

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Vaaja, Tero

**Title:** Expressivism, Self-Knowledge, and Describing One's Experiences

**Year:** 2014

**Version:**

**Please cite the original version:**

Vaaja, T. (2014). Expressivism, Self-Knowledge, and Describing One's Experiences. *The Journal of Mind and Behavior*, 35(3), 151-166.

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

## Expressivism, Self-Knowledge, and Describing One's Experiences

Tero Vaaja

*University of Jyväskylä*

In this article, I defend an account of self-knowledge that allows us a considerable first-person authority regarding our subjective experiences without invoking privileged access. I examine expressivism about avowals by contrasting it with “detectivist” and “constitutivist” accounts of self-knowledge, following the use of these terms by David Finkelstein. I proceed to present a version of expressivism that preserves some of the valid motivating insights of detectivism and constitutivism as essential parts. Finally, I point out how my account views self-knowledge as a cognitive and conceptual ability that can be cultivated; the account construes self-knowledge as a process.

Keywords: expressivism, first-person authority, avowal

Each of us is normally the best person to ask when it comes to our own feelings and experiences. Speaking about one's own mental states is generally held to carry a special epistemic authority. Moreover, this authority belongs exclusively to the first person; others are not admitted to have a similar claim to know someone's experiences even if they are extremely well-informed and familiar with them. I take these to be facts on first-person authority as they appear in the practice of human life quite universally.

Such authority has a central place in social life; denying it can easily (and legitimately?) be taken as an offence. However, it might be that philosophers have historically been overconfident about the special security of our knowledge of our own minds. Carruthers (2011) argues that self-knowledge is interpretive and prone to confabulation. Schwitzgebel (2011; Hurlburt and Schwitzgebel, 2007) claims that we might be regularly wrong about even quite fundamental features of our conscious experience. Therefore it is important to be clear

about the nature of first-person authority, and the conditions in which it may be legitimately challenged.

In this article, I seek to give a modest account of self-knowledge that still respects the special status of the subject as a knower of her own mental states. I treat commonsensical first-person authority as an explanandum, setting aside accounts that seek to dethrone the notion altogether. I start by presenting two contrasting views about the nature of self-knowledge and the basis of first-person authority. I point out how each of these views, “detectivism” and “constitutivism,” is unsatisfactory and how expressivism about avowals, an idea inherited from Wittgenstein (1953), can be seen as preferable to them. I owe the terms detectivism and constitutivism, as well as the main drift of the argument in the first half of this paper, to Finkelstein (2003). Another way to refer to these two contrasting views would be to call them (species of) empiricism and rationalism about self-knowledge, as is done in Gertler (2011). I proceed to present a version of expressivism that incorporates some of the good insights made by detectivism and constitutivism. As explained in the conclusion, I hope my view to be meritorious in respecting commonsensical first-person authority without invoking privileged access, i.e., an idea of a special epistemic channel that makes self-knowledge unproblematic to come by. I also seek to do justice to the meaning of “self-knowledge” as a process that has to do with the personal development of one’s conception of oneself.

### Detectivism

What is it that makes psychological self-ascriptions, or avowals, especially secure?<sup>1</sup> One way of answering is to appeal to introspection, combined with some form of privileged access. The idea is simple: people come to know what their own mental states are like because they are the ones who directly feel or perceive those states. We are assumed to have an “inner sense,” or some naturally evolved capacity that enables us to inwardly monitor our mental states. These are forms of what Finkelstein (2003) calls detectivism: the view that the source of self-knowledge is a perceptual or quasi-perceptual act of detecting that allows us to find out our own mental states.

So, one possible explanation for first-person authority is a combination of two ideas: first, there is a special way of detecting one’s own mental states; and second, that way of detecting is remarkably reliable. Maybe subjects are not completely infallible about everything that goes on in their conscious experience,

---

<sup>1</sup>I will use “avowal” as an umbrella term to refer to any sincere utterance whereby the subject speaks about her mental condition. This liberal use is not a standard one. According to more restricted uses of the term, what I will later refer to as primitive avowals and intellectual self-ascriptions would not necessarily qualify as avowals.

but they have such a propensity of being right about those things that it cannot be paralleled by any other person.

It is hard to deny that in an obvious sense, the subject of a painful sensation is in a better position to observe that particular pain than anyone else. But it is still far from obvious that this is what grounds the typical way in which first-person authority is granted to subjects, or if this is a good account of what self-knowledge is. Next, I attempt to illustrate the issue by an example; my chosen example in this paper will be a case of describing a sensation of pain.

### *Example 1*

I have an abdominal pain that I need to describe to a physician. I am able to point out its location and give an evaluation of its intensity on a scale of 1 to 10. I will also describe its qualitative character by a few adjectives. After careful consideration and some effort to find the right words, I say (at time  $t_1$ ) that my pain is located about ten centimeters up from my waistline, on the left side of my middle abdomen, its intensity is 6, and it is stinging, sharp, distressing, and penetrating.

When I have finished giving my description, I overhear the word “rip,” or someone suggests it to me. I say (at time  $t_2$ ): “Ripping! Yes. That’s what my pain is like. That’s right; I could not come up with it myself.”

When I eventually say that my pain is ripping, I presumably say it with first-person authority. The fact that I needed help in finding the word might give reason for an interlocutor to not take it completely at face value; a question like “Are you sure that is the right word?” might be justifiable. But if I say sincerely and after careful consideration that “ripping” describes my pain perfectly, it is unclear what could ground the claim of someone who insists that I must nevertheless be wrong. In this kind of a situation, any doubt that another person might harbor about the appropriateness of my pain-description will more naturally target my adeptness in the use of the word, rather than the accuracy of my introspective act.

According to detectivism, my statements about my pain are based on perceptual or quasi-perceptual observing. In this case, I am supposedly monitoring my sensation of pain and detecting a ripping quality in it. But detectivism makes it hard to see why my eventual description of my pain as ripping should carry any special authority. It was, after all, based on the same introspective observation that I had already done at  $t_1$ , without at that time judging my pain to be ripping. We can make the example clearer by stressing that my sensation of pain stays the same from  $t_1$  to  $t_2$ : I am not judging my pain to be ripping at  $t_2$  because it started as non-ripping and then suddenly turned into ripping. Someone could suggest that at  $t_1$  I did not attend to the pain as completely as I did at  $t_2$ ; the suggestion could be that upon hearing the word “rip,” I introspectively probed the pain again to see if the new word fits it, and found a novel ripping quality in it. But it is possible that I would sincerely deny that too, and

testify that my pain features in my experience exactly in the way as it did at  $t_1$ . The quality that made me describe it as ripping was in my awareness from the start; I merely came up with a better description of it.

I think it is fairly plausible that in this situation, where I explicitly admit that I do not derive my eventual pain-description from any distinct introspective act, few people would feel that the authority of my avowal diminishes from  $t_1$  to  $t_2$ . This suggests that detectivism is not adequate to explain the basis of first-person authority.

Maybe we should waive the detectivist idea and state that inward perceptions are not the source of the authority of my avowals. Instead, it could be suggested that first-person authority is only a matter of mastering a language. Adult persons who are competent language-users have learned a stock of everyday phenomenological vocabulary, and they are considered to be beyond criticism in their psychological self-ascriptions just by virtue of the fact that they generally use that vocabulary in a coherent and consistent manner, without regularly coming into conflicts with other competent language-users. Upholding the first-person authority might be seen as a mere pragmatic or social convention.

If we think this way, how unassailable a subject's descriptions of her conscious experiences are will be a function of her adeptness in using experience-vocabulary. The descriptions of a fully competent adult will be authoritative, the descriptions of a young child or a non-native speaker less so. However, what should we do in situations where two people, both perfectly competent in introspecting and describing conscious experiences and who we have independent reasons to believe to be undergoing a similar experience, nevertheless describe that experience in mutually inconsistent ways? Do we then have to assume that at least one of them makes an introspective error? Are we then entitled to waive the first-person authority of one or both of them? For Schwitzgebel (2011), cases like that form the basis of one group of arguments to the effect that people are not in general reliable judges of their own conscious experiences.

### Constitutivism

If Schwitzgebel is right, much of the first-person authority that we normally grant to competent adult people is based on false prejudice. However, there is an alternative view of self-knowledge that denies that describing our experience is essentially a matter of having an accurate perception of one's inner episodes, which is then translated into words. This view, called "constitutivism" by Finkelstein (2003), is also friendlier to first-person authority than detectivism ends up being. Its central idea is that our judgments concerning our inner episodes play a constitutive role in determining what those inner episodes are.

Constitutivism seems insightful especially concerning propositional attitudes like beliefs. When we self-ascribe a belief, it seems that we most typically do

that by *rationally committing* ourselves to a belief, via judging that something is the case. Self-ascribing a belief seems to be the act of forming a belief or settling on a belief, rather than finding one via introspection. As the so-called transparency theories of self-knowledge have emphasized, self-ascriptions of attitudes need not involve any judgment turned inwards, so to speak; they are rather part and parcel with the judgments we make of the outside world.

So at least in some cases, my non-introspective judgments may *constitute* my mental states. Also in the case of descriptions of sensations, my authority may be thought to be “not like the authority of an eyewitness [. . .], but rather like] that of an Army colonel when he *declares* an area off limits” (Finkelstein 2003, p. 28; emphasis in the original). A slightly adapted example will illustrate the point:

### *Example 2*

Two people have an abdominal pain that they describe to a physician. It has been established that their pains are caused by a similar medical condition; they are of the same age, gender and build, the patterns of activation in their nervous systems are highly similar, and their pain-descriptions agree for the most part. In short, we have good independent grounds for believing that they are describing qualitatively similar experiences.

One person describes her pain as sharp and ripping. The other person disagrees, saying: “I don’t think it is ripping at all, not really sharp either. It’s more like crushing and suffocating.”

Maybe we always have to leave some room for the possibility that, despite all clues to the contrary, the subjective experiences of the two people are, after all, different. But even if we assume that the experiences are similar and the subjects are just giving mutually incompatible descriptions of the same pain, we can interpret this as a case of faultless disagreement.

We can suggest that what the subjects are doing is not that they observe by introspection features of their inner experiences accurately or inaccurately. Instead, they are making spontaneous applications of concepts, and in doing this they engage in *defining* what their experiences are like. They are flagging a certain description as the correct thing to say about their experience. First-person authority, according to this view, is a matter of being in the unique position of choosing how experience-vocabulary is to be applied to one’s subjective experience. What ultimately makes it the case that a subject’s pain is ripping is the fact that the subject *judges* it to be ripping. Even if there is another, incompatible description of a qualitatively identical experience — even if the description of the first subject is highly anomalous — there is no need to ascribe error to any party. The deviant description can be treated just as a different application of experience-vocabulary, an application that is within the subject’s rationality to make, and which may be psychologically interesting in itself. It does not force us to waive the first-person authority of any speaker involved. First-person

authority is the acknowledgement that subjects' statements about their experiences are (treated as) true in their own conversational context.

Constitutivism, in the case of describing my pain, would be friendly to first-person authority by holding that my sincere testimony is the primary court of appeal which determines what my pains are like. The fact that I judge my pain to be ripping plays a constitutive role in making it the case that my pain *is* (rightly characterized as) ripping. First-person authority exists, according to this view, because the primary way of establishing the character of someone's experiences is to refer to that person's sincere avowals about those experiences. For that reason, my judgment to the effect that I have a ripping pain is essential in making it the case that my pain is indeed ripping, as opposed to crushing or suffocating. Of course, there will be constraints on how I can describe the pain; I cannot normally characterize my pain as "dark green" or "prestigious," for example. But it can be argued that this would not be because those descriptions are erroneous in light of some independent standard, but because they violate some conversational maxims; I would normally know that those words are probably uninformative to others as pain-descriptions. Insofar as I want to communicate, I should not use unhelpful concepts, but otherwise I am free to describe my pain in whatever way seems to me most suitable. In determining what is true to say about my experiences, those avowals of mine will be the primary point of reference. First-person authority just reflects this state of affairs.

Is constitutivism preferable to detectivism? Two points of criticism are important. First, it cannot really be praised as an account of self-knowledge. Instead, it makes it hard to characterize my pain-descriptions and other avowals as instances of (self-)knowledge at all. Knowledge conceptually requires some kind of systematic avoidance of error. Roughly speaking, if something counts as an instance of knowledge, it should involve a judgment that succeeds in representing some state of affairs *correctly*, in virtue of some laudable systematic method. If constitutivism generally holds, and truths about persons' inner states are primarily determined by referring to their avowals, then there will be no such thing as the cognitive achievement of getting a psychological self-description *right*. It will be no more of a cognitive achievement than launching an arrow into a wall and drawing a bulls-eye around its head is an archery achievement.<sup>2</sup>

Second, constitutivism seems to make us responsible for mental facts about ourselves in a way that is not plausible across the board. Here it becomes evident why constitutivism fits better together with accounts of beliefs and other similar attitudes. When we consider the latter, constitutivism seems advantageous, because we generally want to be personally responsible for the contents of our

---

<sup>2</sup>I believe that something like this thought is behind those remarks of Wittgenstein that suggest a "non-cognitive thesis of avowals," as Hacker (1975) calls it.

beliefs and desires. But sensations are different in this regard. According to constitutivism, what ultimately makes it right to say that my pain is ripping instead of crushing is the fact that I judge it to be ripping instead of crushing. But in many cases, I will be unable to accept this account from my own viewpoint. It will at least usually, if not always, strike me as false to say that my pain is ripping because I judged it to be ripping. In a typical situation, I say my pain is ripping because my pain calls for exactly that word, and I will be inclined to insist that I really have no rational control over that matter. If I complain of a sharp pain, no one can seriously suggest to me: “Learn to *judge* it to be dull instead, and then it will not be sharp anymore!” Not all conscious experiences, as they appear to me in first person, leave room to intellectually decide the most appropriate verbal characterization for them. Some experiences do not let me rationally judge what I want to say of them; they will rather take control of me, and demand an expression. This uneasiness from the first-person viewpoint should justify looking for a better account to surpass both detectivism and constitutivism.

### Expressivism

Finkelstein (2003), Bar-On and Long (2001), Bar-On (2004), and Rodríguez (2012) have examined expressivism as a superior alternative for making sense of our relation to our own inner sphere. This view develops a point inherited from Wittgenstein (1953), saying that much of psychological talk in the first person is not descriptive in nature; it does not stem from an observation of an inner object. Instead of merely rejecting detectivism, however, Wittgenstein insisted on continuity or at least a possible connection between verbal avowals and primitive, “natural” expressions:

How do words refer to sensations? [. . .] The question is the same as: how does a human being learn the meaning of the names of sensations? — of the word “pain” for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behavior. “So you are saying that the word ‘pain’ really means crying?” — On the contrary: the verbal expression of pain replaces crying and does not describe it. (Wittgenstein 1953, §244)

According to the possibility Wittgenstein points out, the avowals that we use to talk about our experiences work in the same way as pre-verbal grunts and cries. The point of the avowals is not to be parts of fact-stating discourse, but to give voice to wants and needs in social interaction. The avowals can also be drawn out of me against my will, like primitive expressions. This is a point in favor of expressivism against constitutivism, as the latter threatened to over-intellectualize the subjective sphere.



For the question of why my descriptions of my own experience carry a special authority, expressivism offers a deflationary answer. According to it, avowing is not a matter of describing one's pains or feelings at all. Avowals only superficially look like descriptions. Actually they are sophisticated and cultured expressive behavior: utterances that are in the business of reacting to my surroundings, and thereby doing other things, such as eliciting pity or asking for help. This was Gilbert Ryle's view in his *Concept of Mind*:

[M]any unstudied utterances embody explicit interest phrases, or what I have elsewhere been calling "avowals," like "I want," "I hope," "I intend," "I dislike," "I am depressed," "I wonder," "I guess," and "I feel hungry"; and their grammar makes it tempting to misconstrue all the sentences in which they occur as self-descriptions. But in its primary employment "I want..." is not used to convey information, but to make a request or demand. [...] Nor, in their primary employment, are "I hate..." or "I intend..." used for the purpose of telling the hearer facts about the speaker; or else we should not be surprised to hear them uttered in the cool, informative tones of voice in which we say "he hates..." and "they intend..." We expect them, on the contrary, to be spoken in a revolted and a resolute tone of voice respectively. (Ryle 1949, pp. 183–184)

However, even if Ryle's view of the primary employment of avowals is correct, he realizes that he cannot boldly generalize this point. The existence of a "primary" employment implies that there are one or more secondary employments. Surprising or not, sometimes "I hate..." and "I intend..." are uttered in a cool and measured manner, in order to give a self-description. The view that avowals are simply expressive and lack truth-values is rightly met with suspicion (Hacker 1975; see also Malcolm, 1954). Obviously, if this simple view is what expressivism amounts to, it will explain (apparent) first-person authority, but it will not be an account of self-knowledge. According to it, my verbal avowals are no more instances of self-knowledge than distinctive grunts and gestures are. On the other hand, those avowals cannot be meaningfully corrected by another person, but this is for the trivial reason that they have no factual content to disagree on.

Wittgenstein (1953, II, ix) plausibly acknowledged that avowals can play the role of both expressions and descriptions, or something in between. A non-naive version of expressivism holds that my speech about my own mental states is fundamentally continuous with my natural bodily expressions, but such speech still linguistically expresses or "manifests" facts about my thoughts and feelings so that it is capable of stating truths or falsehoods about me. Bar-On (2004) has developed such a version and labeled it "neo-expressivism." Sophistication is clearly necessary, because it is hard to deny that avowals are in some sense also in the business of stating facts about their speaker. Avowals have contents that can feature in logical inferences, they can be contradicted by other statements, and so on. It seems that expressivism has to face an objection that is parallel to the Frege–Geach problem for metaethical non-cognitivism (for a summary, see Sinclair, 2009): How can this way of talking be fundamentally expressive, when it evidently in many contexts functions like descriptive, fact-stating talk?

In what follows, I will present a development of expressivism to shed light on the nature of avowals, the first-person authority associated with them, and the limitations of that authority. I attempt to combine a number of what I take to be valid insights. First, I will endorse a view that I attribute to Wittgenstein: avowals can function as expressive utterances but also as descriptions, and there is no categorical line separating the two cases. Second, I agree with Rodríguez (2012) in holding that Bar-On's (2004) influential expressivist account has the undesirable feature of taking apart avowals as expressive *acts* and avowals as the linguistic (truth-evaluable) *products* of those acts. I suggest that the putting forward of a linguistic description of one's experience is a single expressive act, whose expressive quality and truth-value are assessed in an interdependent fashion. Third, I seek to integrate detectivism and constitutivism in the picture, by highlighting the kinds of cases where each works best.

### Primitive Avowals, Intellectual Self-Ascriptions, and Deliberations

For heuristic purposes, I will distinguish between three different types of psychological self-ascription. These are not meant as rigid categories. Instead, they represent the end and middle points of a scale on which avowals, and interpretations of avowals, can move. One extreme is a purely expressive, spontaneous avowal; another extreme is a detached, cool self-ascription done as if from a third-person perspective. Between these, there is a vast range of avowals that express the speaker's state of mind *by* asserting something about it. A good label for these latter cases is hard to come by; I will call them *deliberations*, owing the word, and some of my inspiration, to Moran (2001).

Other advocates of the expressivist view have made the point that (some) avowals have a special epistemic authority because of their peculiar expressiveness. They are taken to be immediate, non-judgment-involving airings of the subject's mental states. My aim is to qualify, and clarify, this point by suggesting that some avowals (deliberations) have a special epistemic authority when they are expressive in a certain spontaneous and unstudied way while *also* being honest attempts of a revisable self-description.

#### *Primitive Avowals*

First, I endorse Wittgenstein's point about verbal expressions of feelings being able to take over and extend the function of primitive, non-verbal expressions. Assuming that more articulate and considered expressions can build on simple primitive expressions, I will propose a way of seeing these as a procession on a single, continuous scale. Primitive, natural expressions like cries and smiles are devoid of cognitive content. They are not attempts to convey factual information. They may be expressions of attitudes, means of drawing attention, devices of eliciting reactions from others or otherwise communicative, but they are not

statements or descriptions of the subject's mentality. They can be called purely expressive acts. The simplest form of verbal avowals can be equated with them. Cases where "It hurts!" is used spontaneously and passionately to serve the same function as would be served by a scream, or a case where a spontaneous "I feel so good!" takes the same communicative role that could be taken by an exhilarated smile, can be called purely expressive avowals. These have a character of naturalness and spontaneity; they are drawn out of a person, rather than formulated and put forward by the subject in a controlled fashion. This is one end of my proposed spectrum.

### *Intellectual Self-ascriptions*

On the other end of the spectrum, there are self-ascriptions of mental states that are purely descriptive. Whereas purely expressive avowals are not descriptive to any extent, the self-ascriptions of mental states at the other end of the spectrum are not expressive to any extent. The latter are instances where the subject takes a detached, third-person perspective toward her own mentality, and produces a studied verdict from that perspective. She may or may not like the contents of that verdict; she may even want to disown it. I will call these intellectual self-ascriptions. They will include a case where I reluctantly admit, after a lengthy work to sort out my thoughts, that I am angry with my father because of his strictness as a disciplinarian, while at the same time admitting that I should not and do not want to be angry with him. In another case, I notice my slowing pace of work and carelessness and conclude that I must be tired and frustrated, although I do not feel like saying that I am either of those things; but my physical and behavioral condition force me to make that conclusion anyway. I know, after all, that lethargy and carelessness are objective criteria for a person's being tired.

At this latter end of the spectrum, it can be legitimately said that I come to know my own mental states by *detecting* them in myself, although that detection is not necessarily carried out by inward glances of introspection. In any case, in these instances I attribute a mental state to myself as a result of self-observation of some kind, and this observation has no special claim of authority over anyone else's word. My self-observation can be mistaken for the same mundane reasons as any observation can be mistaken; it will make perfect sense to ask me to do my self-observation more carefully or more attentively, in order to avoid error. It is possible that I mistake the symptoms of a medical condition for symptoms of tiredness, or that I misidentify as repressed anger something that further reflective work reveals to be some other complex feeling. In short, this is a class of cases where I am sufficiently alienated from my own mental state to treat that mental state as an external object of scrutiny. The account of detectivism, while not easily generalizable, fits well here. This kind of self-scrutiny was what

Ryle (1949), who rejected privileged introspective access as the basis of self-knowledge, eventually treated as the paradigm case of real self-knowledge.

### *Deliberations*

There is a purely expressive case of avowal; these I have called primitive avowals. There is also a purely descriptive case of avowal (according to my liberal use of the term); these I have called intellectual self-ascriptions. Now I will distinguish a third case, which is *the speech act whereby the subject puts forward an expressive linguistic utterance to serve as a self-description*. I believe that many, maybe most, avowals in typical human communication can be seen as instances of this type. They are characterized by a desire of the speaker to strike a balance between saying something that can be taken to be an objectively accurate description of her, and voicing her own impulses and wants, all in a single speech act. They are expressive utterances of the subject, but these expressive utterances acknowledge that they are attempts at manifesting a mental event that is an object of scrutiny also from the subject's own perspective; an event for which the giving of an adequate description is a cognitive challenge. They are instances where the subject assesses two things at once: first, what she wants to say about her experience; and second, how objectively plausible her statement is as a self-description. I call these avowals *deliberations*. One more modified example will serve to illustrate the point.

### *Example 3*

I have an abdominal pain that I need to describe to a physician. I am asked to assess my pain's intensity on a scale from 1 to 10. I have used the pain-scale before, and I consider the guidelines I associate with different numeric degrees of pain. I judge that my pain is of the level of 7. Then I am asked to think carefully:

"You describe your pain otherwise in the same way as in those earlier instances when you have judged it to be 5 or 6. You also don't show signs of greater distress over it. Are you sure that 7 is not too much?"

I answer: "Yes, I understand that, but I just feel that this week it is harder to bear. I'm not sure if it is the pain itself that intensifies or if I am just depressed, but 6 would be too small a number now. I'm saying 7."

Here, I am doing several things at once. First, I am giving a description of my pain. My utterance of "7" occurs in a context of giving a description; it is meant to inform the other about a certain feature of my conscious experience, to go down in my medical record as a true proposition about my condition. Second, I am using words (or rather, numbers) expressively: the point of my saying "7" is to let the other know how I feel about my pain, to voice my sentiment. Third, however, in this particular example I acknowledge that I have

some reservations about whether my avowal accurately describes a change in the pain itself or in the overall quality of my mental condition (“I’m not sure if it is the pain itself that intensifies or if I am just depressed”). Here I admit that my decision to say 7 instead of 6 might be borne out of my growing concern over my pain, my overall feeling bad physically and mentally, or something like that. In a way, I give the hearer some freedom to evaluate what conclusions to draw from my utterance.

Now, it seems to me plausible to agree with constitutivism to an extent. My description of my pain has a unique claim to being true. This is because my avowal has a special status in determining what is deemed right to say about my pain’s intensity. My honest avowal of my pain as 7 is a central criterion for it being the case that my pain indeed has the intensity of 7. I am the only one who can apply pain-vocabulary to myself in the first person, so my judgments about my pain are crucial in determining how pain-vocabulary is to be applied to me in particular cases. However, my avowals are not the only criterion for determining what my pain is like; there are bodily and behavioral criteria for different kinds of pains too (as the interlocutor in Example 3 notices).

In light of this, I suggest that my avowal is a complex communicative act: it is, in effect, *a request for others to accept my pain-avowal as a valid description of my pain*. It has a double nature. It is put forward as a description of my state of consciousness, but it is also a kind of an act of pleading: an expression of my want to make others treat my pain as a pain of the level 7. Most of the time, my avowal will be accepted as a valid and authoritative description without a scruple, insofar as people generally accept the first-person authority of subjects over their own mental states. But sometimes there will be room for scrutiny, as in Example 3.

In Example 3, I am saying that my pain is 7 in circumstances where, as far as any onlooker can see, I could as well say 5 or 6. So why am I saying 7? If this unusual question would be put to me, I could approach it in a number of ways.

(a) First, I could try to ground my judgment in some objectively available behavioral evidence. “Look, I may not show signs of greater distress over my pain just now, but there are some signs anyway: it distracts me more than before, it is harder to concentrate on anything else, I am constantly more stressed about it than before . . . . It must have intensified from 6 to 7.” Here, I am taking a more detached position toward my pain by allowing that it is a matter of evidence to decide whether my pain is 6 or 7. This is to move my avowal more in the direction of what I have called an intellectual self-ascription. I loosen my claim to first-person authority somewhat, by allowing that my judgment about my pain might be wrong according to some standards that can override my own statement.

(b) Second, I could (in principle) decide to be a hard-headed constitutivist. “I just feel like saying 7. It seems to me to be the correct application of the

pain-scale to what I am feeling right now. And I am automatically right in this, because it is me who gets to decide how pain-vocabulary is applied to my inner experiences. End of story." I think it is evident that by these words, the subject would make her avowal sound in a certain way suspicious. It seems that her attitude toward her pain is not an attitude of a person who wants to communicate something about her pain to others. It is rather the attitude of a person who is merely interested, for one reason or another, to ensure that the hearers withhold further inquiry and accept her statement. Concerns about the honesty of the avowal would be raised, and there would be some hesitation about whether, or to what extent, her utterance can be taken seriously as an avowal. It would be sensible to protest that the subject does not get to *decide* whether his pain is 6 or 7 just like that.

(c) Finally, there is the option that seems to be natural and plausible: "I just feel that I have to say 7. I cannot help it. It just feels worse today." What I acknowledge here is that my avowal shares the nature of a primitive expression: the number 7 is drawn out from me, somewhat in the way spontaneous grunts or smiles are drawn out from me, rather than rationally decided to be my chosen number for the pain.

Now, I suggest that the first-person authority of my self-ascription is at its strongest when it has a nature like that described in (c). When it is in this way akin to a spontaneous, primitive expression, then the subject has a special force behind her request that her pain-avowal is treated as a valid description of her pain. Her avowal will then represent her genuinely best effort to give a linguistic expression to an event of her consciousness that does not allow for just any arbitrary expression.

In other words, I am suggesting that the authority of an avowal as a self-description is dependent on whether the avowal is taken to share in the nature of a primitive avowal. But I am also arguing that it is necessary for an avowal to be plausible as a description from a detached perspective, if it is to work in its role as an avowal. Once more, I will illustrate by an example. Let us imagine that, in Example 3, I am struck with a sudden fear and anguish over my pain, and I start to feel my constant, familiar pain as so unbearable that I want help with it immediately at any cost. Then, when asked about the intensity of my pain, I will respond "10." Now, it seems that another person would have a good reason to say to me: "Look, you cannot really say that. I know you feel bad, but 10 is the highest point of the scale, it is meant to represent a pain that is so unmanageable that you have never experienced anything worse than it. A person with a pain that has the intensity of 10 would be incapacitated, which you clearly are not." In a way, my "10" would be a failed avowal; it could not be taken seriously as an avowal.

In the previous case, I am uttering "10" as a kind of a purely expressive call for help that does not even purport to be a measured attempt of self-description.

This kind of an avowal will be appropriate in some conversational contexts, but faulty in many others. In particular, it will be unhelpful for the physician, or at least it will put the physician in a position where she has to contemplate how to interpret my utterance. It will not be a fully functional avowal in its context.

I take these considerations to show the following. Insofar as my (deliberative) avowals are descriptions of my mental state, they are also requests for others to accept my description as valid. But the acceptability of my avowal as a valid self-description is largely dependent on whether my avowal is taken to be expressive in the right way (i.e., in the way of a primitive, unstudied expression). And my avowal, however honestly expressive, will not be fully taken seriously as an avowal unless it at least attempts to be a descriptive act (i.e., is constrained by my aspiration to inform others about what my pain is like, and not only by what I want to say about it). The descriptive and expressive aspects of an avowal are interdependent.

### **Conclusion: Avowals, Self-Knowledge, and the Nature of First-Person Authority**

I will now conclude by spelling out some consequences for the issues of self-knowledge and first-person authority that can be drawn from my discussion. First, it seems to me that a crucial part of what is commonly called “self-knowledge” is manifested in a person’s ability to reflect on her use of the different modes of avowals, and to some extent choose between them. Avowals are called for in many different communicative situations. Sometimes, when another person asks me “How do you feel?,” what is expected from me is just a spontaneous manifestation of my feeling of pain, affection, or anxiety. Then, it is an exercise of self-knowledge to be able to recognize and let out my spontaneous and unstudied reaction, suppressing any need to take a detached perspective and survey my state of mind as a part of my objective personal psychology. At other times, it will be necessary for me to study my psychology as if from a third-person perspective, in order to uncover biases or unconscious motivations, acknowledging that my own assessment of my psychology is nothing but an assessment by a fallible human being. Then, it will be necessary to contain my spontaneous and unstudied reactions, and to keep in mind the possibility that my first thoughts about my pains, affections or anxieties might not be the (whole) truth about them. (“I feel like saying that this pain is 7; but don’t I usually have a low pain threshold? Maybe most other people would call it 6, or even 5? And I admit that I am feeling depressed; maybe that is affecting all my judgments more than I realize.”) Understanding that my unstudied expressions and correct descriptions of my psychology (according to some standards that I myself can accept when speaking in third person) can come apart, and finding out how they can be expected to come apart in diverse situations, is a vital part of my

self-knowledge. In deliberations, I talk expressively, and in so doing I manifest my wants and needs to characterize my mental life in certain ways, but at the same time I am subjecting my avowals to interpersonal assessment by presenting them as *descriptions* of myself. Seeing how those expressively grounded descriptions manage with and against those descriptions of me that are given from the perspective of another person helps me to cultivate an important kind of self-knowledge. I am learning how my conception of myself plays together with other people's conception of me.

This characterizes self-knowledge in a sense in which it is a process. It is a sense of self-knowledge that is easily overlooked if the crucial expressive function of first-person psychological talk goes unnoticed. In deliberations, how I can plausibly describe myself constrains how I should feel appropriate to express myself, and at the same time how I need to express myself constrains how I describe myself. Competent use of deliberative avowals might be characterized as communication that is at the same time both self-studying and self-defining — a remarkable feat of human thought.

Second, pointing out the combination of expressiveness and descriptiveness in avowals produces a modest and commonsensical view of first-person authority and its limitations. There is little motivation to assume that individuals have magically accurate introspective powers, so that they would be uniquely authoritative judges of their own mental states in a detectivist manner. But what people do have is a subject's perspective to those mental states, and a desire to define and characterize those states from that perspective. Conscious attempts of persons to work out what their subjective experiences are like — what I have called deliberative avowals — have a special epistemic status insofar as they are properly expressive honest utterances while also being attempts of self-description. A description that I give of my own experience is authoritative when, and insofar, it is based on an expressive act that is ungrounded and natural in the same way as a primitive bodily reaction is. The subject is the only one who is in a position to give a description with this peculiar basis; therefore, naturally, an avowal of this kind carries special weight. When moving away from deliberations toward primitive avowals, or toward intellectual self-ascriptions, motivation to demand a special authority for the avowals wanes: in the case of primitive avowals, because they are not issued or interpreted as statements with factual content, and in the case of intellectual self-ascriptions, because they are not made from the special perspective of the subject-position.

First-person authority is, first and foremost, recognition that each person has a unique status as a generator of knowledge about her own mental reality. Properly expressive deliberative avowals have a special epistemic job to do. They are not infallible, not always even highly reliable, but they are acts of giving voice to a personal experience: they are the subject's applications of concepts to her personal experiences in a certain situation and at a certain time, and as



such they have a constitutive role. They serve as the starting point of inquiry into her experience, and enjoy a certain amount of resistance to corrections. The role of such avowals as (partly) self-defining acts also means that a subject can, in principle, decide to stick to her self-description even when it is anomalous from the perspective of an outside observer. If a subject is truly brought to see her self-description as erroneous, this must happen by eventually bringing her to revise her avowal in such a way that she can, after the revision, own it as her honest self-expression, not only as a third-person description of her forcibly given from outside. This seems an essential characteristic of an autonomous, self-standing subject. Consequently, first-person authority has an ethical dimension in addition to an epistemic one. Respecting it is to grant to other people an authoritative voice in telling what their experiences are like. Disregarding it is to say that it is in principle possible to overrule a subject's self-expressing voice by a third-person, more authoritative account of what her experiences are *really* like. It is doubtful whether those who are subjected to the latter treatment have a chance of seeing themselves as subjects in the full sense.

### References

- Bar-On, D. (2004). *Speaking my mind: Expression and self-knowledge*. Oxford: Oxford University Press.
- Bar-On, D., and Long, D.C. (2001). Avowals and first-person privilege. *Philosophy and Phenomenological Research*, 62, 311–335.
- Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.
- Finkelstein, D.H. (2003). *Expression and the inner*. Cambridge, Massachusetts: Harvard University Press.
- Gertler, B. (2011). *Self-knowledge*. London: Routledge.
- Hacker, P.M.S. (1975). *Insight and illusion: Wittgenstein on philosophy and the metaphysics of experience*. London: Oxford University Press.
- Hurlburt, R.T., and Schwitzgebel, E. (2007). *Describing inner experience? Proponent meets skeptic*. Cambridge, Massachusetts: MIT Press.
- Malcolm, N. (1954). Wittgenstein's Philosophical Investigations. *The Philosophical Review*, 63, 530–559.
- Moran, R. (2001). *Authority and estrangement: An essay on self-knowledge*. Princeton, New Jersey: Princeton University Press.
- Rodríguez, Á.G. (2012). How to be an expressivist about avowals today. *Nordic Wittgenstein Review*, 1, 81–101.
- Ryle, G. (1949). *The concept of mind*. New York: Barnes & Noble.
- Schwitzgebel, E. (2011). *Perplexities of consciousness*. Cambridge, Massachusetts: MIT Press.
- Sinclair, N. (2009). Recent work in expressivism. *Analysis*, 69, 136–147.
- Wittgenstein, L. (1953). *Philosophical investigations* (third edition). [G.E.M. Anscombe, Trans.]. Oxford: Blackwell.