

Why avowals must be assertions

(Published in *Philosophical Investigations*—please cite published version where possible)

Ning Fan, University of Tennessee, Knoxville, nfan1@vols.utk.edu

Abstract

In *Philosophical Investigations* §244, Wittgenstein suggests that we understand avowals (first-person psychological utterances) as manifestations or expressions of the speaker's mental states. An interesting philosophical theory, called *expressivism*, then develops from this Wittgensteinian idea. However, neo-expressivists disagree with simple expressivists on whether avowals are at the same time assertions, which are truth-evaluable. In this paper, I pursue the expressivist debate about whether avowals must also be viewed as assertions. I consider and reject three neo-expressivist objections against simple expressivism. Then, I offer my own account of why we should understand avowals as assertions.

Introduction

When a person sincerely uses the first-person pronoun “I” to declare what is going on in her mind, there is a presumption that she is not mistaken. This presumption is, of course, not so strong as to render all first-person declarations about one's own mental states free from errors. In some exceptional cases, a person's self-declarations can be overridden when others' judgments about her mental state are taken into consideration. Another notable feature regarding self-declarations of mental states is that they are *usually* not evidence-based. A normal person, in most cases, speaks about her mind without relying on the observation of her own behaviours. Here “first-person authority” is a standard term used to refer to these two

special features associated with the first-person announcements of mental states.

If we accept that first-person authority is not an illusion, that is, the above-mentioned characteristics do exist, then the question would be: how to account for this phenomenon? or, more specifically, what is it that differentiates the first-person declaration from the third-person judgment about the same mental state such that the former enjoys a certain authority that the latter lacks? In *Philosophical Investigations* §244, Wittgenstein hints at a possible answer to this question:

A child has hurt himself and he cries; then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour.

“So you are saying that the word ‘pain’ really means crying?” — On the contrary: the verbal expression of pain replaces crying, it does not describe it.¹

Wittgenstein suggests that we understand the utterance “I am in pain” as a kind of pain-behaviours just like crying, or wincing. Understood in this way, the utterance “I am in pain” *just* is another way of expressing or manifesting pain. So it makes no sense to say that I can be mistaken in declaring that I am in pain when the declaration itself *is* the manifestation of pain. What follows from this is that a person’s relation to her pain is not a relation of knowing, but of ownership. As Wittgenstein goes on to say in the following section, “I cannot be said to learn of them [my pains]. I *have* them”.² This explains why a person needs no evidence to learn about her pains—she is not learning anything at all!

It seems that the phenomenon of first-person authority can be explained by using Wittgenstein’s idea of equating first-person utterances of mental states with the expression of those states. However, instead of closing the discussion, such an explanation opens another inquiry into the nature of expression. Some philosophers who are interested in the problem of first-person authority turn their attention to questions such as: is the utterance “I am in pain”

¹ Wittgenstein (1953: §244).

² Wittgenstein (1953: §246, emphasis in original).

truth-evaluable? or, can the utterance “I am in pain”, while being a manifestation of a person’s pain, also be an assertion about one’s pain *at the same time*? An interesting philosophical theory, called “expressivism”, thus emerges from the consideration of these questions. Expressivists in general subscribe to the idea that first-person present-tense psychological utterances are expressions or manifestations of the utterer’s mental states, and the term “avowal” is specially used to designate such first-person utterances. However, expressivists disagree with each other on whether avowals can also be assertions, something capable of being true or false. Neo-expressivists, who hold that avowals are at the same time assertions, criticize the simple expressivists, who deny avowals have such double functions, by presenting difficulties for their view. They argue that simple expressivists cannot respond to their objections without accepting the idea that avowals must also be assertions.

In this paper, I pursue the question of whether first-person propositional avowals³ must also be assertions. After clarifying the main difference between simple and neo-expressivisms in the first section, I go on, in the second section, to examine three objections to the simple expressivist view. I show that these objections can be answered without the simple expressivists giving in to neo-expressivism; thus, they are not sufficient to motivate the departure from simple expressivism to neo-expressivism. However, this does not mean that we do not have good reason to hold that avowals must also be viewed as assertions, which is the central thesis of neo-expressivism. In the final section, I offer my own account of why we should understand avowals as assertions. My argument will draw on Richard Moran’s view about the nature of assertion as assumption of responsibility. And, in warding off the potential objection to my argument, I will make use of Wittgenstein’s idea that speaking a language is living a form of life. If my argument is successful, it will show that the reason why we must

³ When a person stubs his toe, his utterances “Ouch!” or “I’m in pain!” can both be regarded as avowals of his pain. But only the latter is propositional, and has the potential to be an assertion. For this reason, my discussion focuses on the avowals whose content is propositional.

not understand avowals as mere expressions is that such a view implies an unhealthy and incorrect interpersonal relationship between speakers and their interlocutors.

1. Simple expressivism and neo-expressivism

Readers might have the impression that simple expressivists and neo-expressivists are two groups of people who propound their own theories and argue against each other. However, this is not the actual development of the debate. The contrast between simple and neo-expressivism is first proposed by Dorit Bar-On,⁴ who proclaims herself a neo-expressivist, and who uses the term “simple expressivist” to refer to the earlier commentators of Wittgenstein in order to differentiate her position from theirs. So, the so-called “simple expressivists”, in fact, never explicitly refer to themselves in this way. For this reason, in discussing the debate between simple and neo-expressivism, it would be helpful to consider what neo-expressivists maintain and what view they attribute to their rivals.

The following passage by David Finkelstein nicely summarizes the main difference between simple and neo-expressivisms:

These readers [simple expressivists] have clung to an assumption that has no foothold in Wittgenstein’s writings—an assumption that might be stated as follows: “I can let another person know the state of mind that I’m in *either* by expressing it *or* by saying something true about it. But I cannot, in a single speech act, *both* express my state of mind and say, truly, that I’m in it.”⁵

According to Finkelstein, neo-expressivism is the view that avowal typically performs dual functions—it not only expresses (in the sense of manifestation) one’s mental state but also asserts that one is in the very state.⁶ Presumably, what simple expressivists maintain is the opposite of the neo-expressivist thesis. So, the simple expressivist view, as suggested by the

⁴ Bar-On (2004).

⁵ Finkelstein (2010: 194, emphasis in original).

⁶ Finkelstein (2003: 101).

above paragraph, should be understood as that no avowal is at the same time expression and assertion.

However, there are two readings that are compatible with the above understanding of simple expressivism. One is that, according to the context of their use, avowals can sometimes be an expression and sometimes be an assertion (but never be both).⁷ The other is that avowals can *only* serve as expressions in any context.⁸ Apparently, the latter interpretation is less flexible and, thus, presumably, more vulnerable to the criticisms from neo-expressivists. So, when I am examining the objections against simple expressivism in the next section, I will consider how to defend the less flexible version of simple expressivism against them. If my defence is successful, then we have reason to believe that simple expressivism as such is immune from those objections.

2. Three unsuccessful objections to simple expressivism

The objections I am going to discuss can be found in the relevant literature, especially the writings of Finkelstein and Bar-On. However, when I am rephrasing these objections, my aim is not primarily concerned with giving an accurate summary of what these authors have said. Rather, I focus on whether some reason can be used as an objection against simple expressivism and whether it is a plausible one. So, even if it turns out that the objections I consider are not the same as the ones originally raised by the philosophers working in the relevant area, I believe these objections are still worth considering as potential objections.⁹

2.1. Objection 1: logical Inference

⁷ Such a view can be attributed to Hacker (1986, 1993); but Hacker's position changes in his later work (Hacker 2006) and becomes closer to neo-expressivism.

⁸ David Rosenthal seems to attribute this view to Wittgenstein. See Rosenthal (1993).

⁹ The third objection I consider is the one that I construct from Finkelstein's writing (2003). Finkelstein doesn't explicitly use it in his argument against simple expressivism.

According to simple expressivism,¹⁰ avowals are not assertions about the utterer's mental states, so they are not something that can be either true or false. But in daily conversations, avowals, like assertions, still carry information. That is, we learn *something* from a person's avowals and use what we learn as a basis to infer or determine what we should do in a particular situation. However, one might object that avowals, as mere expressions, are incapable of being true or false and cannot by themselves enter into the logical inference. Consider the following remark by Bar-On:

Very briefly, if avowals do not involve genuine, truth-evaluable ascriptions of mental states to particular individuals, then they cannot have truth-conditional equivalents with which they can be legitimately interchanged in certain contexts, and they cannot serve as legitimate premises in logical inference.¹¹

To see the force of this objection, let us imagine a situation in which my friend, George, says, "I want an ice cream!", and I thus infer that I should get him an ice cream. Proponents of this objection will claim that, in order for me to infer what I should do about George's desire, I have to consider *not* George's utterance "I want an ice cream" but its counterpart "George wants an ice cream". This is because George's utterance "I want an ice cream" is merely the expression of his desire. So, it is like exclamations, such as "Ouch!", or "Uh?", that do not have truth value. Since the third-person present-tense sentence "George wants an ice cream", unlike the avowal by George, is an assertion or description that is truth-evaluable, I can use it as a premise, along with the conditional "If George wants an ice cream, then I should get him an ice cream", to draw the conclusion that I should get George an ice cream. However, in making this move, it is assumed that George's original expression "I want an ice cream" is truth-conditionally equivalent to "George wants an ice cream". But it is not clear why something

¹⁰ As I mentioned in the previous section, in discussing the objections to simple expressivism, I uniformly understand simple expressivism as the position that avowals can *only* function as expressions.

¹¹ Bar-On (2004: 232).

that is not capable of being true or false can be truth-conditionally equivalent to something that is. Thus, either avowals cannot be used as premises in logical inferences and thus loses their role in conversational exchanges, or they rely on their counterparts to play this role, but it is mysterious why they and their counterparts are truth-conditionally equivalent.

Let me consider this objection in some detail. First, if the classical truth-functional logic is the only logic that we employ in everyday life, then Bar-On's objection would be fatal because it entails that avowals play no role in the inferences we make every day. However, it is not true that the classical truth-functional logic is the only logic that we use. The reason is that the classical logic fails to capture *every* inference that we ordinarily regard as valid and frequently employ in daily life. Consider the following example:

Premise: This T-shirt is black.

Conclusion: This T-shirt is not white.

Let us use the letter "B" as the translation of "This T-shirt is black" and "~W" as that of "This T-shirt is not white". Then the underlying argument form would be

B

~W

Such a form is not a valid argument form according to sentential logic.¹² However, in everyday

¹² Of course, the defenders of the formalist program in logic might argue that this inference is an enthymeme that rests on an implicit premise—namely, that this T-shirt is black if and only if it is not white. In this way, they can say that this argument is still expressible in classical logic. This is a complicated issue, and a satisfactory treatment of it would exceed the competence of this paper. My tentative response to this formalist objection is that their translation changes the form of the original argument and thereby renders it different from the argument that we are talking about in the first place. Thus, in so doing, they still do not change the fact that the original inference "B/ \therefore ~W" is inexpressible in classical logic.

life, we frequently appeal to this kind of inference, and our intuition is that this is a valid argument, even though its validity cannot be captured by the classical logic.

Besides, we also make inferences from behaviours. I infer that my dog is hungry by observing that she is wagging her tail, or infer that my little nephew wants a doll by looking at the direction he crawls. We can also infer that it is a cold day by noticing that many people wear jackets. So, even if avowals are regarded as mere expressive behaviours, this does not prevent them from entering into the inferences we make in daily life. For this reason, I conclude that Bar-On's criticism of simple expressivism is limited because avowals, although not being able to serve as premises in truth-functional logical inferences, can still contribute to many inferences we make every day.

2.2. *Objection 2: disagreement*

The second objection I would like to examine comes from Finkelstein. It questions simple expressivism's ability to explain some phenomena of disagreement between people. Let me begin with a moderately lengthy quotation from Finkelstein:

Imagine the following scenario: my friend Tom is moving from one apartment to another a few blocks away. Rather than hire movers, he has asked everyone he knows to help with loading and unloading a rented truck. I have agreed to participate, but on the day of the move, I phone him and say: "I'm sorry. I've wrenched my back, now it's hurting a lot. I just don't think I can help with the move today." Tom, who knows that I dislike lifting large objects, say: "Oh, please. Your back doesn't hurt; you just don't feel like getting off the couch." I reply: "No, Tom. I'm telling you the truth; it hurts like hell." Given the sort of expressivism that Wittgenstein is widely thought to espouse, when I say, "I'm telling you the truth; it hurts like hell", I seem to be either mistaken or conceptually confused. For even if I am in awful pain, when I say, "it hurts" I cannot, on this view, be saying something true. Moreover, given simple expressivism, Tom and I should not be understood as even *disagreeing*: he has made a factual claim to the effect that I am not experiencing back pain, and I have made no claim at all. But it seems undeniable that we *are* disagreeing. Hence the unattractiveness of simple expressivism.¹³

¹³ Finkelstein (2010: 192, emphasis in original).

There are two criticisms that can be extracted from the above passage. And both of them aim at the same conclusion: accepting simple expressivism would make what happens between Finkelstein and Tom seem inexplicable. The first objection can be summarized as follows: given simple expressivism, it seems odd for someone to say “I’m telling you the truth; it hurts like hell”, when his original utterance “it hurts” is merely expressive and not truth-evaluable. So simple expressivists, in order to preserve their theory, must conclude that the person in this scenario is conceptually confused. But this is a counterintuitive explanation of what happens to that person.

I think Finkelstein’s first criticism is misguided because it overly emphasizes the literal meaning of the utterance “I’m telling you the truth” and doesn’t pay sufficient attention to how it is *used* in that context.¹⁴ When a person says “I’m telling you the truth”, sometimes he does mean that he is telling you something (or, precisely, some proposition) that he takes to be true, but sometimes what he means is simply that he is not feigning or pretending. Simple expressivists can argue that in the imagined scenario above, what Finkelstein really means by his utterance “I’m telling you the truth” is not that his previous utterance “It hurts” is *true*, but rather that he is not faking his back pain, that is, his avowal is sincere. So, the conceptual confusion is only *apparent*.

The second criticism developed from the above quotation is that simple expressivists fail to interpret the situation between Finkelstein and Tom as a disagreement, but our intuition is that they are disagreeing with each other. In reply, I argue that Finkelstein fails to appreciate what a genuine disagreement requires. It should be noted that two people can disagree about X *only if* there is some observable phenomenon related to X, or an objective criterion of X such that they both can refer to and use it to settle their dispute. In this sense, people must stand in

¹⁴ See Wittgenstein (1953: §11, emphasis in original): “Of course, what confuses us is the uniform appearance of words when we hear them in speech, or see them written or in print. For their *use* is not that obvious”.

the epistemic relation to the phenomenon on which they disagree with each other. However, in the case of pain, a person who is in pain does not stand in such a relation to his pain. As Wittgenstein points out, it makes no sense for a person to say that he knows he is in pain (unless it is a joke); thus, it makes no sense to say that a person can disagree with others about whether he knows he is in pain.¹⁵ Besides, there is no objective criterion to identify something as a pain except the person who is in pain says so.¹⁶ So, there is no common standard available to both the experiencer and the observer in order for them to settle their “disagreement” over whether the former is really in pain.

Once we clarify what is required for genuine disagreement to occur, it becomes clear that Tom and Finkelstein in the imagined scenario are not really disagreeing. A more plausible understanding of the situation should be that Tom is unwilling to acknowledge that Finkelstein is sincerely avowing his pain. And this interpretation is compatible with simple expressivism.

2.3. *Objection 3: first-person authority*

Now, I want to consider an objection that can be developed from Finkelstein’s writings, and then show how it can be met by simple expressivism. As I mentioned in the introduction, expressivism in general is inspired by Wittgenstein’s remarks, and one of its attractions is that it can account for the phenomenon of first-person authority. In this sense, the explanation of first-person authority is the desideratum that both simple and neo-expressivists strive for. However, if it turns out that simple expressivists fail to account for how first-person authority is possible, then it would be a strike against their view.

Traditionally, when we talk about first-person authority, we think of someone as having authority over what she *says* of her own mental states. (That is, we focus on a person’s self-

¹⁵ Wittgenstein (1953: §246).

¹⁶ Wittgenstein (1953: §290).

ascriptions of mental states, and such ascriptions are truth-evaluable.) Normally, we would not say that a person has authority over her cry or smile because crying or smiling is not something that can be either true or false. However, if avowals, as simple expressivists understand, are mere expressions, then it is hard to see how simple expressivism can account for first-person authority in the traditional sense. Neo-expressivists, on the other hand, will not face this difficulty because they contend that avowals typically play two functions—expressing and asserting, and that it is the combination of these two functions that makes the explanation of first-person authority possible. In the following passage, Finkelstein gives a nice neo-expressivist account of first-person authority:

But an avowal of happiness typically performs two functions: it expresses the speaker's happiness, and it says something true—that the speaker is happy... This is why—although you should attend both to my facial expressions and to what I say about myself, if you wish to learn my state of mind—we speak of *authority* only in connection with what I say about myself... It is only when we take note of the way in which mental state self-ascriptions are akin to smiles and wincings that first-person authority can come into focus as an unsurprising concomitant of the fact that one of the ways in which a person may express her state of mind is by commenting on it.¹⁷

Finkelstein's account of first-person authority does a great job at accounting for first-person authority conventionally understood. If simple expressivists insist on staying in this game, then I admit that they are unlikely to do any better than what a neo-expressivist like Finkelstein does. But simple expressivists can opt out of this game; that is, they can refuse to accept the traditional understanding of first-person authority, and instead provide their own understanding of what first-person authority is, and an account of how such authority is possible.

Hacker proposes that we should strive for an account of first-person authority that is different from the traditional one and belongs uniquely to avowals:

They [avowals] are authoritative...not because they are assertions of something the agent knows, but because they are *manifestations* of the agent's feeling, thinking or intending

¹⁷ Finkelstein (2003: 101-102, emphasis in original).

whatever he feels, thinks or intends. So their truthfulness normally guarantees their truth. Hence, the ‘authority’ they have is, in certain cases, akin to the evidential authority of expressive non-linguistic behaviour”¹⁸

Although Hacker is widely thought to be a simple expressivist, I suspect his position, as shown in the quoted passage, has leaned towards neo-expressivism. One evidence is that, in saying “their [avowals] truthfulness normally guarantees their truth”, Hacker seems to imply that avowals are something capable of being true or false. However, setting aside the issue of whether or not Hacker is a bona fide simple expressivist, I believe we can extract an interesting notion of first-person authority from his remarks by modifying the aforementioned sentence as “the truthfulness of avowals normally guarantees the truth of *others’* beliefs”. That is, a person’s authority over her own mental states exhibits itself in such a way that other people can reliably learn about what she feels, wants, or thinks by attending to her avowals.¹⁹ And, simple expressivists can argue, what makes this possible is that her avowal, as an expression of her mental state, is good evidence for people to believe that she is in that particular state. A different view about first-person authority can be formulated as follows:

A person has authority over her mental state, M, in that her avowal of M can serve as a ground²⁰ for others’ believing, via attending to her avowal of M, that she is in M.

I do not intend this account as a rival account to Finkelstein’s. Rather, I mention it only to show that simple expressivism, despite failing to account for how a person can have authority over what she *says* of her mental states, can still accommodate some other aspects of first-

¹⁸ Hacker (2006: 246, emphasis in original).

¹⁹ Finkelstein also mentions this aspect of first-person authority, but his position is different from the one I develop here since Finkelstein thinks our knowledge about a person’s psychology gained in this way is not always inferential or evidential. See Finkelstein (2003: 101).

²⁰ Ground is different from reason, something capable of true or false. Sensory experience as of P can serve as a ground for believing the proposition that P but not as a reason for believing that P. For more about the relation between reason and ground, see Audi (2020: 72-86).

person authority (for example, the authority of being evidence for other beliefs). Thus, we can understand simple expressivism as a less ambitious theory regarding the explanation of first-person authority but we cannot thereby dismiss it on the ground that it is incapable of such explanation.

3. Why avowals must be assertions

That first-person psychological utterances—avowals—typically express the utterer’s state of mind is accepted by both simple and neo-expressivists. The disagreement between these two theorists consists in whether avowals are also assertions. In the last section, three objections against simple expressivism have been shown to be unsuccessful because they can be met without invoking the neo-expressivist thesis, namely, that a person, when avowing, not only expresses but also asserts what’s going on in her mind. More specifically, those objections are unsuccessful because they fail to provide a satisfactory answer to the question of why avowals *must* be viewed as assertions. If we don’t necessarily have to regard avowals as assertions, then simple expressivists can safely maintain their position.

In this section, I provide a new argument to support the claim that avowals must also be regarded as assertions. This argument relies on what assertions can do beyond the communication of information. I believe only when we look outside the communicative function can we really appreciate the substantial difference between expression and assertion, and use it as a basis to build an argument against the simple expressivist understanding of avowal.

3.1. Assertion and responsibility

In discussing the epistemology of testimony, Richard Moran offers an interesting account of what a person does when she asserts something to her interlocutor. This account is commonly

known as the *assurance* view about assertion. Let us begin by considering the following remark by Moran and see how it can contribute to our thinking about the non-communicative function of assertion.

[T]he speaker, in presenting her utterance as an *assertion*, one with the force of *telling* the audience something, presents herself as *accountable* for the truth of what she says, and in doing so offers a kind of guarantee for this truth.²¹

From the speaker's perspective, what a person does with her assertion that P is not simply to communicate her belief that P, but also, by the very assertion, present herself as accountable for the truth of P. So understood, what is being offered by the speaker is not just a piece of information but also a guarantee from the speaker herself about the truth of that information. From the audience's perspective, the audience gains a reason to believe what the speaker says "because he sees the speaker as presenting herself as accountable for the truth of P, and asking...that this offer of assurance be accepted".²²

Many objections to the assurance view focus on the relation between testimonial justification and the speaker's assumption of responsibility (or offer of assurance). One standard objection goes like this: "Suppose the speaker is an epistemically unreliable agent. Then it is implausible to think that the assurance she offers by asserting or testifying that P counts as a reason for the audience to believe that P."²³ I admit this is a powerful objection against the assurance view, but I don't think we need to worry about it here. One reason is that the assurance view, presumably, is a view about all assertions, whose content can be about what happens in the external world or about the mental states of the speaker. However, since I am concerned with the assertions that are themselves expressions of the mental states, their content is exclusively concerned with the asserter's mental states. And given Wittgensteinian

²¹ Moran (2018: 51, emphasis in original).

²² Moran (2018: 61, emphasis in original).

²³ For a brief version of this objection, see Lackey (2011: 322).

idea that a person's relation to her own mental states is not the relation of knowing,²⁴ it makes no sense to say that a person is unreliable in *knowing* (and thus testifying) her mental states. For this reason, we can reject the hypothesis that a person can be an *epistemically* unreliable agent with regard to what's going on in her mind, and the objection cannot get off the ground.²⁵

Of course, more arguments are needed to show how the speaker's offer of assurance can enhance the testimonial justification an audience has. But this is not the concern of my project. I am interested in the normative relation that is highlighted by the assurance view. And the objection considered here does not challenge the possibility of understanding assertion as assumption of responsibility. So, even if the assurance view is vulnerable to some similar objections regarding whether the speaker's assumption of responsibility contributes to the audience's testimonial justification, I should not worry about them because I only focus on the former idea itself.

To see why it is important to regard the speaker as being responsible for what she says, we have to first consider the assignment of responsibility in communication through avowals under the simple expressivist picture. For simple expressivists, communication through avowals typically goes like this: the speaker does her best to avow (in the sense of manifesting her mental states), and the interlocutor uses her avowal as a basis to infer what is going on in her mind. In this case, avowals functions like evidence; what a speaker has to do is to generate evidence—that is, manifest her mental states—for her audience. But so understood, the speaker is not responsible for whether the audience correctly understands her, insofar as her manifestation serves as evidence for her mental state.²⁶ On the other hand, it is the audience

²⁴ Wittgenstein (1953: §246).

²⁵ Of course, Wittgenstein does not deny the possibility of knowing one's own mental states by observation; see Wittgenstein (1953: §586). Thus, it might be argued that if there is a person who always learns about her mental states by observation, and if she is also notoriously unreliable in doing so, then in this sense we can say that she is an epistemically unreliable agent regarding what is going on in her mind. However, such a scenario has been rejected by Moran as impossible. The reason is that the person in such an extreme self-alienation would have no psychology to be known about in the first place. See Moran (2001: 84).

²⁶ One might think that there is some way that we can legitimately hold a speaker responsible even if she is

who shoulders the responsibility of communication since the audience has to make the correct inference from the evidence given to him. If the speaker is misunderstood, the fault lies on the audience who fails to draw the right conclusion from the evidence. In this scenario, the speaker is free from being blamed, and the audience has no right of complaint if miscommunication occurs.

Now, I argue that this simple expressivist picture of communication through avowals is distorted. It unfairly burdens the audience with the job of interpreting what the speaker expresses (in the sense of manifestation) and fails to respect the rational agency of the speaker by treating her as mere provider of evidence. However, neo-expressivists will not face this problem because they regard avowals as at the same time assertions. With the theoretical resources of assurance view in hand, neo-expressivists can say that by avowing, a speaker does not merely generate behavioural evidence and leaves the burden of communication to her audience. Rather, they can argue, what a speaker does when she avows is that she also assumes the responsibility for the truth of what she avows. In this way, the defect in the simple expressivist picture of communication through avowals can be remedied. Once we understand the speaker's avowals also as assertions by which the speaker takes responsibility for what she avows, we can change the normative relation between the speaker and the audience in communication.

Let me use a concrete example to illustrate why simple expressivists misplace the burden of responsibility:

Suppose my friend, Robert, and I are planning a trip together. Browsing some famous tourist spots, Robert says, "The museum in the downtown is so amazing. I would really love to go there!". Hearing this, I respect his desire and include the museum as one of the

understood as a mere generator of evidence. I will consider and reject this possibility in the next section.

spots we will be visiting. However, when we are actually visiting the museum, Robert shows little interest in the artworks and keeps grumbling that the place is boring. His unexpected behaviours ruin the whole trip. Despite being upset with him, I recall that Robert is an arty person: whenever he says something related to art or music, he seems only to show people that he has good taste in life. So, based on my past interactions with Robert, I should not have interpreted his original utterance as expressing his wanting to go to the museum. I am the one who is responsible for understanding Robert as wanting to visit the museum.

Given the simple expressivist model of communication, it is questionable whether I have the right to complain about Robert's behaviour because it is me who fails to correctly interpret his avowal "I would really love to go there!". Also, Robert cannot be blamed for such avowing, given that his avowal does provide evidence regarding his pretentiousness (and it is me who failed to draw the right conclusion).²⁷ However, this reading of the situation is extremely counterintuitive. Normally, we would think that Robert must be held accountable in this situation. That is, we must view him as someone who is responsible for what he avows. Thus, the simple expressivist view about avowal has to be rejected.

At this point, one might wonder: isn't it to the advantage of the speaker that she doesn't have to assume the responsibility for the truth of what she says and leaves the task of interpreting her utterance to her audience? In the case of Robert, he can avoid being held accountable for the consequences following his avowals. So, what is it that motivates him to

²⁷ In this scenario, Robert can be understood as manifesting his pretentiousness either by making an insincere avowal or by making a spontaneous avowal (i.e., Robert cannot help but saying such things). But as long as we accept the idea that in the communication the job of the speaker is to provide evidence about what is going on in his mind, we have no good reason to blame Robert for so acting because it is me who fail to interpret Robert's action correctly or fail to recognize that he is not being sincere. This line of thought echoes a similar point made by Moran: "If this were all that is involved in telling someone something, then the liar or the verbal misleader really would be able to say that she cannot be blamed because all she did was provide evidence for what her beliefs are and leave her audience to draw his own conclusion from this" (Moran 2018: 89).

assert what he wants—that is, to assume the responsibility for what he wants? I propose that the answer to these questions lies in the fact that a person will want to be treated as a rational agent capable of being responsible for what he avows. Suppose that a month later Robert and I plan another trip together. During the preparation, he tells me that he wants to go to the concert. This time I do not take his word for it but instead interpret his utterance as mere manifestation that he feels he is better than people because of his good taste in music. In this scenario, Robert is clearly disrespected by me because I treat him as if he were a child or a dog who only manifests their inner worlds and leaves evidence for people to interpret them. I am the one who is in charge of interpreting what Robert expresses and taking up the responsibility for the truth of what is going on in his mind. Despite being absolved from potential blame, Robert is not treated by me as an equal. Of course, Robert might not actually know that he is treated by me in this way, and I might not in fact have any condescending attitude when I am interacting with him in this way. But this does not change the fact that he is disrespected by me, because my way of viewing him is not a proper way of viewing a rational agent like myself.

Thus, I conclude that simple expressivism entails an implausible characterization of interpersonal relationships. The seemingly harmless thesis inspired by Wittgenstein that “avowals typically manifest one’s mental states”, when paired with the assumption that “avowals *cannot* also be assertions”, yields the striking result that “the speaker, when avowing, is not being treated as a rational agent”. In order to get out of this predicament, I suggest we accept the thesis that avowals should also be viewed as assertions and understand one of the functions an assertion has as bringing the speaker who asserts into the normative relation she deserves.

3.2. *Interpretation and shared form of life*

My argument relies essentially on Moran’s assurance view of assertion, specifically on his

view that assertion should be seen as the asserter's assuming responsibility and offering assurance for her audience. From this, I derive the idea that treating a speaker as responsible for what she asserts is a way to respect her agency, and argues that this idea further constitutes the reason why a person's avowals must also be regarded as her assertions. Now, I want to close this paper by considering a possible objection to my argument.

Some people might be dissatisfied with the solution I offer to remedy the interpersonal relationship exhibited between Robert and me. They think it is not necessary to regard Robert as making assertions in order to bring him into the ideal normative relation (that is, hold him responsible for what he avows). Rather, it suffices to treat Robert's avowal as generally under his own control and to regard Robert as the kind of speaker who understands what inferences his audience is likely to draw from his expressions or avowals. Consider Robert's original avowal "I would love to go there!". If this utterance is not something that is just blurted out of his mouth, then it seems that we have reason to hold Robert responsible for this avowal given that he should be aware of the significance such uttering can have on others. Now, if this objection holds, my answer to the question of why avowals *must* be assertions would be unsatisfactory because holding a speaker responsible for what he avows does not necessarily involve treating the speaker as making an assertion.²⁸

In reply, I agree that by viewing a person as being generally in control of his avowals and as being aware that others may rely on his avowals in forming plans, we can avoid viewing him as if he were a child who is generally regarded as lacking these abilities. But I object that a person considered in this way is still minimally responsible for the communication, since the burden of successful communication still relies on how his audience interprets his avowal. In the case of Robert, it does not make a difference whether I treat his avowal as coming from a self-controlled agent *if* I continue to view his avowal as a piece of evidence awaiting my

²⁸ I owe this objection to Richard Eldridge.

interpretation. When viewed as a mere source of evidence, Robert's contribution to my understanding of his avowal remains limited and passive. It does not even matter whether Robert is intentionally misleading me when avowing, because in this situation I am still provided with something that has evidential value, and from which I can infer what he is really thinking. Whether Robert is viewed as a self-controlled agent or not, if I still rely primarily on the approach of taking his avowals as behavioural evidence, Robert can never be included in the normative relation pertaining to our communication.

However, assuming that I understand the meaning of Robert's utterance, if I view Robert as making an assertion—that is, as *presenting something as true*, I can straightforwardly understand what Robert intends to tell me by his avowal without further interpretation. Suppose I know that Robert presents the utterance “I want to go to the concert” as true, and suppose that I know what Robert means by “I want to go to the concert”. Then, I can understand what Robert wants to tell me without interpreting him. However, if I treat Robert's utterance “I want to go to the concert” as a piece of behavioural evidence about his mental state, then I am restricted to the method of interpretation when it comes to knowing what Robert is thinking. In this sense, I cannot understand him without relying on my interpretation of him. Nevertheless, if I see Robert as making an assertion when he is avowing, I am provided with something that does not require further interpretation from my end. When viewed in this way, Robert is the one who is fully responsible for what I believe through his avowal, since I no longer depend on my interpretation of what he avows. Therefore, I argue that merely regarding a person as a self-controlled agent who is aware of the implications of his avowals is not enough to hold him responsible for what he avows. If we want to include the speaker in the normative relation exhibited in communication, what we have to do is to regard the speaker as making an assertion, as presenting his utterance as true. Only in this way, we

have a chance of giving up the reliance on our own interpretation when understanding our speaker.

So far, my argument relies on the assumption that the speaker and the audience share the same meaning so that the audience does not have to interpret the speaker in order to understand her assertions. And one might question what reason I have to make such an assumption. Isn't it the case that sometimes we interpret other people's words in order to understand them? If my argument denies this interpretive attitude as the proper attitude towards people, then respecting people's agency seems to be purchased at the price of making them unintelligible.

First of all, I admit that we *sometimes* do adopt the interpretive approach towards our interlocutors in conversations for the purpose of understanding what they mean. Donald Davidson at one point suggests that the method we use in radical interpretation is the same as the method we use in ordinary conversations. Consider this vivid characterization of how to determine what your friend means by "yaw!":

If you see a ketch sailing by and your companion says, 'Look at that handsome yawl', *you may be faced with a problem of interpretation*. One natural possibility is that your friend has mistaken a ketch for a yawl, has formed a false belief. But if his vision is good and his line of sight favourable it is even more plausible that he does not use the word 'yawl' quite as you do...²⁹

This kind of scenario no doubt happens in our life, but I argue that it is only a local phenomenon.³⁰ Most of the time, we can understand our speakers without consciously interpreting their words. This is because, as Wittgenstein suggests, speaking a language "is part

²⁹ Davidson ([1974] 2001: 196, emphasis added).

³⁰ I believe Davidson would agree with me on this point. This is because according to the principle of charity advanced by Davidson, we have to assume that our speaker has a system of beliefs and other attitudes very similar to our own, otherwise the interpretation would not be possible. Here I take this principle to imply that interpretation is possible only against a background of something that needs not be interpreted, and "without a vast common ground, there is no place for disputants to have their quarrel" (Davidson, [1977] 2001: 200) or for a person to question whether his friend means ketch by "yawl".

of an activity, or of a form of life”,³¹ and in most cases we share the same form of life with our speakers. Undeniably, sometimes we need to use the interpretation method to determine how our speakers use a specific sentence or word. But the shared form of life guarantees that we would not have to appeal to interpretation every time we engage in a conversation with our speaker. Consider the primitive language-game that Wittgenstein discusses where A calls out “block”, “pillar”, “slab”, “beam”, and B brings the stone to A as he requests.³² Presumably, in most cases, B responds directly to A’s orders without interpreting them. Of course, our actual life would be much more complex than the life exhibited between these two workers. But the complexity of the actual life does not preclude the possibility that such a complex life, along with the more complex language pertaining to it, can still be *shared* by people.³³ So, it is not implausible to hold that in most circumstances we respond to other people’s utterances without interpretation.³⁴ For this reason, I believe that radical interpretation, the consideration of which no doubt sheds light on the problems associated with human language, still remains a philosophical thought experiment, and that it is extremely rare, if ever, for us to *systematically* treat a person’s utterances as evidence *in order to understand him*.

Here, one might worry if my conclusion implies that in some situations where we have yet to understand the meaning of our speakers, we have no choice but to treat what they say as evidence. And by doing so, we unavoidably disrespect their agency by viewing them as mere providers of evidence who are not responsible for what they say. I would say this worry is unnecessary. It should be noted that in those exceptional situations, we take the evidential view

³¹ Wittgenstein (1953: §23).

³² Wittgenstein (1953: §2).

³³ In *Philosophical Investigations* §199, Wittgenstein suggests, “To follow a rule, to make a report, to give an order, to play a game of chess, are *customs* (usages, institutions)” (Wittgenstein 1953: §199, emphasis in original). If we add that to speak a language is also a custom, then we can agree that speaking a language must be an activity that involves more than one individual.

³⁴ Finkelstein also endorses this view in his discussion of rule-following. He denies that between a rule and its application there always exists a gulf that requires an interpretation to bridge. Therefore, Finkelstein thinks in most circumstances, we can understand what other persons mean without interpretation. See Finkelstein (2003).

towards the utterances of our speakers because we have not yet possessed a more straightforward way of understanding them. As audiences in such situations, we only temporarily shoulder the responsibility of communication but do not mean to permanently exclude our speakers from the normative relation. We treat what they say as evidence *only* to give up this very stance towards them in the future conversations. Once the shared meaning between our speakers and us is established, we then become capable of sharing the normative relation that we should have with our speakers.³⁵ Hence, the talk of disrespecting others' agency makes sense *only* when we have established the shared meaning with them but insist on viewing what they say as evidence awaiting interpretation.³⁶

Now, given that we generally share the same form of life with our speakers, there is no point in systematically adopting an interpretation method towards the utterances of our speakers. Of course, it is still *possible* for a person to widely adopt this approach to understand others' utterances (including their avowals), since *systematically treating people's utterances as evidence* is a kind of language-game human beings *can* have. But what I have shown is that such an approach has no use in daily conversations, so, presumably, nothing would motivate a person to engage in this sort of language-game. Besides, we have reason to avoid entering this sort of language-game: we have to treat people as speaking subjects and thereby respect their rational agency. So, it should become clear that the answer to the question of why avowals must also be viewed as assertions is that we *ought* to treat them as such. If we don't do so, we will always rely on our interpretation to understand other people's mental states and will thus fall into an undesirable interpersonal relationship that no rational agents want to have.³⁷

³⁵ For this reason, I think the account I develop here is not in conflict with Davidson's theory of interpretation. Davidson's account can be viewed as something that makes my account of respecting people's agency possible because it provides a way to establish shared meaning.

³⁶ This statement rules out the possibility that we disrespect a dog by treating her behaviours as evidence, since it is not possible for us to establish shared meaning with her in the first place.

³⁷ I am grateful to professor Richard Eldridge for his encouragement and his constructive criticisms on the earlier draft of this paper.

References

- Audi, R. (2020). *Seeing, Knowing, and Doing: A Perceptualist Account*. New York: Oxford University Press.
- Bar-On, D. (2004). *Speaking My Mind: Expression and Self-Knowledge*. Clarendon Press.
- Davidson, D. (1974). "On the Very Idea of a Conceptual Scheme". Reprinted in Davidson (2001), pp. 183-98.
- Davidson, D. (1977). "The Method of Truth in Metaphysics". Reprinted in Davidson (2001), pp. 199-214.
- Davidson, D. (2001). *Inquiries into Truth and Interpretation* (2nd ed.). New York: Clarendon Press.
- Finkelstein, D. H. (2000). "Wittgenstein on Rules and Platonism". In A. Crary and R. Read (eds.), *The New Wittgenstein*. London: Routledge, pp. 53-73.
- Finkelstein, D. H. (2003). *Expression and the Inner*. Harvard University Press.
- Finkelstein, D. H. (2010). "Expression and Avowal". In K. D. Jolley (ed.), *Wittgenstein: Key Concepts*. Durham, UK: Acumen Press, pp. 185-98.
- Hacker, P. M. S. (1986). *Insight and Illusion: Themes in the Philosophy of Wittgenstein*. Oxford: Oxford University Press.
- Hacker, P. M. S. (1993). *Wittgenstein: Meaning and Mind*. Oxford: Blackwell.
- Hacker, P. M. S. (2006). "Of Knowledge and of Knowing that Someone is in Pain". In Pichler, and Säätelä, S. (eds.), *Wittgenstein the Philosopher and his Works* (2nd ed.). Frankfurt: Ontos Verlag, pp. 244-76.
- Lackey, J. (2011). "Testimonial Knowledge". In S. Bernecker and D. Pritchard (eds.), *The Routledge Companion to Epistemology*. London: Routledge, pp. 316-25.
- Moran, R. (2001). *Authority and Estrangement: An Essay on Self-knowledge*. Princeton

University Press.

Moran, R. (2018). *The Exchange of Words: Speech, Testimony, and Intersubjectivity*. New York: Oxford University Press.

Rosenthal, D. (1993). "Thinking that One Thinks". In M. Davies and G. W. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays*. Oxford: Basil Blackwell, pp. 197-223.

Wittgenstein, L. (1953). *Philosophical Investigations*, trans. G.E.M. Anscombe, P.M.S. Hacker and J. Schulte. Oxford: Wiley-Blackwell.