

Expressivism, Self-Knowledge, and Critical Rationality

I. The Problem of Self-Knowledge: Some Asymmetries Between Self- and Other-Ascriptions

Our first-person present tense ascriptions of contentful mental states (for example, of belief, desire, intentions), and phenomenal states (such as pains and the like) are thought to differ in a number of significant and fundamental ways from our ascriptions of those states to others. For example, when a person ascribes a mental state that *p* to another, she must do so on the evidence provided by the utterances and actions of the other. However, it at least appears that typically she need not do so when ascribing such states to herself. In other words, there is an *immediacy* to a first-person present tense ascription of, for example, a belief (that is, of the form ‘I believe that *p*’) that third-person and past-tense ascriptions (‘He believes that *p*’) lack. In addition, unlike other-ascriptions, self-ascriptions are typically taken to be *groundless*, in the sense that demands that we justify our self-ascriptions, or explain how we know that we are in the mental states we self-ascribe, are generally deemed inappropriate. Furthermore, assuming sincerity on the part of the person, such self-ascriptions (those not ascribed on the basis of behavioural evidence) are highly likely to be correct. This likelihood of correctness is not thought to extend to her ascriptions of similar mental states to others, or to past-tense ascriptions to herself. Thus, it is said, persons appear to possess a level of *authority* with respect to certain of their self-ascriptions that, while it falls short of infallibility, is far greater than that which they enjoy with regard to their attribution to others.

Amongst those who accept that these asymmetries obtain two general explanatory paths have been taken. In more recent times, an *epistemically deflationary* approach has gained some currency. According to this view, the authority and immediacy generally granted to certain kinds of self-ascriptions are not to be explained in terms of any privileged position the subject occupies with respect to the perception of her mental states, nor in any advantage she might enjoy with respect to the amount or quality of evidence she might have for them. Instead, it is based on some other non-epistemic feature of self-ascriptions. Still, for most philosophers the question remains an epistemic one. The task as these philosophers see it is to show how we may incorporate the asymmetries into an account that explains how self-ascriptions express knowledge, that is, count as a form of justified true second-order belief about first-order mental states.

In what follows I will first briefly explain one deflationary explanation of the asymmetries, namely the expressivist account found in Rockney Jacobsen’s “Wittgenstein On Self-Knowledge and Self-Expression” (1996). I shall then look at one opposing trend amongst several philosophers to explain the asymmetries by drawing an essential connection between our capacity for self-knowledge and our status as critical reasoners and rational agents. I shall argue that this approach suffers from a number of significant problems that not only undermines it, but also reinforces the sort of Wittgensteinian account of authoritative self-ascription for which Jacobsen argues.

II. An Expressivist Explanation of Authoritative Self-Ascription

According to Jacobsen's understanding of Wittgensteinian expressivism, the non-evidential basis and reliable truth of authoritative self-ascriptions are explained by the fact that such utterances ascribe the very beliefs they express. In other words, the essential claim is that, perhaps contrary to appearances, utterances of '*p*' and 'I believe (or desire, intend, etc.) that *p*' typically express the same mental state of belief (or desire, intention, etc.) that *p*. However, it remains that, as indicated by their differing truth conditions, they mean different things. That is to say, in the case of a special class of self-ascriptions, meaning and expressive content diverge. The basic argument is this. If my statement of 'I believe that *p*' serves to ascribe to me the belief that *p*, it follows that my *utterance* will be true if and only if I do in fact have that belief. But according to the expressivist thesis I also express the belief that *p*. Consequently, if I am sincere in my utterance of the self-ascription (i.e., I have the belief I express), then it follows that my utterance must be true. And this accounts for why, when I utter sincere self-ascriptions of my mental states, I will always get them right. In uttering 'I believe that Wagner died happy' I ascribe to myself (again, as indicated by its meaning) the very belief that my utterance expresses; assuming I am sincere, I will then have the belief I ascribe to myself. Thus, my sincere self-ascriptions will be true. The additional fact that they are expressions of mental states, and not assertions about them (that is, knowledge claims derived from some sort of cognitive act, for example some form of inner self-observation), explains why we can make them immediately and effortlessly, that is, without appeal to any evidence.

III. Self-Knowledge and the Supervisory Model of Rational Agency

In recent times a number of philosophers (for example, Sydney Shoemaker, Tyler Burge, Akeel Bilgrami) have linked self-knowledge claims (that is, authoritative self-ascriptions of mental states) to the critical rationality and rational agency understood to be essential to the first-person perspective. While their approaches differ in various respects, each argues that (1) self-ascriptions express second-order beliefs about first-order mental states, and that (2) the explanation of the truth of, and warrant for, these beliefs that qualifies them as knowledge is to be found in the requirement for self-knowledge that the possibility of critical rationality demands. Each subscribes to what I shall call the *supervisory model* of rationality. Key to this model is what David Owens calls the idea of "reflective control," that we exercise control over our first-order mental states through normative second-order judgement about the probative force of the reasons we have for those states (Owens, 4). In this way we posit our freedom with respect to those states, and are justifiably held accountable for them.

We see this idea at work in Shoemaker's claim that deliberation on what to believe and do involves agency, and "that the agency involved in deliberation essentially involves self-knowledge" (Shoemaker, 28). As he puts it, we are not merely the passive subjects of our beliefs; through reviewing and deliberating on the quality of the reasons we have for them we decide what it is we ought and shall believe. Burge states the idea in more detail:

As a critical reasoner, one not only reasons, one recognizes one's reason as reasons. One evaluates, checks, weighs, criticizes, supplements one's reasons and reasoning. Clearly, this requires a second-order ability to think about thought contents or

propositions and rational relations among them. ... For reasoning to be critical, it must sometimes involve actual awareness and review of reasons; and such a reviewing standpoint must normally be available. ... [T]o be fully a critical reasoner, one must be able to – and sometimes actually – identify, distinguish, evaluate propositions as asserted, denied, hypothesized or merely considered. (Burge, 246-247)

It is in virtue of this capacity to review and reasonably confirm and correct our own first-order attitudes and reasoning by reference to rational standards that we are epistemically responsible (ibid., 258).

Bilgrami also argues that authoritative self-knowledge is tied to the reflective control we exercise over our mental states. Indeed, for him the very idea of first-person authority is a fundamentally normative one, arising as it does out of the justification required for the practices we engage in associated with our holding one another to account for our mental states. He begins with the fact that we generally hold one another responsible for our own mental states, the responsibility of which is predicated on our having reflective control over them. And this, it is argued, presupposes authoritative, rationally necessary self-knowledge – for as Shoemaker and Burge point out, one could not exercise control over one's states through deliberation on them if one had no idea what they were.

IV. Responsibility, Reflection, and Responsiveness to Reasons

According to the supervisory model, self-knowledge is essential for maintaining rational coherence in one's mental life. Furthermore, given that it is in virtue of our capacity to exercise reflective control over our mental states that we can be held responsible for them, self-knowledge is also essential to rational agency. The proponent of reflective control argues not only that our second-order beliefs about the reasonableness of our first-order states may serve as reasons for those states, but that, as far as our status as rational agents is concerned, they are the primary reasons that "rationally motivate" those states.¹ This is not to say that, on this view, rational belief formation must always involve second-order reflection on the soundness of the reasons for it. A subject's belief that *p* – say, that a mouse has taken up residence in her house – may be based on a first-order awareness of pieces of evidence – a hole chewed in a bag of rice, what appear to be mouse droppings on the shelf – that serve as reasons that motivate and provide sufficient justification for the belief that *p*. However, if she is to be held responsible for her first-order state, the subject must be capable of forming a judgement concerning whether or not it is justified through second-order reflection on the justificatory force of the first-order belief and reasoning that supports it. Her focus is not on whether or not the hole in the bag and droppings were caused by a mouse, but rather whether or not her evidentiary beliefs about these things warrant a belief in its presence. And this higher-order judgement must determine whether or not she holds the belief that a mouse is indeed in the house.

¹ The term 'rationally motivate' is also borrowed from Owens. It is meant to "register the fact that reasons for belief produce belief ... by explaining their product in a way that makes sense of it"(Owens, 17).

What does this involve? Say a subject believes that p for reasons q and r . First of all, if she is to reflect on her belief that p and her reasons for it she must know what that belief and reasons are – she must form true second-order beliefs about them. She then deliberates on the soundness of the first-order belief by examining those beliefs that serve as reasons for it, as well as the reasoning that connects them to it. This includes judging if they themselves are justified, whether any fallacies in reasoning have been committed, and if the evidence represented by those beliefs is sufficient to support the belief they are taken to motivate.² Having successfully applied her knowledge of epistemic norms to her reasons and reasoning, she may either (1) find that everything meets the epistemic mark, upon which case she endorses the belief as one she ought to have and maintains it, or (2) find some fault in her reasoning and judges that she ought not hold the belief, at which point she changes her mind. In this way the subject assumes responsibility for her belief.

Even if taken at face value, this picture faces difficult questions. I shall focus on two. First of all, let us suppose that such second-order judgement is possible. The proponent of reflective control claims that a subject's second-order judgement that her first-order *prima facie* reasons and reasoning in support her first-order belief that p are in order is what ultimately motivates her belief that p . So on this view, what directly motivates a subject's first-order belief that p for which she may be held responsible is not her first-order judgements about the world, but rather her second-order belief that the normative constraints on belief have been met. In other words, in light of her second-order judgement that the belief that p is sound and ought to be believed she decides to believe it. But, it may be asked, can such second-order judgement really play the motivational role envisioned for it?

Owens points out that in order to reflect on the reasonableness of her belief that p , the subject must already have a first-order awareness of the reasons that prompt that belief. In exercising reflective control over her mental states, she engages in second-order judgement the purpose of which is to ensure her reasonability by explicitly acknowledging through that second-order judgement the normative force of the reasons she already has. But what do the subject's higher-order judgements that she has those reasons, and that they suffice for the reasonableness of her belief, add to the motivational equation? How do they exert an independent rational influence on – count as reasons for – her belief? As Owens puts it, “[i]f you already have a non-reflective awareness of the reasons which ought to motivate you, how does the judgement that you ought to be moved by them help to ensure that you are so moved? Such judgements”, he concludes, “look like an idle wheel in our motivational economy...” (Owens, 18). Indeed, this would seem to be the conclusion not only because such judgements merely confirm what is already the case. The mechanism of reflective control is second-order judgement about first-order states. The picture is of a mind turned inward, focused entirely on the rational standing of its own contents. How is it that the product of this inner inquiry – a mental state that refers to the epistemic fitness of other mental states – can serve as the primary reason to hold a first-order belief about the world?

The above questions about the motivational efficacy of second-order judgement about the epistemic standing of one's first-order belief presupposes that such judgements are possible – it is assumed that, through reflection on the reasons that rationally motivate

² There is a threat of an infinite regress of justification here “However, I shall ignore this issue, as it is not the matter, which generates the biggest problem for the view under consideration.

her belief that p , the subject may arrive at a second-order belief that those reasons are (or are not) sufficient for that belief. And this second-order belief is what ultimately motivates the belief that p for which she may be held responsible. However, if a reasonable belief that p is one that is motivated by an awareness of sufficient evidence for that belief, then a problem arises. For, as Owens points out, reflection on strictly evidentiary beliefs that justify the belief that p may not determine whether or not those evidentiary beliefs are sufficient to rationally motivate the belief that p (Owens, 25). We may agree that the formation of a rational belief that p or not- p should be determined by the balance of evidence for or against p . However, what determines what constitutes a sufficient level of evidence cannot be decided by deliberation on evidence alone. Rather, the point at which one judges that evidence to be sufficient for the formation of a belief will be determined partly by the subject's non-reflective sense of non-evidential considerations – for example, of the importance to the subject of the matter in question, or how much of his cognitive resources he is willing to devote to it. The fly in the ointment for the proponent of reflective control is that reflection on such justifying reasons (that is, that play a role in determining the rationality of a belief that p) cannot rationally motivate that belief. One cannot rationally motivate oneself to believe that p by reflecting on one's beliefs that time is running out and that it is important that a decision on whether p gets made.

Now to the second objection, which concerns what at least appears to be the threat of an infinite regress contained within the supervisory model. According to this account, the justification for, and reliable truth of, a subject's second-order beliefs and judgements about the rational standing of her first-order mental states is explained by the regulative role they play in the maintenance of her rationality. To quote Burge again:

if one lacked entitlement to judgements about one's attitudes, there could be no norms of reason governing how one ought to check, weigh, overturn, confirm reasons or reasoning. For if one lacked entitlement to judgements about one's attitudes, one could not be subject to rational norms governing how one ought to alter those attitudes given that one had reflected on them. If reflection provided no reason-endorsed judgements about the attitudes, the rational connection between the attitudes reflected upon and reflection would be broken. So reasons could not apply to how the attitudes should be changed, suspended, or confirmed *on the basis of* reasoning depending on such reflection. But critical reasoning just is reasoning in which norms of reason apply to how attitudes should be affected partly on the basis of reasoning that derives from judgements about one's attitudes. So one must have an epistemic entitlement to one's judgements about one's attitudes. [Furthermore], if reflective judgements were not normally true, reflection could not add to the rational coherence or add a rational component to the reasonability of the whole process. It could not rationally control and guide the attitudes being reflected upon.... (Burge, 249-250)

On the one hand, the justification for and security of self-knowledge is explained by the role it plays in critical rationality, which itself is required to regulate one's first-order mental life. The reflective second-order judgements we arrive at regarding our first-order mental states could only fulfill their regulative role if they counted as knowledge. That is, if the second-order judgement that one ought or ought not believe that p was not itself true and warranted, it could not serve as a reason to form, maintain, or discard the belief that p . So, second-order judgements count as knowledge in virtue of the regulative role

they play. But, we might ask, what explains the fact that our second-order judgements and reasoning about the first-order states we judge ourselves to have are so reliably correct? Burge argues that the rationality of a subject's first-order states is maintained by the supervisory function of second-order judgments. But this presupposes that those second-order judgements are themselves in accord with the norms of reason. But what explains this? That our first-order states are in accord with reason is explained by the supervisory activity of our second-order judgements. But what explains how those second-order judgements are normally sound? That they must be is dictated by the role they are said to play. However, to make this point is not to explain how they remain so. If our second-beliefs were not rational, they could not serve the regulatory role that Burge assigns them; but if their rationality is explained in the same way that the rationality of our first-order beliefs is explained, then a third-order of belief will be needed to regulate our second-order beliefs. But since our third-order beliefs can only perform that regulatory role if they are themselves rational, the supervisory account is set off on an infinite regress. Looked at from another angle, if our second-order judgements do not themselves require a third overseeing level that regulates and ensures their remaining in accord with reason, then they (somehow) remain in accord with reason without any higher-order supervision, and so such supervision cannot be necessary for rationality. So the appeal to higher-order supervisory intervention need not be required to account for the rationality of our first-order states.

V. First-Order Expression and the Rational Adjustment of Mental States

The denial that second-order deliberation and the self-knowledge it presupposes may figure in the rational motivation of our mental states is consistent with the expressivist understanding of the character of self-ascriptions outlined earlier. On this view, we need not engage any second-order cognitive faculty to reliably self-ascribe our mental states; rather, this capacity is explained by first-order linguistic expressive *know-how*. The suggestion is that just as our ability to self-ascribe our mental states is what one might call a first-order accomplishment, so too is our ability to maintain a rationally coherent mental life.

Suppose one believes that p , but that one is confronted with evidence that constitutes a *prima facie* reason for a contrary belief that q . According to the supervisor model of self-knowledge and rationality, the adjustment of beliefs in light of this new evidence would require a host of second-order beliefs, among them beliefs that one has these competing beliefs, beliefs about their inconsistency, and beliefs about what changes in those beliefs would be required to resolve the discrepancy between them (Shoemaker, 33). Once again, instead of being directly responsive to the first-order reasons to believe that q that the evidence presents (by either being convinced of the truth of q and thereby relinquishing the contradictory belief that p , or taking that evidence as reason to engage in further first-order inquiry concerning whether p or q), the subject must elevate his focus onto his own psychological states, as opposed to the state of the world, and through consideration of them alone arrive at a judgement about what first-order belief he ought to hold. Now, as mentioned in the earlier discussion of Burge, one problem here is that this presupposes that the subject's higher-order reasons to which he is responsive are themselves sound, and so the proponent of this view owes an explanation of how this

might be ensured that does not trigger an infinite regress of additional layers of overseeing judgement. A second problem is that, even assuming they were true, whatever (supposed) second-order beliefs I arrive at regarding the content of my first-order beliefs and how they relate to one another would be irrelevant to the determination of whether or not the particular first-order beliefs in questions were true. And if what justifies me in holding a belief must be something relevant to whether the belief is true, then those second-order beliefs ought (also) to be irrelevant to whether I am justified in holding that first-order belief and – to that extent – to its rationality.³

VI. Concluding Remarks

It has been argued that our supposed second-order judgement about the content and rational standing of our first-order states by which we are said to exercise control over those states could not do the job assigned to it. Since such second-order belief could not serve as a reason to hold a first-order state, it would be, as Owens puts it, an idle wheel when it comes to the rational motivation of that state. But neither should we expect that we need such second-order belief playing such a role. For, it has been claimed, our deliberations about what we ought to believe, desire, and intend should be guided by our understanding of the first-order reasons for them. And this, I suggest, is consistent with a non-epistemic expressivist account of authoritative self-ascription that denies that we have the sort of self-knowledge thought to be necessary by the advocate of the supervisory model of rationality.

³ One other related problem is worth mentioning here. Even if we were to accept the supervisory model and its understanding of the role of second-order belief in rationality, that model would still retain a limited attractiveness as an explanation of self-knowledge. While it could account for the authoritative self-ascription of those states such as beliefs and desires that are responsive to reasons, unlike the competing expressivist account it would have nothing to say in explanation of the similar degree and kind of authority we have with respect to self-ascriptions of sensations of pain, emotions, and the like.

Bibliography

- Bilgrami, A. (1999) "Self-Knowledge and Resentment." In C. Wright, B. Smith, C. MacDonald (eds.), Knowing Our Own Minds (pp. 207-241). Oxford: Oxford University Press.
- Burge, Tyler. (1998). "Our Entitlement to Self-knowledge." In P. Ludlow and N Martin (eds.), Externalism and Self-Knowledge (pp. 239-263). Stanford: CSLI Publications. Originally published in Proceedings of the Aristotelian Society, New Series, Vol. 96 (1996), pp. 91-116.
- Jacobsen, R. (1996). "Wittgenstein On Self-Knowledge and Self-Expression." The Philosophical Quarterly. Vol. 46, No. 182, January: 12-30.
- Owens, D. (2000). Rationality Without Freedom. London: Routledge.
- Shoemaker, S. (1996). "On Knowing One's own Mind." In S. Shoemaker, The First-Person Perspective and Other Essays. Cambridge: Cambridge University Press.