutive principle*" of mentalistic discourse, which "enters primitively into the conditions of identification" of the subject's mental states.<sup>43</sup>

In contrast to avowal expressivism, however, the default view does *not* deny truth-evaluability to avowals. Rather, it proposes a reconstrual of the truth-conditions of all psychological ascriptions. The truth-conditions of such ascriptions are governed by a special constraint, which Wright formulates as follows: "unless you can show how to make better sense of her by overriding or going beyond it, [a subject's] active self-conception, as manifest in what she is willing to avow, must be deferred to" (Wright 1998, p. 41). Instead of denying truth-conditions to avowals, then, the default view

<sup>40</sup> Wright 1998, p. 38. I address these difficulties in sections of Chapters 8–10 of Bar-On 2004.

<sup>41</sup> Wright 1987, p. 402.

<sup>42</sup> Wright 1998, pp. 41–3.

<sup>43</sup> *Ibid.*, p. 42.

maintains that mastery of the truth-conditions of all mentalistic ascriptions requires recognizing that, unless there are good reasons to the contrary, one must accept what a subject is willing to avow, and only make ascriptions to her in accordance with what she herself is willing to avow. The epithet “default view” is apt, since the view presents “a subject’s opinions about herself” as “default authoritative and default limitative.”<sup>44</sup>

The default view takes the practice of psychological ascription to provide a kind of ‘bedrock’ in accounting for asymmetries between avowals and other ascriptions. On this view, the presumed truth of avowals and the fact that they are not subject to ordinary epistemic assessment are to be regarded as simply *constitutive* of the concepts involved. The default status of avowals is part of what defines our concepts of mental states; it is not a product of non-conceptual facts about subjects’ mental life or their relationships to their mental states. Psychological ascriptions are all ultimately answerable to the subject’s own avowals *because this is how psychological concepts work*. This is the sense in which the default view presents the epistemic asymmetries between avowals and other ascriptions as a consequence of ‘grammar,’ rather than substantive epistemological differences. At the same time, however, the default view, unlike the simple avowal expressivism, can accommodate the semantic continuities between mental ascriptions to oneself and to others. For there is nothing in the default view as described so far to stand in the way of taking the surface grammar similarities between these two types of ascriptions at face value. For example, nothing seems to stand in the way of taking my self-ascription “I am in pain” to share truth-conditions with your ascription “She is in pain.”

By making the asymmetries out to be a matter of a conceptual constraint built into psychological discourse, the default view can offer a clear contrast between psychological and non-psychological ascriptions. As Wright intends it, the default view is an *anti-realist and non-factualist* view of psychological discourse, which presents our practices of psychological ascriptions, in contrast to our practices of non-psychological ascriptions, as ultimately “not accountable to any reality.”<sup>45</sup> However, we must recall that the epistemic asymmetries that give rise to our puzzle include ones that arise *within* psychological discourse; they pertain to differences between avowals and other *psychological* ascriptions—specifically, ascriptions of mental states to others, or to oneself at other times, and most notably ‘reportive’ present-tense “I”-ascriptions that are based on evidence, inference, etc. We still need to hear what distinguishes applications of psychological concepts in one’s *own* case, when made in the course of *avowing*. On the face of it, it doesn’t look as though an account that builds first-person authority into the truth-conditions of *all* psychological ascriptions is equipped to capture those

<sup>44</sup> Ibid, p. 41. The default view is also discussed in Wright 1987, pp. 401–402, 1989, sec. IV and 1991, sec. IV. Note that it is not obvious how the above constraint captures the *truth*-conditions of avowals, as opposed to describing our practice of *accepting* them as true. It is far from obvious how to work out the link between the fact that the linguistic practices surrounding avowals are governed by the constraint, on the one hand, with the conditions under which avowals are *true*. (Thanks to Dylan Sabo here.)

<sup>45</sup> Wittgenstein 1974, §133 as quoted in Wright 1998, p. 39.

asymmetries that arise *within* psychological discourse, and that seem sensitive not (or at least not only) to the *semantic content* (or truth-conditions) of avowals, but to the special way in which they are issued.

In offering the default view as an alternative to the Cartesian ‘privileged access’ view, Wright is partially motivated by an analogy to other areas of discourse where it seems misguided to conceive of successful judgments as tracking a completely independent reality—e.g. color judgments, and, on some views, ethical judgments. In these areas, Wright thinks it may be explanatory to suppose that there is *judgment-dependence*: what is true about an object’s color, for example, may not be entirely independent of the color-judgments of a well-placed perceiver.<sup>46</sup> In the case at hand—the mental or psychological realm—the claim would be that whether, for example, the avowal “I feel awful” is true or not, and whether or not the self-ascriber does feel awful, is not entirely independent of what she *thinks* about her state. Since this dependence on the subject’s own verdict constrains any ascription of a mental state to her, if someone said of me: “She feels awful,” the truth of her ascription would presumably also not be independent of what *I* take to be the case about my state. If so, it should be no surprise that there is such a fit between what a subject says (or thinks) about her condition and the truth of the matter, and ultimately no point in seeking a substantive explanation of the asymmetries between avowals and other ascriptions; hence, the default view.

However, the mentalistic case seems different from paradigm cases of judgment-dependence. In the former case, what calls for explanation are (partly) the *asymmetries* between a subset of mentalistic “I”-ascriptions (avowals) and other mentalistic ascriptions, including certain mentalistic “I”-ascriptions. Yet the notion of judgment-dependence by itself does not seem apt to capture these asymmetries. In paradigm cases (such as that of color), the idea of judgment-dependence is invoked to capture the intuition that a suitably well-placed judge can presumably not go wrong when making the relevant judgments. But in the case of avowals the intuition is not simply that *any* well-placed judge is authoritative regarding any arbitrarily chosen mental fact. Rather, the intuition is that subjects may be well-placed *only with respect to their own present mental states* and *only when avowing*, so that being well-placed will systematically depend on whether the relevant mental state is the subject’s or someone else’s. But then we may still be under explanatory pressure: *why* is it that subjects are to be systematically presumed *better* placed than their observers to pass ‘truth-determining’ judgments on certain facts in the mental realm, only some times, and only when avowing?

A proponent of the default view may object that the above complaint misses the point. The whole idea behind the view is to deny that there is any substance to the idea of first-person authority, anything ‘behind’ the special status assigned to avowals, beyond the conventions of mentalistic discourse. If so, the above complaint simply serves to expose how radical—and ultimately unsatisfactory—the default view really is.

<sup>46</sup> For discussion of this idea, see e.g. Wright, 1989: sec. 3 and 1992, Chapter 3, appendix.

To adopt the default view is in effect to deny that “the features of avowals . . . which seem to betray something remarkable about self-knowledge” reflect any deeper facts about the subject-matter of avowals, or subjects’ relationship to it (Wright 1998, p. 39). The view is that there is nothing *to* explain concerning the special security of avowals. We can note the security as marking a constitutive feature of psychological discourse; but it is futile to try to offer a substantive explanation of it. But then, as Wright himself remarks in criticism of the view, adopting the default view may be nothing more than “merely an unphilosophical turning of the back” on issues surrounding avowals and privileged self-knowledge.<sup>47</sup>

Like the view we labeled “psychological expressivism,” the default view attempts to solve our puzzle regarding the status of avowals by focusing on conceptual or ontological features that separate psychological discourse as such from ordinary descriptive discourses with unproblematic subject-matters. But while our puzzle does encompass contrasts between mental and bodily ascriptions, it is in the first place a puzzle that concerns the contrastive *epistemology* of avowals, and it also encompasses contrasts arising *within* psychological discourse and even within the class of mental “I”-ascriptions. For this reason, I have argued, neither psychological expressivism nor the default view is equipped to offer a satisfactory solution to it.

## 5. A Neo-Expressivist View of Avowals

Let us take stock. We have rejected the materialist introspectionist view on the grounds that, in its eagerness to put mental and non-mental ascriptions on an ontological par, it failed to respect those features of avowals that set them apart from certain nonmental present-tense self-ascriptions. Simple avowal expressivism addresses that problem by claiming that avowals, unlike all other ascriptions, are simply expressions of avowed mental states rather than descriptive reports about them. But the price it pays is that of compromising the semantic continuities between avowals and other ascriptions. Our study of psychological expressivism and the default view revealed that it is misguided to seek an account of the contrast between avowals and other ascriptions that is fashioned after anti-realist accounts in other areas.

Returning to our earlier separation of the positive *expressivist claim* from the negative *ontological claim*, as well as from the negative *semantic claim*, we can see in outline what shape an adequate solution to our puzzle may take. The positive expressivist claim seems to hold much promise for explaining the epistemic asymmetries between avowals and other ascriptions. But the temptation to put it to the service of an anti-realist agenda, by committing to the negative ontological claim, should be resisted, on pain of incoherence. And, to avoid compromising the semantic continuities, the

<sup>47</sup> Wright 1998, p. 41. For further discussion of the default view, see Boghossian 1989, McDowell 1998, Fricker 1998, Bilgrami 1998, and Moran 2001. The discussion here follows closely my discussion in Bar-On 2004, pp. 347–50, 410–12.

expressivist claim must be decoupled from the negative semantic claim.<sup>48</sup> My main task in this section is to flesh out this outline and present what I regard as a viable *neo-expressivist* account of the special security of avowals.

As a preliminary, let us return to the case of self-ascriptions issued through proprioception or kinesthesis. We should note that such self-ascriptions, which surely represent genuine truth-evaluable reports, nevertheless share a certain epistemic feature with avowals. Proprioceptive and kinesthetic self-reports are “identification-free”: epistemically speaking, they do not rest on a recognitional identification of the subject of the utterance or thought. That is to say, in normal circumstances, if I say (or think): “My legs are crossed” or “I’m spinning around,” my utterance/thought does not rest on my *recognizing* some individual as myself. I do not identify someone as being me and take that person’s legs to be crossed. I have no more reason for thinking that *someone’s* legs are crossed than whatever reason I have for thinking that *my* legs are crossed. This is what renders bodily self-reports of this kind “immune to error through misidentification” (to use terminology originally employed by Gareth Evans and Sidney Shoemaker).<sup>49</sup>

The following features of ascriptions that are immune to error through misidentification (IETM) are crucial for our present purposes. (i) The class of ascriptions that are IETM is heterogenic; it includes mental self-ascriptions (“I have a toothache”), perceptual self-reports (“I see a canary”), *as well as* a variety of bodily self-reports (“I am sitting down,” “I am bleeding,” “My legs are crossed”). (It also includes some demonstrative ascriptions, such as “That is moving fast.”) (ii) When a self-ascription of the form “I am F” is IETM, although I may fail to be F, I cannot have reason for thinking that *someone* is F that isn’t at the same time reason for thinking that *I* am F. And I cannot as much as doubt that it is *me* who has the relevant properties without doubting that *someone* has them. Thus, although my self-ascription may be in error, the error is not one of recognitionally *mis-taking* one individual for another; it is not an *error through misidentification*. (iii) Whether or not a self-ascription is IETM depends *not* on its semantic content, but rather on its “epistemic pedigree.” One and the same (semantically individuated) self-ascription can be IETM or not, depending on the *basis* on which it is made. For example, if I recognize that my legs are crossed by seeing someone in a mirror and taking her to be me (perhaps mistakenly), my self-ascription will *not* be IETM. (iv) In the case of avowals, though, I seem to enjoy additional, *ascriptive* security: not only are avowals immune to error of misidentification, but unlike all nonmental self-reports (as well as mental third-person ascriptions or past-tense self-ascriptions or evidential mental self-ascriptions), I submit that they are also *immune to error through misascription*. When avowing “I’d like some water,” for example, not only do I have no

<sup>48</sup> Wright, too, separates off the positive expressivist claim as one that a realist need not deny, see Wright 1988, p. 30.

<sup>49</sup> For discussion of the phenomenon of immunity to error through misidentification, see Wittgenstein 1958, 66–7, Evans 1982 esp. Chapter 7, sec. 2, Shoemaker 1968, and Wright 1998, pp. 18–20. I discuss the phenomenon and its relevance to explaining the security of avowals in Bar-On 2004, Chapters 4 and 6.

reason for thinking that *someone* wants some water, other than whatever reason I have for thinking it's *I* who wants water. I also have no reason for thinking that I want some substance that isn't also a reason for thinking that I want water. And I have no reason for thinking that I have some attitude or other toward water that isn't also a reason for thinking that I want water. As I self-ascribe a mental state in the normal, 'avowing' way (e.g. declaring "I have a terrible headache!," walking into a room and saying "I am so annoyed at this mess!," looking at the sky and saying "I think it's going to rain," etc.), I do not latch on to a present state of mine and *take* it to be M (a headache, or annoyance, or thought that p). I have no reason for thinking that I am in *some* state or other, over and above or separately from whatever reason I have for thinking I am in M.<sup>50</sup> If my avowal is false, this will not be due to a recognitional error of *mis-taking* one of my mental states for another. I *can* make such recognitional errors, however, if I were to self-ascribe M on the basis of some evidence, observation, interpretation, therapy, etc.

To say that avowals are immune to error through misascription is, then, not to say that they are absolutely infallible or incorrigible. It's just to say that they are protected from a certain array of epistemic errors (and thus corrections)—a *much wider* array than all other ascriptions, *including*, specifically, proprioceptive and kinesthetic self-reports, as well as evidential mental self-ascriptions. If I say or think: "My legs are crossed" (in the normal way), then even if I cannot be misidentifying who it is whose legs are crossed, if my legs are *not* crossed, my self-report will be mistaken, precisely because I have misidentified the state of my legs: I will have mistaken one state of my limbs for another. In the case of avowals, though, I suggest, I am protected from epistemic error in the ascriptive part as well. To reiterate, my avowal is not guaranteed to be correct and forever beyond correction. As we saw, in certain circumstances we may even be led to question an avowal of pain. But I would argue that these are not circumstances in which the avower is simply falsely ascribing to herself a state on the basis of a mistaken belief that she is in some *other* state. (After all, in the circumstances, she would be equally disposed to emit a spontaneous yelp. And, on pain of regress, it is implausible to suggest that the yelp would be the result of a false belief based on how her present state *appears* to her.)

The characterization of avowals' distinctive security as a matter of their unique immunity to error through misascription provides a suitably tempered interpretation of the claim that avowals are incorrigible and indubitable, an interpretation which does not require invoking either Cartesian privileged access or, indeed, any distinctively secure epistemic basis on which avowals supposedly rest. Though on this characterization avowals are not portrayed as absolutely incorrigible or indubitable, we can better understand the specific ways in which they are not open to epistemic challenge, doubt, and criticism, unlike other empirical ascriptions. The characterization also gives a more precise sense to the intuition that avowals are epistemically more immediate or direct

<sup>50</sup> If, as Wright says, the avowal is groundless, then I have *no* reason for thinking that I am in M. Alternatively, it may be suggested that I do have a reason—though it isn't a *justifying* reason: it is simply my *being* in M that gives me reason. (I consider this suggestion in Chapter 9 of Bar-On 2004.)

than other ascriptions. However, by itself, the characterization does not give us a full explanation of avowals' special security. For one thing, we need to understand the *source* of the additional immunity to error that avowals enjoy.<sup>51</sup> Why *is* it that avowals are not only immune to error through misidentification? Moreover, we have seen that immunity to error, in general, is no guarantee of *truth*. Yet avowals contrast with all other ascriptions in being very strongly presumed to be true.

It is at this point that I think we should co-opt the key insight from simple avowal expressivism: the idea that avowals' security is to be explained by appeal to their *expressive character*, rather than by appeal to the security of the epistemic basis on which they are made (be it introspection, or any other kind of epistemic basis). Like simple avowal expressivism, *avowal neo-expressivism* maintains that the epistemic simplicity and immediacy of avowals is best explained by seeing them as direct *expressions of subjects' self-ascribed mental states*. But in contrast to simple avowal expressivism, neo-expressivism maintains that avowals—unlike natural expressions, and like various mental reports—represent genuine, truth-evaluable ascriptions. In what follows, I briefly introduce certain key distinctions that would allow us to give more precise sense to the neo-expressivist claim that avowals are similar to natural expressions in being directly expressive of subjects' mental states, even though, semantically speaking, they are truth-evaluable self-ascriptions. I will then return to the question of the explanation of avowals' security.

The first distinction is one borrowed from Wilfred Sellars 1969, between two different senses (or kinds) 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 off and collecting oneself.)

*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  $[p]$ .<sup>52</sup>

<sup>51</sup> In the case of proprioceptive reports, the source of their immunity to error through misidentification has to do with our possessing special mechanisms for obtaining information concerning our *own* bodies. See Evans 1982, Chapter 7.

<sup>52</sup> Sellars also speaks of expression in the *causal* sense, where 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 shaking hands may express in the causal sense one's pain or nervousness, respectively.) For present purposes, we can set aside 'c-expression,' though it plays a role in the full neo-expressivist account. See Bar-On 2004, pp. 216f., 248f., and *passim*.

As we consider the respects in which avowals may resemble natural expressions, we should be focusing on natural expressions that fall under a-expression—where the relevant mental state is the rational cause of, or reason for, one’s expressive behavior. Furthermore, as we consider these cases, we must distinguish between the *act* of expressing and its *product*. Like other nominals in English (and other languages), such as “statement,” “assertion,” “ascription,” “report,” as well as, indeed, “expression,” etc., “avowal” admits of two different readings. It can be read as referring to someone’s act of avowing, which is an event in the world with a certain causal history and certain action-properties. But it can also be read as referring to the result or product of such an act—a linguistic (or language-like) token, an item with certain semantic properties. The important similarities between avowals and other intentionally produced expressions (intentionally produced natural expressions included) 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 such as a facial expression, a gesture, a bodily movement, a demeanor, an inarticulate sound, when understood as products, and however produced—do not on their face 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, or a cowering demeanor stands for feeling threatened.<sup>53</sup> Thus there are notable differences between avowals and acts of natural expression in terms of their products. But I think that the expressivist insight regarding avowals should be understood, in the first instance, as a claim about the relevant *acts*, not about their products. The claim is that there are notable similarities between the act of avowing a state and the act of giving it a (so-called) natural expression; e.g. between the act of saying: “I really like this!” (or, alternatively, “This is wonderful!” or “Yippee!”), and the act of giving an ear-to-ear smile, or a firm handshake. And this claim could be true, even if there were systematic differences between the products of acts of avowing and the products of naturally expressive acts.

<sup>53</sup> Although, of course, there are gestures and other non-verbal expressions that are governed by socio-cultural conventions. The conventions do not, however, assign semantic content to the gesture, but rather set up a (pragmatic) connection between the making of the gesture (e.g. tipping one’s hat) and being in the relevant state, or having the relevant attitude, sentiment, etc. (e.g. feeling respect).

Also, a case can be made that, e.g., an animal’s alarm call has semantic content—representing, e.g., *threat from above* or whatever. Still, unlike an avowal, it cannot be plausibly taken to express in the semantic sense a proposition *about* the relevant mental state of the animal issuing the call.

Thus consider a paradigm case of expressive behavior: a small child, Jenny, eagerly reaching for a teddy bear. Jenny simply wants the toy, and her reaching for it gives direct expression to her desire, quite independently of any judgment on her part to the effect that she wants the teddy. Jenny wants the teddy and her (non-verbal) behavior shows it; she does nothing to hide her desire for teddy. And in normal circumstances, her audience will be directly responsive to the state they perceive her to be in; barring contravening reasons, they will simply hand her the toy. But now consider another episode in which Jenny emits a certain sound (“Uh!”), or calls out: “Teddy!” as she reaches for the toy. And finally, consider an episode in which she avows: “I want Teddy,” perhaps with no reaching at all. Intuitively, the verbal emissions, just like the reaching behavior, which we would consider a natural expression of Jenny’s desire, ‘come directly from’ the child’s desire. They seem equally ‘pressed out’ from her, and they appear no more driven by a prior deliberation, consideration, or determination regarding how things are presently with her. The verbal utterances seem to be equally expressive of her desire for the teddy. Granted: when saying “I want Teddy!” Jenny makes intentional use of conventional verbal means that are acquired. If her saying “I want the Teddy” itself constitutes expressing her desire, the expression will not be a natural one (at least in one sense of the term). But notice that the same would hold of her saying “Daddy!” to express her happiness at seeing Dad, or her saying “Ouch!” to express her pain. Very early on in our lives, we begin to use acquired, conventional means—both words and gestures—to give expression to our states.

To vary the example, consider a linguistic utterance such as “It’s so good to see you!” It typically serves to express the speaker’s joy. But, it seems, one can equally express her joy by avowing “*I am so glad to see you!*” Armed with the act/product distinction, as well as the distinction between a- and s- expression we can capture the similarities and differences between the two utterances. They can both be seen as instances of the same type of act, and they both a-express the same type of mental state. But the products of these two acts, though both linguistic, are (semantically) different. The first sentence token s-expresses the proposition that it is very good to see one’s hearer, in virtue of the rules of English. The second sentence token is self-ascriptive: it s-expresses the proposition or thought that the speaker is very glad to see her hearer. By contrast, the product of an act of giving your friend a cheerful hug is not governed by any syntactic or semantic rules of English, and it does not s-express anything. Still, it can equally be seen as an act of the same type, which a-expresses your joy at seeing your friend.

What this illustrates is that our ordinary everyday notion of expression applies readily to non-verbal and verbal expressions alike, and to verbal expressions of a wide variety. The commonsense notion is that of an act, whether linguistic or nonlinguistic, whose point is none other than either to make public, put forth, share, or air, or else just to give vent or voice, rather than to offer a descriptive report, make an assertion, or inform someone about the speaker’s present thoughts, feelings, emotions, or attitudes. The notion seems to apply even when the expressive vehicle used is not ‘natural’ but is rather introduced in some conventional way and is acquired. It is also applicable to acts

performed in the absence of audience, as when someone spits at a photograph or caresses an object while alone, as well as to mental acts such as saying to oneself: “Oh my god!” or “That bastard” or “What a mess!”

On avowal neo-expressivism, then, avowals emerge as a certain class of expressive *acts*; acts in which a subject gives linguistic vent to present mental states. These are acts in which subjects *speak their minds*, instead of giving non-linguistic expression to it, and the linguistic vehicles they use are *self-ascriptive*. Acts of avowing are epistemically unmediated; but for all that, they have products (sentence tokens, in the case of speech, articulate thought tokens, in speechless cases) that can be true or false. And we have seen that this is not without precedent. When I exclaim “This dog is scary!” I may be giving voice to my present fear, using a linguistic vehicle—a sentence token—that is true or false, depending on whether the designated dog is scary or not. Similarly, when I avow “I am scared of this dog,” I may be giving voice to my fear, using a sentence token that is semantically about me and a state of mine, and is true or false depending on whether I am scared of the designated dog or not.<sup>54</sup>

We are now in a position to return to one of our earlier questions. We have noted that, even though what avowals s-express are not a priori, or necessary propositions, and although when avowing subjects make epistemically unmediated, ‘groundless’ claims, avowals are quite rarely challenged or corrected, and not because they cannot be negated (like moans or verbal expletives), it is not simply that avowals go unchallenged; they are also strongly presumed to tell us *true* things about their subjects. On the neo-expressivist account, the reason is this. When I avow, I a-express the *very same state* whose presence *makes true* the proposition s-expressed by the avowal (understood as product). To take me to be avowing, then, is to take it that I *am* in the designated state, which is to take the avowal to be *true*.<sup>55</sup> (We have already touched on the question how the expressivist idea helps explain the various epistemic asymmetries.)

Our puzzle regarding avowals was this:

Why is it that avowals, understood as true or false ascriptions of contingent states to an individual, are so rarely questioned or corrected, are generally so resistant to ordinary epistemic assessments, and are so strongly presumed to be true?

This puzzle invites a search for a distinctive feature(s) of avowals that would explain why they enjoy the level and kind of security that we ordinarily assign to them. If we are not to turn our back on the puzzle, as the default view arguably does, yet at the

<sup>54</sup> The observation that is crucial for the neo-expressivist account is that, in point of its potential to express the subject’s (self-ascribed) present state of mind, a self-ascription such as: “I am so annoyed at you” (produced in speech or in thought) is no different from other verbally expressive acts, e.g.: “This is very annoying,” or “Enough!” or even “Ugh!” In Bar-On 2004, Chapter 7, I argue that it is *further* plausible that these latter expressions can be expressive of a subject’s mental state in the same sense as intentional acts of natural expression, such as letting out a sound or making a face or a gesture that are not conventional. But the further claim is not strictly part of the core neo-expressivist thesis.

<sup>55</sup> The complementary question that then arises for the neo-expressivist is how to understand the possibility of *false* avowals. I discuss this issue at length in sections of Chapters 7 and 8 of Bar-On 2004.

same time improve on the materialist introspectionist view of first-person authority, we must show why it should be *reasonable* for us to treat avowals as protected from certain *kinds* of epistemic mistakes and criticisms, even if not absolutely infallible and incorrigible, and to assign avowing subjects inalienable and non-transferable first-person authority. On the neo-expressivist story, it is reasonable for us to assign inalienable and non-transferable first-person authority to subjects, because this authority is assigned on the strength of recognizing their avowals as acts in which they directly express *the very states they self-ascribe* through the products of those acts at the same time as they self-ascribe them.

## 6. Expressivism, Truth, and Realism

As mentioned earlier, Wright does not think avowal expressivism is ruled out of court by the Geach point, but he has other worries about the expressivist proposal. I believe avowal neo-expressivism is equipped to address the remaining worries Wright raises for avowal expressivism.<sup>56</sup> But my concern here is rather with Wright's reasons for taking expressivism in general to represent a 'faux pas.' Deploying the distinctions that (I believe) allow avowal expressivism to rise from the dead, might we not be able to articulate a *neo-expressivist* paradigm that is properly expressivist but neither requires denying truth-evaluability nor is driven by an antecedent anti-realist agenda?

First, let us note that avowal neo-expressivism, like simple avowal expressivism *properly construed*,<sup>57</sup> is neither motivated by nor committed to psychological anti-realism. The view is strictly designed to explain the *epistemic* asymmetries between avowals and other ascriptions. *Like* simple avowal expressivism, the neo-expressivist view does not explain avowals' distinctive features by appeal to their secure epistemic basis, but instead appeals to their expressive character. However, unlike the simple view, it does not take the security to be a matter of the *semantics* or 'grammar' of avowals understood as products; for avowals *qua* products are said to share truth-conditions with other ascriptions (e.g. "I am tired" avowed by me is true iff DB is tired). The neo-expressivist positive claim is confined to what avowals a-express and does not concern what avowals (or rather their products) s-express. It does not get coupled with a negative semantic claim (a denial of truth-conditions and thus of truth-evaluability).

But suppose we now ask: What are the 'truth-makers' of avowals—what renders an avowal true or false? The answer, as far as neo-expressivism is concerned, is straightforward: facts about the mental lives of subjects. There is no presupposition that such facts are problematic, nor even that there is any systematic ontological contrast between psychological and other facts. Moreover, facts about the mental life of any subject can in principle be reported by anyone, *including* the subject herself. It's just that when

<sup>56</sup> See footnote 28.

<sup>57</sup> See above, Section 3.

*avowing*, subjects do not set out to issue reports about their states of mind. Suppose I avow: “I am tired” (instead of just stretching and giving a big yawn), or “I’m feeling upset about my brother’s not showing up to my party” (where presumably I *couldn’t* engage in any non-verbal behavior expressive of the relevant state, though I *could* express my state by saying: “It’s very upsetting that my brother didn’t show up to my party”). I can turn around and make a past-tense descriptive report made true by the same state of affairs: “I was tired then,” or “I felt upset about . . .”. Indeed, at the time I issued an avowal, I might have issued a present-tense evidential self-report instead. Had circumstances been different, I might have said “I am tired” on the basis of seeing my reflection in the mirror, or I might have concluded that I am feeling upset about my brother’s absence on the basis of a therapy session, discussion with a friend, etc. The neo-expressivist explanation of avowals’ distinctive security focuses directly on epistemological and action-theoretic differences between avowals and their corresponding self-reports, rather than on any alleged ontological differences between the mental and other realms.

Unlike the answer given by wholesale psychological expressivism to the (essentially epistemological) puzzle about avowals’ security, the neo-expressivist answer avoids ‘teetering into incoherence,’ since it does not presuppose that there are no states for *any* psychological sentences to describe. Moreover, unlike the default view, it is not committed to the strong dependence of psychological facts on the judgments of the subjects who harbor psychological states. On the default view, *being* in a mental state is partially constituted by thinking that one is in the mental state, so that facts in the mental realm are determined in part by what self-ascribers take them to be. But there are good reasons to avoid tying the presence of all mental states to the presence of thoughts *about* them. For one thing, ‘basic’ mental states, such as hunger, thirst, pain, etc. are ones that we seem able to share with non-human animals and pre-cognitive children, who are presumably not equipped to entertain thoughts about their own mental states. And when it comes to sophisticated psychological states, such as an occurrent desire to have one’s political rival be embarrassed on a given occasion, it seems quite possible for someone to *have* that desire without realizing it. Moreover, it seems perfectly possible for one to form the judgment that one is in some psychological state even though one is not. Passing judgment on a contingent matter of fact is a risky business, in the case of mental states as elsewhere. However reliable one is, one can get things wrong. There are all sorts of ways of coming to make a judgment that one feels depressed, or hopes for something, or wants something, or even is in pain, and they can lead us astray. There may well be particular mental phenomena that do fit Wright’s characterization of the judgment-dependence of mental ascriptions, where one’s judgment that one is in the mental state actually makes it the case that she is, or at least partially constitutes her being in the state.<sup>58</sup> But I do not believe this to be the

<sup>58</sup> See Taylor 1985, pp. 100–1, Moran 1988, and Tanney 1996.

general case. In general, I think we should allow that judging that one is in some mental state does not make it so.

All this tells against *conceptually* tying the very existence or presence of mental states to subjects' judgments. And nothing in the neo-expressivist view of avowals requires such a conceptual link. Not only does the view allow for the possibility of being in a mental state without forming any judgment about it, and of forming false judgments about one's present mental states on this or that basis, but the view also allows for the possibility of false avowals. Unlike views that explain avowals' security by appeal to the 'logical grammar' of avowals in particular or of mental discourse more generally, neo-expressivism does not associate with avowals any conceptual guarantee of truth. It also does not require accepting that mental states in general are of subjects' own making.<sup>59</sup>

In the particular case of avowals, there may be, then, good reasons for the neo-expressivist to avoid strongly anti-realist views of mentality, without yet committing to any particular view of the nature of mental states. Taking the lead from the neo-expressivist view of avowals, we may generalize, and consider the following combination:

- (i) Endorsing a positive expressivist claim concerning a given discourse D, i.e. maintaining that utterances/thoughts that involve the 'signature' terms or concepts of D characteristically (and perhaps essentially) serve to a-express certain mental states of the speakers/thinkers.
- (ii) Rejecting the negative semantic claim, i.e. maintaining that utterances/thoughts in D are truth-evaluable in that they s-express propositions that can be true or false.
- (iii) Rejecting the negative ontological claim, i.e. refusing to commit to there being *no* D-facts of the matter, and no D-objects, properties, or states to constitute a proper subject-matter for discourse D.

In order for this combination to present a viable option, the expressivist claim has to be put forth for some particular explanatory purpose *other than* the traditional expressivist purpose of explaining how utterances of a given discourse can have a point even though there are no facts for them to be about. It would be particularly instructive, I think, to sketch such a neo-expressivist option for the case we originally contrasted with the case of avowals: that of ethical discourse.<sup>60</sup>

According to a view we may dub *ethical neo-expressivism*, traditional ethical expressivism was correct to identify an important, perhaps essential, expressive function of ethical discourse and reflection: ethical claims (whether made in speech or in thought) serve to express certain of our conative attitudes. Armed with the distinctions introduced when presenting avowal neo-expressivism, we can retain this key idea without

<sup>59</sup> For further discussion, see Bar-On 2004, Chapters 8 and 10.

<sup>60</sup> The application of the neo-expressivist framework to ethical discourse is the topic of Bar-On and Chrisman 2009.

compromising the palpable semantic continuities between ethical and non-ethical claims. As in the case of avowals, we can distinguish between the act of making an ethical claim and the product of this act. An ethical claim considered as a product—i.e. a sentence- (or thought-) token that essentially employs ethical terms (or concepts)—can be said to *s*-express a true or false proposition. However, as with avowals, what is *s*-expressed by a claim does not settle what mental state is characteristically *a*-expressed by acts of making the claim. With Hume, we may think that purely cognitive states (such as beliefs or judgments) cannot by themselves motivate or explain action, and furthermore, we may think there is a non-accidental connection between sincerely making an ethical claim and being motivated to act in its accord. The expressivist claim tells us that ethical utterances (in my terminology) *a*-express the very same states whose presence is required for understanding the perceived motivational force of such utterances.<sup>61</sup> As Ayer already saw, this expressivist insight is best captured without supposing that (again, in my terminology) the products of ethical claims *s*-express propositions that self-ascribe those states.<sup>62</sup> Thus, as long as we're talking about expressive acts, we can agree with Ayer that ethical claims betray conative attitudes, not because they report them, but because they (*a*-) express them directly. What the *neo*-expressivist goes on to add is that we need not go on to accept Ayer's claim that ethical claims (understood as products) are not truth-evaluable because they do not (*s*-) express *any* true or false propositions.

We can present ethical *neo*-expressivism as a combination of the following three claims:

- (i) Ethical claims understood as *acts* are different from ordinary descriptive claims in that agents making them (in speech or in thought) essentially *a*-express conative attitudes.
- (ii) Ethical claims understood as *products* are semantically continuous with ordinary descriptive claims in being truth-evaluable, embeddable in conditionals, negation, logical inferences, etc. This is because ethical claims *s*-express true or false propositions.
- (iii) It is a matter for further metaphysical investigation to determine what (if anything) *makes* ethical propositions true or false: whether there are ethical properties or states of affairs, and if there are, what their nature is and to what extent they are mind- (or judgment-) dependent remains to be told.<sup>63</sup>

<sup>61</sup> They do so, *whether or not* they *also* express a belief whose content is given by the *s*-expressed proposition. For some discussion, see Bar-On and Chrisman 2009.

<sup>62</sup> Ayer 1946, pp. 104–8.

<sup>63</sup> As noted above, in the ethical case, unlike in the avowal case, we are not courting incoherence by leaving open the possibility that there are no 'truth-makers' for the relevant propositions. One may wonder, however, what we are to make of the proposition allegedly *s*-expressed by an ethical claim, if we are to leave open the possibility that there are no ethical properties or states of affairs. For our purposes here, we can specify this proposition disquotationally. So, for instance, we'll say that the claim "Tolerance is a virtue" *s*-expresses the proposition that tolerance is a virtue. (To say this is not to commit to a disquotational theory of the *truth* of ethical (or other) claims. It is just to avoid committing to the possibility of any paraphrastic semantic analysis thereof.)

Ethical expressivists, it seems, should welcome the possibility of decoupling the key expressivist insight from unwelcome semantic commitments. But why should anyone be drawn to expressivism about ethical discourse once it is divorced from the denial of ethical facts? The answer already suggested is that the expressivist claim may be explanatory of certain puzzling features of ethical discourse. Quite independently of any ontological qualms about ethical facts or properties, we may find it illuminating to recognize that ethical claims *a*-express conative states, since that may explain the apparently non-accidental connection between ethical claims and motivation (at least within a Humean framework). On the ethical neo-expressivist proposal, to take someone as making a genuine ethical claim is *ipso facto* to regard her as giving voice to a state whose presence explains why she is suitably motivated when making the claim.

Finally, I think Wright should welcome the introduction of a new expressivist paradigm, since it may help capture several insights he himself has offered in his extensive discussions of expressivism, realism, and truth.<sup>64</sup> In particular, the new paradigm seems very much in the spirit of Wright's minimalist proposal, which allows us to preserve truth-evaluability and truth talk wherever we find the right assertoric trappings, while leaving it open whether we should come down on the side of realism or anti-realism. What the new paradigm adds to the mix is the positive expressivist claim; a claim that, I hope to have shown, we often have reason to invoke.<sup>65</sup>

## Bibliography

- Austin, J. L. 1961 "Performative Utterances," in J. O. Urmson and G. J. Warnock (eds.) *Philosophical Papers*, Oxford, Clarendon Press, pp. 233–52.
- Ayer, A. J. 1946 *Language, Truth, and Logic*, 2nd edn., London, Gollancz.
- Bar-On, D. 2004 *Speaking My Mind: Expression and Self-Knowledge*, Oxford, Clarendon Press.
- Bar-On, D. 2012 "Externalism and Skepticism: Recognition, Expression, and Self-Knowledge," in A. Coliva (ed.) *Self-Knowledge and the Self*, Oxford, Oxford University Press, pp. 189–211.
- Bar-On, D. and Chrisman, M. 2009 "Ethical Neo-Expressivism," in R. Shafer-Landau (ed.) *Oxford Studies in Metaethics*, Volume IV, pp. 133–66.
- Bilgrami, A. 1998 "Resentment and Self-Knowledge," in C. Wright, B. C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*, Oxford, Oxford University Press, pp. 207–41.
- Blackburn, S. 1984 *Spreading the Word*, Oxford, Oxford University Press.
- Blackburn, S. 1993 *Essays in Quasi-Realism*, Oxford, Oxford University Press.
- Blackburn, S. 1998 "Wittgenstein, Wright, Rorty and Minimalism," *Mind* 107, pp. 157–81.
- Boghossian, P. 1989 "Content and Self-Knowledge," *Philosophical Topics* 17, pp. 5–26.
- Davidson, D. 1973 "Radical Interpretation," *Dialectica* 27, pp. 313–28.
- Davidson, D. 1984 "First Person Authority," *Dialectica* 38, pp. 101–11.
- Dennett, D. 1987 *The Intentional Stance*. Cambridge, Mass., MIT Press.

<sup>64</sup> See Wright 1992 and the essays collected in Wright 2001 and Wright 2003.

<sup>65</sup> I wish to thank Matthew Chrisman, Peter Hanks, and Ram Neta for helpful comments on previous drafts. Many thanks to Jussi Suikkanen for both his comments and his bibliographical help.

- Evans, G. 1982 *The Varieties of Reference*, Oxford, Clarendon Press.
- Fricker, E. 1998 "Self-Knowledge: Special Access vs. Artefact of Grammar: A Dichotomy Rejected," in C. Wright, B. C. Smith, and C. Macdonald (eds.) *Knowing Our Own Minds*, Oxford, Oxford University Press, pp. 155–206.
- Geach, P. 1960 "Ascriptivism," *The Philosophical Review* 69, pp. 221–5.
- Gibbard, A. 1990 *Wise Choices, Apt Feelings*. Oxford, Clarendon Press.
- Gibbard, A. 2003 *Thinking How to Live* Cambridge, Mass., Harvard University Press.
- Hacker, P. 1993 *Wittgenstein: Meaning and Mind*, Oxford, Blackwell.
- McDowell, J. 1998 "Response to Crispin Wright," in C. Wright, B. C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*, Oxford, Oxford University Press, pp. 47–62.
- Mackie, J. 1977 *Ethics: Inventing Right and Wrong*, London, Penguin.
- Moran, R. A. 1988 "Making Up Your Mind: Self-Interpretation and Self-Constitution," *Ratio* NS 1, pp. 135–51.
- Moran, R. A. 2001 *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton, NJ, Princeton University Press.
- Ridge, M. 2006 "Ecumenical Expressivism: Finessing Frege," *Ethics* 116, pp. 302–36.
- Rorty, R. 1970 "Incorrigibility as the Mark of the Mental," *Journal of Philosophy* 67, pp. 399–424.
- Sellars, W. 1969 "Language as Thought and as Communication," *Philosophy and Phenomenological Research* 11, pp. 506–27.
- Shoemaker, S. 1968 "Self-Reference and Self-Awareness," *Journal of Philosophy* 65, pp. 555–67.
- Stevenson, C. 1944 *Ethics and Language*. New Haven, Conn., Yale University Press.
- Tanney, J. 1996 "A Constructivist Picture of Self-Knowledge," *Philosophy* 71, pp. 405–22.
- Taylor, C. 1985 "The Concept of a Person," in *Human Agency and Language: Philosophical Papers*, Cambridge, Cambridge University Press, pp. 97–114.
- Wittgenstein, L. 1953 *Philosophical Investigations*, ed. G. E. M. Anscombe and R. Rhees, trans. G. E. M. Anscombe, Oxford, Blackwell.
- Wittgenstein, L. 1958 *The Blue and Brown Book*, New York, Harper & Row.
- Wittgenstein, L. 1974 *Philosophical Grammar*, Oxford: Blackwell.
- Wright, C. 1987 "On Making Up One's Mind: Wittgenstein on Intention," in P. Weingartner and G. Schurz (eds.) *Logic, Philosophy of Science and Epistemology, Proceedings of the XIth International Wittgenstein Symposium*, Vienna, Holder-Pickler-Tempsky, pp. 391–404.
- Wright, C. 1988 "Realism, Antirealism, Irrealism, Quasi-Realism," in H. Wettstein, T. Uehling and P. French (eds.) *Realism and Anti-Realism, Midwest Studies in Philosophy*, vol. XII, Minneapolis, University of Minnesota Press, pp. 25–50.
- Wright, C. 1989 "Wittgenstein's Rule-Following Considerations and the Central Project of Theoretical Linguistics," in A. George (ed.) *Reflections on Chomsky*, Oxford, Blackwell, pp. 233–64.
- Wright, C. 1991 "Wittgenstein's Later Philosophy of Mind: Sensation, Privacy, and Intention," in K. Puhl (ed.) *Meaning Scepticism*, Berlin, de Gruyter, pp. 126–47.
- Wright, C. 1992 *Truth and Objectivity*. Cambridge, Mass., Harvard University Press.
- Wright, C. 1995a "Truth in Ethics," *Ratio* 8, pp. 209–26.
- Wright, C. 1995b "Can There Be a Rationally Compelling Argument for Anti-Realism about Ordinary ('Folk') Psychology?," *Philosophical Issues* 6, pp. 197–221.
- Wright, C. 1996 "Précis of *Truth and Objectivity*," *Philosophy and Phenomenological Research* 56, pp. 863–8.

- Wright, C. 1998 “Self-Knowledge: The Wittgensteinian Legacy,” in C. Wright, B. C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds*, Oxford, Oxford University Press, pp. 13–46.
- Wright, C. 2001 *Rails to Infinity: Essays on Themes from Wittgenstein’s Philosophical Investigations*. Cambridge, Mass., Harvard University Press.
- Wright, C. 2002 “What Could Anti-Realism about Ordinary Psychology Possibly Be?,” *The Philosophical Review* 111, pp. 205–33.
- Wright, C. 2003 *Saving the Differences: Essays on Themes from Truth and Objectivity*, Cambridge, Mass., Harvard University Press.