Neo-Expressivism: (Self-)Knowledge, Meaning, and Truth

Dorit Bar-On, University of Connecticut

1. 'Saving the Differences'

In the Preface to his collection of essays, *Saving the Differences*, Crispin Wright introduces the Wittgensteinian concern with

...differences in the role and function of superficially similar language games ... which those very similarities encourage us to overlook, thereby constituting a prime cause of philosophical misunderstandings and confusions.¹

One instance of this concern is 'the idea that statement-making and susceptibility to ordinary propositional logic may be a relatively superficial aspect of discourses, masking differences in point of objectivity and the manner in which they relate to the real world.' Wright immediately goes on to say:

That idea calls for a philosophy of truth and truth-aptitude which allows the application of those notions to tolerate such deep differences, and an account of wherein the deep differences themselves consist.²

In other words, attempts at 'saving the differences' face a dual challenge:

a. to explain *contrasts* between discourses of interest – ethics, mathematics, theoretical sciences, mentalistic discourse – on the one hand, and discourses deemed unproblematic, on the other, while at the same time trying

b. to accommodate the 'relatively superficial' logico-semantic *similarities* between the target discourses and the unproblematic ones.

¹ Crispin Wright, Saving the Differences (Cambridge, MA: Harvard University Press, 2003), vii.

² ibid.

This can prove challenging, because certain strategies for accomplishing the first task make accomplishing the second task more difficult, and vice versa. Just to give the flavor: *emotivism* may provide a way of capturing a rather sharp contrast between ethical and other claims by portraying ethical claims as non- truth-apt emotive expressions (like emitting a 'Boo' or 'Hurray'). But in so doing it makes it impossible to accommodate discourse continuities between ethical and non-ethical claims. And *deflationism* allows us to preserve truth evaluability for ethical (or other potentially problematic) claims, since it maintains that there is nothing to thinking or claiming that it is *true that* p over and above thinking or claiming that p. But in so doing it may make it difficult to insist that, though ethical claims can be true, can state facts, can be involved in logical inferences, and so on, nevertheless, ethical discourse contrasts with other discourses in being non-objective or meriting an anti-realist construal.

Furthermore, certain strategies for addressing a. (i.e., for explaining deep differences) risk being self-undermining, because they globalize. (Thus, for one recent historical example, recall Dummett's well-known attempt to provide an anti-realist construal of mathematics by giving a pair of arguments to the conclusion that our *understanding* of mathematical statements could not consist in apprehending verification-transcendent truth-conditions. These were the Manifestation and the Acquisition arguments.³ It did not take long to realize that, if these arguments were right, they would generalize: *no* statement that we understand could have verification-transcendent truth-conditions. Global anti-realism followed. This meant that, if we were still to insist on saving the differences, we would have to look elsewhere: the idea of non- verification-transcendent truth-conditions – later replaced in Dummett's work by *assertibility* conditions – cannot be used on its own to draw the requisite contrasts.)

^a Michael Dummett, Truth and Other Enigmas (Cambridge, MA: Harvard University Press, 1980).

Things get even more complicated, when it is appreciated that different allegedly problematic discourses exhibit *different* contrasts with discourses deemed unproblematic. So, for example, mathematics (and perhaps theoretical science too) primarily raise an *ontological* challenge: How to place mathematical (or theoretical) entities - if such there be - in the observable world as we ordinarily know it. (Ontological problems can of course bring in their train, or be motivated by, epistemological concerns. How could we come to have *knowledge* of abstract or theoretical entities?⁴) Ethical discourse can also raise ontological challenges, but there is at least one type of challenge it raises that faces even those who are prepared to set aside ontological qualms about 'queer' ethical properties or facts. This is a *moral-psychological* challenge: how to explain the apparently essential connection between ethical judgments and motivation to act. Ethical beliefs appear different from ordinary beliefs in that they cannot be motivationally inert. And this poses a challenge for *both* those who are committed to naturalism and those who are not. Turning to mentalistic discourse, those who wish to resist Cartesian dualism face the challenge of explaining how, if mental states simply constitute a sub-class of our bodily states, we could possibly have privileged 'first-person' knowledge of our own present mental states.

My own philosophical prejudice is that attempts that I would describe as *semantic* – that is, attempts to save the differences in terms of either meanings or truth – either globalize or else fail to preserve logical-semantic continuities. I cannot here argue for this general claim. Instead, I indicate some general features of what I take to be the desired alternative. The alternative would:

(i) preserve logico-semantic continuities without committing to deflationism

⁴ Cf. Paul Benacerraf, 'Mathematical truth', Journal of Philosophy, 70 (1973), 661 – 679.

(ii) avoid tying the differences to differences in the 'types of propositions' expressed in different discourses.

(iii) avoid postulating a plurality of truth properties, each operating in a different range of discourses

(iv) explain differences on a case-by-case basis, to avoid globalizing

(v) embrace the possibility of metaphysical or epistemological pluralities (depending on the contrasts to be explained)

In the next Section (2), I consider the case of mentalistic discourse as a kind of case study. I review some of the philosophical puzzles that have led philosophers to think of mentalistic discourse as problematic. These puzzles concern the semantic, epistemological, and metaphysical status of contrasts between first-person present-tense attributions ('avowals') and all other ordinary contingent attributions, which contrasts appear to be constitutive of distinctively *mentalistic* discourse. In Section 3, I briefly present my own, neo-expressivist strategy for addressing the puzzles. Neo-expressivism is a descendant of 'simple' expressivism - which is, in turn, the analogue in the mentalistic realm of ethical emotivism. But neo-expressivism differs from simple expressivism precisely in that it is not proposed as a view of either the (non-) truth-aptness of avowals or about their meanings. Crucially, it accommodates both continuities and discontinuities between mentalistic and other discourses, and yet it does not globalize. At the same time, it is possible to apply the neo-expressivist framework in other areas where the notion of expression is deemed explanatorily useful. I will illustrate this - in Section 4 - by recalling neo-expressivism about *ethical* discourse. In the final Section 5, I make some comments about truth and meaning and tease out some of the commitments of the approach I advocate.

2. Mentalistic Discourse - Three Dilemmas

4

Following Descartes, discussions of the nature of mind often begin with some commonsense observations about striking features of 'the first-person', as contrasted with 'the third-person' (which, for relevant purposes, also includes what we say or think about our own bodies). Such observations lie behind three representative dilemmas concerning mentalistic discourse – one semantic, one more metaphysical, and one epistemological.

2.1 Three Dilemmas

First, thinking in semantic terms, Elizabeth Anscombe (following Wittgenstein) observed that the first-person pronoun '**I**', at least in certain of its uses, enjoys a peculiar guarantee of semantic success.⁵ When using '**I**' as subject' – for example, when using it in an avowal, such as '**I** am in pain', or '**I**'m feeling sad' – one cannot fail to pick out *some* object, and, moreover, one cannot fail to pick out the *right* object (that is, oneself). So, suppose – to embellish on one of Anscombe's thought-experiments – you are hurtling through space in a sensory deprivation tank, afflicted by annesia as well as paralysis, and you think to yourself '**I** really hate this!'. You have no means of recognizing yourself as the object to which '**I**' refers on this occasion (you cannot track yourself through descriptions stored in memory, proprioception, or location relative to other objects). Still, your thought cannot fail to be about *yourself*. But how can that be?

Anscombe argues that, if we suppose that 'I' is indeed a referring expression, then we have to accept that it could not refer to an 'ordinary' object (such as a human body), given the guaranteed referential success noted above. But then it must refer to an *extraordinary* object: an object securely knowable the way no ordinary object can be. (The Cartesian Self or Ego can then

⁵ Elizabeth Anscombe, 'The First Person', in Samuel D. Guttenplan (ed.), *Mind and Language*. (Oxford: Oxford University Press, 1975), 45 – 65.

be seen as a postulated referent for 'I' that possesses the requisite features.) *Anscombe's Dilemma* says:

Either 'I' refers to a peculiar object (=Cartesian Ego) or 'I' does not refer at all. Anscombe goes on to argue that Cartesian Egos would in fact make very poor candidates for objective reference, and so she opts for the second horn, thereby adopting a deflationist No Reference thesis with respect to 'I'.

Surely, however, the No Reference thesis flies in the face of palpable continuities between uses of "I" and uses of other referential terms. And, given that Anscombe wants to restrict the No Reference claim to just statements involving uses of 'I' as subject, her account would face serious difficulties accounting for 'mixed' uses of 'I' – e.g. 'I got very little sleep last night, so I am feeling very tired right now'. What is worse, it doesn't really address the very puzzle that Anscombe uses to motivate the thesis. (That puzzle was motivated by the observation that certain uses of 'I' seem to enjoy peculiar referential security. However, as Evans and others later showed,⁶ uses of 'I' 'as subject', are not restricted to self-attributions of specifically *mental* states. So, if we accept, with Anscombe, that uses of 'I' as subject are *all* non-referential, we are left with the original Cartesian puzzle: what explains the *special* security of 'I' in picking out oneself when thinking, specifically, about one's own mental states?)

Anscombe's strategy illustrates what I would consider an overkill solution, one that attempts to capture certain striking differences (in this case, semantic ones) by sacrificing logicosemantic continuities. It also illustrates the heavy cost that can be incurred by semantic solutions to metaphysical or epistemological puzzles.

⁶ Gareth Evans, Varieties of Reference (Oxford: Oxford University Press, 1982).

Next, thinking more directly in metaphysical terms, about the *nature* of mental states, there is a second dilemma, articulated by Richard Rorty,⁷ who thought that observations related to those behind Anscombe's dilemma strongly constrain our metaphysical view of the mental. States of mind properly so-called would have to be such that their subjects are in a position to issue *incorrigible* reports on them. However, Rorty argued that accepting incorrigibility at face value is inconsistent with materialist realism; it requires commitment to Cartesian dualism. The alternative, it seems, is to deny there is a distinct category of mental states. *Rorty's Dilemma* says:

Either mental states are peculiar states, over and above bodily states (so Cartesian mindbody dualism is true), or we must embrace eliminativism about the mental.

As Rorty saw it, the materialist view of the day – that states of mind are simply states inside our body (specifically, states of our brain) – is inconsistent with the idea that we could make incorrigible pronouncements on our own present states of mind, ones that no one else – not even a qualified brain scientist – could overturn. But it is also inconsistent with materialism to suggest that mental states are 'peculiar' states that afford incorrigibility. Hence the dilemma. Rorty's own solution was to suggest that a committed materialist would have to give up altogether on the mentalphysical separation and on the idea that mental phenomena as such have a distinct character that separates them from all other phenomena. And this, in effect, meant embracing *eliminativism* – an overkill position that Rorty himself had already advocated earlier.⁸

Finally, focusing more on epistemological issues, contemporary philosophers have sought to explain the commonsense idea that, as subjects of mental states, we each enjoy *basic selfknowledge* of those states. This is reflected in striking *epistemic asymmetries* between our

⁷ Richard Rorty, 'Incorrigibility as the Mark of the Mental', Journal of Philosophy 67 (1970), 399 - 424.

⁸ See Richard Rorty 'Mind-Body Identity, Privacy, and Categories', Review of Metaphysics 19 (1965), 24 – 54.

spontaneous self-attributions of present mental states ('I have a terrible headache', 'I wish I could get out of here', etc.) – our so-called *avowals* – and all other pronouncements we make on contingent matters. We typically take avowals at face value, we do not expect avowers to have justification or reasons for them; and, except under unusual circumstances, we do not take ourselves to be in a position to correct them. Yet avowals are taken to represent genuine – indeed privileged – knowledge we have of our own states of mind. Basic self-knowledge, then, appears to be at once *epistemically base-less* and *epistemically privileged*. This gives rise to the following, Basic Self-Knowledge Dilemma:

Either we have a peculiar way of knowing our own minds that is underwritten by a form of privileged access or we do not have any genuine (let alone privileged) knowledge of our own minds.

Contemporary epistemologists standardly assume that, to have genuine knowledge of some range of facts, one must employ some specific *way* of acquiring the relevant beliefs (some *epistemic basis*). But then it follows that, to have genuine basic self-knowledge, there would have to be a distinctive kind of epistemic access we have to our own minds. Note, first, that to accept this is, in effect, to reject the epistemic base-lessness of avowals as merely apparent. But, moreover, to support genuinely *privileged* knowledge, the type of access would have to be superior to all others. So, if our avowals are epistemically base-less they cannot be instances of genuine knowledge. Whereas if they are based on ordinary ways of knowing, they cannot amount to privileged knowledge.

Here, again, a well-known Wittgensteinian response grabs the defeatest horn of the dilemma, maintaining that it is a mistake to think of avowals as representing *any* kind of knowledge. I consider this, too, to be overkill.

Summing up, we have: Peculiar reference or No reference; Peculiar minds or No minds; Peculiar knowledge or No knowledge.

At the root of each dilemma there is a certain presupposition. In the case of Anscombe's semantic dilemma, the presupposition was this: Successful reference to oneself using 'T requires deploying some foolproof means of recognizing or singling out the referent of 'T. In the case of Rorty's Dilemma, the presupposition was that avowals' incorrigibility must be grounded in the peculiarly non-bodily nature of mental states. And, in the case of the Basic Self-Knowledge Dilemma, the presupposition was that we could only possess privileged basic self-knowledge if we had privileged epistemic access to our minds. In each of these cases, rejecting the relevant presupposition may allow us to slip between the horns of the dilemma. I myself have argued that rejecting Anscombe's presupposition should allow us to accommodate the peculiar semantic success of 'T' without invoking a Cartesian Ego as its semantic value.' And I have argued that rejecting the presuppositions behind the other two dilemmas should allow us to accommodate the seemingly peculiar epistemic status of avowals without postulating mental states over and above states of our bodies.¹⁰ So we need not be driven by the dilemmas to referential deflationism about 'T', or metaphysical eliminativism about mental states, or epistemic defeatism about self-knowledge.

3. Neo-Expressivism about First-Person Discourse

In earlier work,¹¹ I proposed an approach to first-person discourse, avowals' distinctive security, and basic self-knowledge that is designed to avoid the above three dilemmas. The approach takes its initial lead from a form of *expressivism* often associated with Wittgenstein. On a

⁹ Dorit Bar-On, *Speaking My Mind: Expression and Self-Knowledge* (Oxford: Clarendon Press, 2004), Chapters 2-3.

¹⁰ Op. cit. note 8. See also e.g. Dorit Bar-On, 'First-Person Authority: Dualism, Constitutivism, and Neo-Expressivism', Erkenntnis 71 (2009), 53 – 71, and 'Externalism and Skepticism: Recognition, Expression, and Self-Knowledge', in A. Coliva (Ed.) Self-Knowledge and the Self (Oxford: Oxford University Press, 2011), 189 – 211.

¹¹ In e.g. op. cit. note 8, 9.

simple version of the view, our avowals only serve to express, and in no way report or describe, expressed states of mind. That view portrays avowals as being fully on a par – semantically, and therefore epistemically – with nonverbal expressive behaviors such as winces and moans. They are neither truth-evaluable nor epistemically assessable. Now, it is easy to see that simple expressivism purchases the contrasts between avowals and other pronouncements at the (heavy) price of denying palpable semantic, logical, and epistemic continuities between them. (In this respect, the view is analogous to Anscombe's No Reference view of 'I'.¹²)

3.1 Neo-Expressivism

The *neo-expressivist* view I have advocated seeks to avoid the obvious difficulties with simple expressivist on the *neo-expressivist* account, avowals such as 'I am so glad to see you!' (spontaneously volunteered), 'I'd love some dessert' (produced in response to a query), or 'Boy, I love this movie' (said to oneself in inner speech) do indeed enjoy a special security. And this security is to be explained by appeal to the fact that they serve to *express the self-attributed states themselves*, rather than to issue descriptive reports that merely convey one's belief about their presence. However, this explanatory claim concerns avowals' character *as acts*. Avowals understood as acts are contextually interchangeable with non-self-attributive verbal expressions of states of mind (e.g., saying, or thinking, 'It's so good to see you!', or 'Dessert would be nice!'). And they are importantly continuous with nonverbal expressive acts (such as giving a hug). Like other expressive acts, avowals (whether made in outer speech *or* silently) give vent to the very states of mind that the avowals understood *as products* (that is, *qua* linguistic or mental representational tokens) ascribe to the avower. That is to say, when avowing, one performs a distinct type of act: one expresses – in the *action* sense – the very state that is self-attributed by the proposition that the

¹² As discussed in op. cit. note 8.

expressive vehicle one uses expresses in the *semantic* sense. Note that, contra Anscombe, the expressive vehicle used in avowing – an articulate 'I'-ascription token – can be taken to have face-value semantic properties. When avowing: 'I am in pain', I use a token sentence that refers to me (and ascribes to me being in pain. By contrast, 'This hurts' does not use a vehicle that explicitly refers to me, and a grimace neither refers to anyone nor ascribes to them any property.

In general, when engaging in expressive acts, individuals give direct voice to present states of mind. Such acts are designed to *show* expressers' states of mind – as opposed to merely telling of them. Importantly, on my view, one can show a state of mind by *speaking one's mind*, where an act of speaking one's mind uses a linguistic vehicle. (So, instead of showing your being happy to see your friend by giving a hug, you could say: 'So great to see you!' *or* avow: 'I'm so happy to see you'. Same state of mind, different expressive vehicles.) Like other expressive acts that employ linguistic vehicles (slurs and pejoratives being one case in point), avowals can rely on semantic and pragmatic features of the relevant vehicles to show the expressed states. In some analogy to explicit performatives, we can take avowing uses of the first-person to be specialized linguistic vehicles designed to show self-attributed states. In creatures like us, some of the communicative roles played by the more visceral showing afforded by natural expressions, such as animals' growls, bared teeth, grimaces, and so on, are taken over by spontaneous, competent use and immediate uptake of linguistic vehicles.

When it comes to the distinctive security of avowals, the neo-expressivist proposal is that what explains the security is the expressive character of avowals, rather than any epistemic basis on which they are made. As expressive acts, avowals – like other expressions, including natural expressions, and unlike evidential reports (whether third- or first-person) – are indeed (by design) both *produced* and *taken* as epistemically base-less pronouncements. It is for this reason that it is

11

inappropriate to expect them to be backed up by justification or reasons. Yet, insofar as avowals use as expressive vehicles truth-evaluable sentences that semantically express propositions about oneself, they are importantly different from other kinds of expressions of the relevant states. For, understood as 'products' avowals are truth-evaluable self-attributions. Such attributions can be evaluated in terms of their status as items of knowledge (privileged or not). Thus, although the neoexpressivist account borrows from traditional avowal expressivism the insight that avowals' distinctive security is due to their similarities to other expressive acts, it departs from the traditional account in highlighting the fact that, like various mental and non-mental descriptive reports, avowals use expressive vehicles – sentence- or thought-tokens – that are semantically complex and are truth-evaluable.¹³

However, in opposition to *epistemic* approaches, the neo-expressivist denies that explaining what renders avowals instances of (privileged) *knowledge* is either necessary or sufficient for explaining avowals' distinctive security. On the contrary, she maintains that the secure status is to be explained by appeal to avowals' expressive character; the explanation of what (if anything) renders avowals instances of (privileged) self-knowledge is to be provided separately. And, in opposition to both dualist and so-called *constitutivist* accounts, the neo-expressivist does not

¹³ Linguistic expressions can supplement *and* supplant non-linguistic ones. When it comes, more specifically, to avowals, it might be argued that they wear the states they express on their linguistic sleeve, as it were. An avowal such as 'I hope we'll get some rain today' explicitly names a kind of state of mind (a hope) and articulates its content (that it rain today), as well as attributing it to a certain individual. By contrast, 'Oh for some rain today!' expresses one's hopeful state without naming it or attributing it to oneself. The two utterances (which may, of course, be produced in *sotto* voce, with no audience present) can be used, in context, to (a-)express the same state of mind. Arguably, however, the former, but *not* the latter also uses a specialized vehicle for expressing a self-belief. This difference may have significant consequences, which I cannot discuss here, for reasons of space. (I discuss these matters in "Avowals and Active Self-Beliefs" (tentative title), in-progress.

For a full development of the account summarized so far and an explanation of the strong presumption of truth governing avowals, see op. cit. note 8, Chapters 6–8. See also Dorit Bar-On, 'Expression: Acts, Products, and Meaning', in S. Gross et. al. (eds.), Meaning Without Representation: Essays on Truth, Expression, Normativity, and Naturalism (Oxford: Oxford University Press, 2015), 180 – 209.

explain avowals' special epistemic status by appeal to the fact that *being* in such states is essentially linked to (or guarantees) having direct knowledge of them.¹⁴

How should base-less yet privileged basic self-knowledge be explained? In earlier work, I canvassed several possible answers to this question, all of which involve rejecting the presupposition that where there is knowledge, there must be a distinctive *way of knowing.*¹⁵ And in work in-progress,¹⁶ I take steps toward developing a substantive epistemology of basic self-knowledge that rejects this presupposition, thereby avoiding the third dilemma.

4. Ethical Neo-Expressivism

It is worth emphasizing that the neo-expressivist view is not intended to apply across the board, to all areas of discourse; it is not a version of what Price calls *global* expressivism.¹⁷ Global expressivism recommends generalizing the idea that the 'point' of a given discourse is to give voice to certain attitudes, and moving towards a revisionary, anti-representationalist ('subject-centered') semantics (which, following Brandom, Price describes as an expressivist semantics). Such a semantics would explain the meanings of statements in *any* given discourse in terms of the states of mind characteristically expressed in making those statements.¹⁸ (And, Price further believes, adopting such a semantics would obviate all 'object-centered' metaphysical questions.) This

¹¹ I criticize epistemic and constitutivist approaches in op. cit. note 8, as well as, e.g., 'Minding the Gap: In Defense of Mind-mind Continuity', in Kevin M. Cahill and Thomas Raleigh (eds.), Wittgenstein and Naturalism (New York, NY: Taylor and Frances, 2018), 177 – 203; 'Belief Self-Knowledge' (with Kate Nolfi), Oxford Handbook Online (2016); and 'Epistemological Disjunctivism: Perception, Expression, and Self-Knowledge' (with Drew Johnson), in Pritchard et al. (eds.), Epistemological Disjunctivism (Oxford: Oxford University Press, forthcoming).

¹⁵ See op. cit. note 8, Chapter 9, and Bar-On and Nolfi, op. cit. note 14.

¹⁶ 'No "How" Privileged Self-Knowledge' (tentative title).

¹⁷ See, e.g. Huw Price, 'Prospects for Global Expressivism', in H. Price et al. (authors) *Expressivism, Pragmatism, and Representationalism* (Cambridge, MA: Cambridge University Press, 2013), 147 – 194.

¹⁸ Several authors have proposed that global expressivism be seen as a *metasemantic* rather than a semantic view. (See, e.g. Michael Ridge, *Impassioned Belief* (Oxford: Oxford University Press, 2014).) I recognize the relevance of this option to the line of thought I develop in this and the following section. However, I think it deserves a separate treatment, which I cannot provide here.

globalizing move is no part of the neo-expressivist view I advocate, insofar as that view can potentially apply piecemeal, to specific areas of discourse. Neo-expressivism exploits the possibility of capturing semantic and logical continuities across different domains by adhering to a relatively neutral - and more conservative - semantics that reflects uniformities in linguistic use. Such a semantics, I would urge, neither settles nor preempts metaphysical debates at the local level. Granted, utterances of 'Snow is white', 'John loves Mary', 'Murder is wrong', 'John buttered the toast with a knife', 'That joke was funny', 'I'm annoved at my friend', 'Rain is likely', all share certain propositional-compositional features that make them apt for various linguistic transformations. A successful semantics should capture these shared features. But I see no specific role for expressivist insights to play in devising such a semantics. Expressivist insights, I would argue, are more suitable for capturing what we do with certain vocabularies, and thus for capturing at least some systematic differences among different areas of discourse. (Some but not all. As we will see later, assuming that metaphysical questions are not ruled out of court, some kinds of differences may be best captured by a more direct appeal to types of objects, properties, or structures that the different discourses are about.) I return to this below.

Still, the neo-expressivist framework can be useful in addressing puzzles that arise in connection with areas other than mentalistic discourse. To see how, consider, first, that the distinction between a-expressing and s-expressing, and between acts of expressing and the expressive vehicles used, apply across all areas of discourse, and regardless of what semantic, epistemological, or metaphysical analysis we adopt for the relevant domain. Regardless of the particular domain with which one is concerned, we can clearly separate the (three-place) expression relation that holds between individuals, their mental states, and some expressive vehicle (that is, a-expression) from the (two-place) expression relation that holds between meaningful

14

tokens and their contents (that is, s-expression). And, of course, *the act of making a claim* – again, regardless of what the claim happens to be about – is something different from what is produced by the act, namely, a sentence- (or thought-) token.¹⁹

So what might things look like when we turn to ethical discourse, for example? In contrast with the case of avowals, in the case of ethical claims, the sentences produced do not s-express selfascriptive propositions. We should not suppose, with subjectivism, that an ethical claim such as 'Hunting for fun is wrong', understood as a linguistic token, has the same semantic content as a self-ascription such as 'I disapprove of hunting for fun'. Still, on analogy with the neo-expressivist treatment of avowals, we can explain one central contrast between such an ethical claim and an ordinary descriptive claim such as 'Hunting for fun is common in the U.S.' by appeal to the expressive character of acts of making ethical claims. Elsewhere,²⁰ Matthew Chrisman and I have put forward ethical neo-expressivism: the view that ethical claims - understood as acts - are 'in the business of a-expressing motivational states (whatever those turn out to be, according to our best moral psychology).²¹ However, the expressive vehicles we use when making ethical claims typically, sentences that use ethical vocabulary – can still be seen as s-expressing truth-evaluable propositions. They can be regarded as having standard *truth-conditions* (more on truth-conditions below, in Section 5). At least when it comes to simple atomic sentences, these propositions can be specified disquotationally; so, for example, e.g., 'Murder is wrong' can be taken to s-express, quite simply, the proposition that murder is wrong. Thus, there is no presumption that a semantic

¹⁹ In cases where nonlinguistic vehicles are used, we can think of what is produced as a token of a type – of, e.g. facial contortion, bodily gesture, vocalization, and so on.

²⁰ Dorit Bar-On and Matthew Chrisman, 'Ethical Neo-Expressivism', in *Oxford Studies in Metaethics*, Vol. V (2009), 132 – 165. See also Dorit Bar-On et al. '(How) Is Ethical Neo-Expressivism a Hybrid View?', in G. Fletcher and M. Ridge (eds.) *Having It Both Ways: Hybrid Theories and Modern Metaethics* (Oxford University Press, 2014), 223 – 247.

²¹ This leaves room for the possibility that ethical claims *also* a-express beliefs – a (hybrid) view defended, in somewhat different terms, by Drew Johnson (dissertation in-progress).

analysis of ethical claims should – or could – yield a paraphrase in some preferred vocabulary that captures the literal meaning of such claims in other terms. And this explains why ethical claims that use simple indicative sentences behave in the same ways (logically and grammatically) as ordinary descriptive claims. (It is a separate issue – to be handled by the *metaphysics* of ethics – to determine what, if any, conditions must obtain in the world for ethical sentences to be true or false. See below, Section 5.)²²

So, armed with the right distinctions, we can retain a key expressivist idea – that engaging in ethical discourse and reflection involves expressing the very attitudes that motivate and explain our actions – while avoiding a host of difficulties that beset traditional expressivism, as well as some later developments of it. Where traditional expressivists have erred, on this way of looking at things, is in thinking that the expressive function of ethical claims is somehow executed through the linguistic *meanings* of ethical sentences. As in the case of avowals, the proposition that is s-expressed does not automatically settle what mental state is characteristically a-expressed by acts of making the claim. In the case of ethical claim, the neo-expressivist proposal is that *acts* of making ethical claims serve to a-express the very same states whose presence is required for understanding the perceived motivational force of such utterances.²³

I think there are very good reasons to resist forging the link to motivation through the literal meaning (or truth-conditions) of ethical claims.²⁴ In general, I think we should locate the

²² It is consistent with the present proposal to maintain that a complete semantic analysis of sentences containing ethical terms such as 'good' can go beyond the disquotational pairing up of sentences with propositions. The key point is to deny that such an analysis must provide a paraphrase of some sort, involving lexical decomposition of the relevant terms, or spelling out (nondisquotationally) necessary and sufficient conditions, for example.

²⁸ Thus, neo-expressivism does not purport to settle the question *which* psychological states qualify as motivational. Humeans will insist that they must be noncognitive; others may demur. Moreover, even if one sides with the Humeans, and insists that one who makes an ethical claim is a-expressing a noncognitive motivational attitude, it's still possible to allow that one is *also* a-expressing a *belief* whose content is given by the proposition that is s-expressed by the sentence used. For some discussion, see Bar-On and Chrisman op. cit. note 22.

²⁴ So doing entails that it is *conceptually impossible* for someone to issue an ethical claim without being motivated to act (or refrain from acting) in accordance with it. A virtue of ethical neo-expressivism is that it can capture the

contribution of the idea of expression *not* in what we do *in* language, but rather in what we do *with* language – not in what ethical sentences *say*, but in what we characteristically do when using these sentences to make ethical claims. (This, I would argue, tells against recent attempts at a wholesale replacement of traditional semantics by an 'ideationalist' semantics.²⁶ If I am right, on a proper conception of expression, which utilizes the distinction between a- and s- expression, and on a suitably modest understanding of propositions, such global semantic attempts may be overkill. It is far from clear how such theories can preserve the character of ethical claims as truth-evaluable statements, capable of participating in logical inferences, embeddable in negation and other truth-functional contexts, etc. And it has proved difficult to develop a thoroughgoing 'expressivist semantics' for logically complex sentences with mixed ethical and non-ethical parts.²⁶)

Accepting a traditional semantic framework for ethical sentences – now understood in terms of the idea that indicative ethical sentences s-express propositions or have standard truthconditions – can allow us to preserve continuities between these sentences and other indicative sentences and to avoid Frege-Geach type problems. However, contrary to what many expressivists and their opponents believe, adopting a truth-conditional semantics for a given discourse need not commit one to a particular view on the nature of truth or to a realist metaphysics underlying the discourse. (Just as one can accept that, e.g. 'I am in pain' is true iff DB is in pain without committing to a dualist view on the nature of pain, so one can accept that 'What she did was morally wrong' iff a particular act is morally wrong without committing to 'spooky', irreducibly

alleged internal connection to motivation without having this implication. In Bar-On and Chrisman (2009) and in Bar-On et al., it is argued that this provides resources for capturing a fairly strong "internal" connection between ethical claims and action, as well as providing a more nuanced array of diagnoses of different ways the connection between making an (apparently) ethical claim and motivation can be broken.

²⁵ See, e.g., Mark Richard, 'What Would an Expressivist Semantics Be?', in Gross et. al. *op. cit.* note 13, 137 - 159.

²⁶ If Mark Schroeder is right, this simply cannot be done. See his *Being For* (Oxford: Oxford University Press, 2008).

normative ethical facts or properties.²⁷ Indeed, maintaining that ethical claims (as products) sexpress propositions is even consistent with embracing ethical anti-realism (as many expressivists do).²⁸ Thus, if there are reasons for those inclined toward expressivism in ethics to resist the neoexpressivist account of motivational internalism, worries about its ontological commitment need not be among them.

5. Meaning, Truth, and Truth-Conditions²⁹

Expressivism traditionally understood is one of a family of views that tries to save differences between a given discourse and others deemed less problematic in terms of a *plurality of meanings.* (Early on I mentioned Dummett as an important source of this general approach, applying it initially to explain important contrasts between mathematics and other discourses.) More recently, Wright (who has been largely sympathetic to Dummett's semantic approach), has argued against locating differences in the types of meanings or propositions expressed by sentences of various discourses. For, Wright thinks that there are systematic difficulties with the Dummettian attempt to provide an anti-realist construal of the meanings of sentences that is apt to replace truthconditional semantics. Instead, Wright suggests, we should recognize a plurality of *truth properties.* The suggestion[#] has given rise to a program known as *pluralism about truth.* The truth pluralist

²⁷ It is helpful to separate, in this connection, various strands in what draws philosophers to expressivism in various domains. In particular, even as regards traditional expressivism, it is perfectly reasonable to separate the *positive expressivist* strand – i.e., the idea that claims in a given domain function to express a distinctive (noncognitive) type of mental state or attitude – from the *negative ontological* strand – i.e., the idea that there are no properties for terms in the relevant domain to denote (or facts for claims in the relevant domain to report or describe). For relevant discussion, see Dorit Bar-On, 'Expression, Truth, and Reality: Some Variations on Themes from Wright', in A. Coliva (ed.) Mind, Meaning, and Knowledge: Themes from the Philosophy of Crispin Wright (Oxford: Oxford University Press, 2012), 162 – 194, and Dorit Bar-On and James Sias, 'Varieties of Expressivism', Philosophy Compass 8 (2013), 699 – 713, Section 2.

²⁸ So, it is notable that, strictly speaking, even traditional expressivism has always been neutral with respect to the realism/anti-realism debate in ethics (though, of course, nearly all expressivists have also been anti-realists).

²⁰ This section overlaps with parts of Dorit Bar-On and Keith Simmons 'Truth: One or Many or Both?', in N. Kellen et al. (eds.), *Phuralisms in Truth and Logic* (London: Palgrave Macmillan, 2018).

²⁰ First made in Crispin Wright, Truth and Objectivity (Cambridge, MA: Harvard University Press, 1992).

rejects the deflationist claim that truth is at best a 'shallow', insubstantial property. She proposes instead that we should recognize a plurality of substantive truth properties – properties such as *superwarrant, superrcohernece,* or *correspondence* – each appropriate for a different range of domains of discourse. An advantage claimed by pluralists for their view is that, unlike quietists views, it may allow us to capture intuitive differences between areas of discourse, while also avoiding the need for rejecting truth-conditional semantics across the board.

In a recent paper – "Truth: One or Many or Both?"³¹ – Bar-On and Simmons argue against what is considered to be the most promising version of truth pluralism: moderate pluralism (due to Michael Lynch³²). We think the view faces serious difficulties, not the least of which is that of accommodating universal generalizations (such as 'Everything John says is true'), and the truth of 'mixed' sentences (such as 'Hunting for sport is prevalent in the Western U.S. and is morally reprehensible'). (Some of these difficulties, note, are highly reminiscent of difficulties touched on earlier, with Anscombe's view of 'I', and with the purely expressivist view of the semantics of avowals and of ethical claims.) More directly relevant to our concerns here, we argue that pluralists have given little reason why we should not adhere to *alethic conservatism*, according to which there is only one way for sentences, propositions, beliefs, etc. to be true - though, when they are, there may be multiple ways things can be to make them so. (Thus, a joke's being funny may be a very different sort of thing from someone's act being wrong, or a number being divisible by 2, or a chair's being brown; and then again an act's being morally wrong may be a different sort of thing from it being politically wrong.) The appeal to a plurality of truth properties, we argue, contributes no explanatory power beyond what can be obtained by focusing on a plurality in kinds of worldly

³¹ Op. cit., note 30.

³² Michael Lynch, *Truth as One and Many* (Oxford: Oxford University Press, 2009).

conditions that are apt to render true claims made in different domains.

So, as a default position, we think that the debates between various realist and anti-realist construals of given domains should be reconstructed *neither* in terms of kinds of truth properties *nor* in terms of kinds of meaning. Instead, we suggest (more conservatively), that the plurality be assigned to the relevant realms of facts – to the *worldly conditions* that could *render* statements in given domains true. Like the semantic conservatism advocated by neo-expressivism, we think that the alethic conservatism we recommend may allow us to save differences without incurring unwanted ontological commitments.

A key move in resisting truth pluralism is recognizing that the notion of *truth-conditions* is invoked in philosophical discussions of truth in two ways that can – and, we submit, *should* – be separated. Briefly, consider the familiar slogan 'The meaning of a sentence is given by its truthconditions' used to characterize the Davidsonian theory of meaning. That slogan aims to capture the idea that we can theoretically characterize the meanings of sentences in a given language in terms of theorems derived from the axioms of a theory of truth for that language. These would be *meaning-specifying* biconditionals, such as

(W) 'Wasser ist nass' is true if, and only if, water is wet.

Here the right-hand side picks out the worldly condition – water's being wet – under which the mentioned left-hand sentence is true. However, as recognized by Davidson,³³ the right-hand side must pick out that worldly condition in a way that is fit to capture the *semantic place* occupied by the mentioned sentence – that is to say, in terms of its relations to *other* sentences. So it would be bad, for example, if our theory yielded the biconditional

³⁸ Donald Davidson, 'Reply to Foster', reprinted in *Inquiries into Truth and Interpretation* (New York: Oxford University Press, 1984), 171 – 180.

(W') 'Wasser ist nass' is true if, and only if, H₂O is wet.

since (W') is, intuitively, not meaning-giving. Thus, in a Davidsonian theory of meaning, it matters *how* the conditions on the right-hand side are picked out. The notion of truth-conditions relevant here is a *semantic* one, to wit:

(i) truth-conditions as they figure in *meaning-giving* biconditionals – worldly conditions
picked out in a way fit for specifying the meaning of a given sentence.³⁴

This notion can be distinguished from a *metaphysical* notion of truth-conditions, to wit:

 (ii) *worldly conditions* (objects, properties, states of affairs – if any) identified by a metaphysician as revealing the underlying nature, ontological constitution, etc., of elements in a given domain.

Truth-conditions in the semantic sense are what Davidson focuses on in his seminal 'Truth and Meaning';³⁵ and he presents them as relatively neutral, metaphysically speaking. He says:

If we suppose questions of logical grammar settled, sentences like 'Bardot is good' raise no special problems for a truth definition. The deep differences between descriptive and evaluative (emotive, expressive, etc.) terms do not show here. ... we ought not to boggle at "Bardot is good" is true if and only if Bardot is good'; in a theory of truth, this consequence should follow with the rest, keeping track, as must be done, of the semantic location of such sentences in the language as a whole – of their relation to generalizations, their role in such compound sentences as 'Bardot is good and Bardot is foolish', and so on.

³⁴ On some views (though not Davidson's), truth-conditions so understood are what competent speakers have mastered (or internalized) and know, at least implicitly. For relevant discussion and references, see Dorit Bar-On 'Anti-Realism and Speaker Knowledge', Synthese 106 (1996), 139-166.

We must here set aside the difficult question whether – and how – a truth theory for a language L can, as Davidson hoped, do all that we may expect of a theory of meaning for L. (For discussion, see Ernie LePore and Kirk Ludwig, *Donald Davidson* (Oxford: Clarendon Press, 2005), Part I.)

³⁵ Reprinted in op. cit., note 34, 17 – 36.

What is special to evaluative words is simply not touched: the mystery is transferred from the word 'good' in the object-language to its translation in the meta-language.³⁶

(And, Davidson might have also said, 'I'm feeling sad' – uttered by DB at time t and place p – is true iff DB is feeling sad at t and p. Here, too, the truth-conditional theory will not capture what is special about feeling sad *or* about avowing – as opposed to reporting – one's sadness.)

Here are some key features we take to characterize the semantic notion of truth-conditions:

(1s³⁷) They are a product of a *logico-semantic* analysis that exhibits the truth-conditions of sentences as a function of the semantic values of their parts, in a way that reveals how they systematically interact with other sentences and sentence parts, how they embed in various constructions (such as conditionals, modal and propositional attitude contexts), and so on.³⁸

(2s) Calling the relevant conditions '*truth*-conditions' signals the involvement of truth in recovering logical structure, entailment relations among sentences, etc. But to play this role 'truth' need not be understood as denoting any robust (or specific) metaphysical property. At the same time, the use of the Tarskian truth schema to specify sentences' meanings in no way *commits* one to deflationism about truth, and on the view of leading deflationists is in fact incompatible with it.³⁹

(3s) A Davidsonian semantic theory is *not* a purely disquotational theory of *meaning*. Davidsonian semantic analysis can uncover covert ambiguities and context-sensitivity, surprising logical forms and entailment relations,⁴⁰ and so on.

³⁶ Op. cit., note 34, 31.

³⁷ 's' for *semantic*.

[®] This, as well as some of the main features below, apply even more clearly in the case of contemporary formal semantic analyses of natural language that make ostensible use of the notion of truth/truth-conditions of the sort offered by, e.g. Irene Heim and Angelika Kratzer, *Semantics in Generative Grammar* (Oxford: Blackwell, 1998). Thanks to Matthew Chrisman for highlighting this point.

³⁰ As is well known, Davidson himself has argued *against* deflationism about truth, for reasons we cannot rehearse here.

⁶⁰ A good example is Davidson's own analysis of action sentences. See Donald Davidson, 'The Logical Form of Action Sentences', reprinted in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 105 – 122.

(4s) At the same time, one should not expect truth-conditional analysis to yield in every case an analytic paraphrase. Meaning-giving biconditionals can be modest, semantically speaking. Often the best theory can do is offer a disquotational truth-condition, and there is no presumption that it will reveal anything interesting about the meaning of semantic atoms such as 'dog', 'mountain', 'walk', 'love', 'blue', let alone 'happy', 'funny', 'wrong', and so on.

(5s) More importantly, as Davidson himself remarks, meaning-giving biconditionals are also *ontologically* modest. Having settled on the meanings of mathematical sentences, for example, it is open to the semantic theorist (who may or may not herself be a metaphysician) to ponder the nature of mathematical facts – whether there are numbers, what kinds of things they are, and so on. *Using* the vocabulary of an area of discourse, and putting on a metaphysician's hat, so to speak, one can ask questions such as 'What in the world (if anything) makes something *beautiful, morally right, funny*?'. These questions are not in general questions about language, but are instead raised in the 'material' mode, *using* language. (There is no apriori reason to expect that the best analysis provided by the semanticist would dictate unique answers to the ontological questions.)

(6s) In the case of areas of discourse thought to involve commitment to ontologically problematic facts, acknowledging that meanings can still be specified by using biconditionals derived from a truth theory allows us to accommodate undeniable logico-semantic continuities between (at least some) allegedly problematic areas and more straightforwardly 'descriptive' ones.

(7s) At the same time, as remarked earlier, the association of truth-conditional meanings with, e.g., ethical sentences does not automatically remove all worries about the problematic character of putative facts in the relevant domain. One can still be an anti-realist about ethics, even if ethical sentences are assigned truth-conditional meanings.

23

This brings us to the second notion of truth-conditions. On the present way of carving things up, these are *worldly conditions* invoked when attempting to answer metaphysical questions about ontology, nature, constitution, and so on – questions such as e.g., What is pain? What is color? What makes a person happy? How are mountains to be individuated? When does S know that p? These conditions have the following important features:

(1m⁴) They are semantically innocent. They are provided – and are offered in response to questions that do not concern the meanings of sentences. Metaphysical questions *can* follow on the heels of assigning semantic truth-conditions to sentences, as when we learn from the semanticist – perhaps disappointingly – that 'Torturing animals is morally wrong' is true iff, well, torturing animals is morally wrong, and we press: but what *makes* a practice morally wrong? However, the question about the *nature* of moral wrongness can also arise prior to, and independently of, recovering the truth-conditional meanings of sentences involving the phrase 'morally wrong'.

(2m) Raising such question, of course, requires making competent use of language. However, except when one's metaphysical inquiry concerns language, the metaphysical search for worldly conditions does not implicate any particular analysis of meaning. (Although, at times, admittedly, metaphysical disagreements *can* be traced to divergent uses of language.) Familiarly, when the metaphysician of mind tells us that pain is a certain configuration of brain states, or essentially a functional state, this is not offered as a meaning analysis. Similarly for the utilitarian reduction of the goodness of actions to their maximization of utility, and various other reductive accounts.⁴²

⁴¹ 'm' for metaphysical.

⁴² As explained in op. cit., note 30, fn XX, it is a misunderstanding to think that semantic externalism gives the lie to the metaphysical neutrality of semantic analysis just suggested. Briefly, all the externalist semantic theory is in a position to claim is that the meaning of 'water' is dependent on the nature of water, *whatever that* is. If water is in fact identical to the chemical substance H4O, then being H4O is constitutive of its metaphysical nature. But there is no expectation that 'H4O' should figure in an externalist semantic account of the term 'water'.

(3m) The worldly conditions that figure in a Davidsonian semantic theory can be seen as *truth*-conditions only in the sense that we can think of them as *making true* some sentences/ propositions/beliefs/etc. and not others. But, so understood, they are conditions that are individuated metaphysically, not semantically. Consider: the worldly condition of H₂O's being wet is, metaphysically speaking, one and the same condition as that of water's being wet. So this worldly condition – described either way – makes true the sentence 'Water is wet'. But, for all that, "'Water is wet" is true iff H₂O is wet' is *not* a meaning-giving biconditional for 'Water is wet'. It will not (or at any rate should not) be a theorem derivable from a truth theory for English. (This may be why deflationists about truth are perfectly happy to allow that we do – and can, consistently with deflationism – speak of worldly conditions that we loosely refer to as truth-conditions.)⁴⁴

Going back to truth pluralism, we think that there is nothing to be gained, explanatorily speaking, by invoking a metaphysical plurality of truth properties over and above whatever plurality is recognized in the worldly conditions that our metaphysicians have identified or proposed, as they investigate different domains of discourse. Consider a metaphysical inquiry into what makes something illegal. Such an inquiry may conclude that the legality of this or that act depends in some systematic way on our legal practices, historical facts, and so on. Perhaps it will conclude that nothing *is* legal that would not be *judged* legal by an ideally placed judge, so that it makes no sense to suppose that the legality of an act could forever elude human judgment. We can summarize the results of this inquiry by saying: In legal matters (in contrast with other sorts of matters), truth is judgment-dependent. But the question is whether putting things this way really commits us to a distinct *truth property* possessed by all and only true *legal* sentences, as contrasted with, say,

⁴⁸ For discussion and references, see Dorit Bar-On et al. 'Deflationism and Truth-Condition Theories of Meaning', reprinted with Postscript in B. P. Armour-Garb and JC Beall (eds.), *Deflationary Truth* (Open Court, 2004), 321 – 352.

sentences of everyday discourse about mid-sized dry goods. The alethic monist eschews this commitment and insists that there is only one way for true sentences, propositions, beliefs, etc. to be true, though there may be multiple kinds of worldly conditions that make them true. And the relevant plurality can in principle be captured without engaging in semantic ascent.⁴⁴

* * *

On my preferred view, the proper job of semantics is to systematize our understanding as logico-semantic sentential and sub-sentential relations, sameness of meaning across different languages, consistency of content across force-stripping contexts, and so on. I take this task to be theoretically independent of settling on the correct metaphysics for a given discourse. (The seeds of this way of thinking of truth-conditional semantics, I pointed out, are already in Davidson's seminal "Truth & Meaning".) It should not be supposed to be part of semantics' job to reveal to us nature's joints or to offer a substantive characterization of the metaphysical conditions that must obtain for given interpreted sentences to be true. Whether or not the propositions s-expressed by claims in some area of discourse have ontologically problematic 'truth-makers' is a matter to be settled by metaphysicians, not semanticists; so is the question what in the world *renders* those claims true or false. As we saw in the case of avowals, as well as ethical claims - assigning Davidsonian truth-conditions to the propositions expressed by sentences of a discourse is also consistent with discerning a distinctively expressive function of the discourse. That explanation of systematic differences across areas of discourse did not require assigning different types of meanings or different truth properties to claims in the different areas. Taking advantage of such

⁴⁴ Of course, given the equivalence of and <<p> is true>, one *can* advert to a 'formal mode' and speak of the *truth* of 'x is red' being a different sort of thing from the *truth* of 'x is divisible by 2' – indeed, sometimes putting things in terms of truth may be indispensable. However, the alethic monist objects to the further move to a plurality of *truth properties*; she denies that that move is forced on us by taking seriously (rather than being quietist about) debates between realists and anti- realists.

non-semantic resources we can hope to transcend superficial similarities and still save the

differences.45

⁴⁵ Earlier versions of this paper were presented at a workshop on *Expressivisms, Knowledge, and Truth*, held at University College, London, October 19-20, 2018, and as a keynote address at the Meetings of the Society for Exact Philosophy at York University, Toronto, May 19, 2019. I wish to thank audiences at these meetings – and especially Matthew Chrisman, Maria-Jose Frapolli-Sanz, Robert Myers, Huw Price, and Claudine Verheggen – for helpful comments and discussions.