Kristina Musholt

Comments on "The Concept Possession Hypothesis of Self-Consciousness" by Stephane Savanah

In his very interesting paper, Stephane Savanah argues for the thesis that concept possession is not only necessary, but also sufficient for self-consciousness (= the CP-hypothesis of self-consciousness).
He defines self-consciousness as the subject's awareness of itself as a self and contrasts this with what he terms primitive self-awareness, in which the organism has access to some of its mental or physical states, but is not conscious of its own existence as such.
Stephane's main thesis is that there is a constitutive link between concept possession and self-consciousness. This thesis rests on the claim that (1) the concept of agency is central to self-awareness, and on the intution that (2) there is a strong correlation between self-consciousness and intelligent behavior. He offers some preliminary support for his thesis by (3) drawing an analogy to perception, but most of the work in support of his thesis is being done by the argument that (4) all concepts bear a relation to the self-concept. I have some worries about each of these four points.
(I also have some worries about Stephane's interpretation of Bermúdez' theory, but I leave these to later, as I don't want to make this commentary too long.)

Before I begin with some more critical remarks on the paper, let me state what I like about Stephane's approach. First, I very much like his overall project of providing a framework for evaluating research on self-consciousness, for it is indeed a problem that the term is by no means used consistently in this area of study. Second, I agree with the view that we ought to distinguish consciousness on the one hand and self-consciousness on the other. Finally, I also agree with the thesis that self-consciousness – in contrast to consciousness – requires conceptual abilities.

I will now turn to my worries with regard to the details of his proposal.

(1) Agency and Self-Concept

The first worry that comes to mind regards Stephane's contention at the beginning of the paper that the concept of agency is central to the notion of self-consciousness. Undoubtedly, agency plays an important role in many theories of self-consciousness, and Stephane names a few examples of such theories. However, most authors don't take the notion of agency to be the single defining characteristic of self-consciousness – Bermúdez, for example, as Stephane himself points out, defines the self-

concept in terms of self-awareness of oneself as a perceiver, bearer of reactive attitudes and agent. A self-concept that relies on agency alone seems rather impoverished in comparison, and Stephane presents us with no reason to focus on agency and dismiss other aspects for an analysis of self-consciousness. Admittedly, he states that he is not going to argue for the centrality of agency for a theory of self-consciousness in the present paper, but given that his arguments rest heavily on this assumption, some hints at least would have been helpful.

(2) Intelligence and Self-Consciousness

Stephane then goes on to point to an intuition underlying his main thesis, namely that there is a strong correlation between intelligence and self-consciousness. First, I would love to hear a little more about how exactly this connection is to be understood, and about what exactly he means by "intelligence". Why is it that, as he points out, the same animals we view as relatively intelligent we also suspect of being self-aware? Is every being that displays flexible (that is, not stimulus-response based) behavior (which seems to be what Stephane takes intelligence to mean) thereby self-conscious? Or is it rather the case that we are inclined to think of certain animals as intelligent because they display signs of self-awareness (such as mirror self recognition)? In other words, what is the explanatory direction (if any) that Stephane has in mind here?

In the next section, Stephane distinguishes between three layers of development. Level 1 is reserved for organisms that display only stimulus-response behavior and cannot be ascribed with conscious awareness of their surroundings. Examples for these are single-celled organisms or bacteria. Level 2 applies to organisms that are conscious and whose behavior is more flexible than that of organisms at Level 1, but that nonetheless is still "based on a stimulus response paradigm". This is to say that although these organisms have mental representations, they are not capable of "understanding" or of inferential reasoning. This latter ability only comes with concept possession, which is reserved for Level 3 organisms. Since Stephane takes concept possession to be a necessary and sufficient condition for self-consciousness, organisms at this level not only possess conscious mental states, but also self-awareness[1].

Now, one might ask oneself what justifies the ascription of conscious mental states to the beings at Level 2? Arguably, reference to mental representation is justified as an inference to the best explanation just in case a subject displays intentional behavior that could not be explained otherwise. But here

[1] Note that self-consciousness and self-awareness are used interchangeably in Stephane's paper as well as in my comments.

there seems to be a tension between this kind of inference to the best explanation on the one hand, and Stephane's claim that Level 2 organisms display behavior that is based on the stimulus-response paradigm on the other hand. For behavior that can be explained on the basis of stimulus-response mechanisms does not seem to warrant the ascription of mental representations to the organism displaying the behavior (because the behavior can be fully explained without appeal to mental representations). Hence, the claim that Level 2 organisms have mental representations seems to imply that they are capable of intentional behavior (that is to say behavior that goes beyond stimulus-response reactions). However, according to Stephane's own reasoning, intentional behavior is a hallmark of intelligence. Thus, if the behavior of these organisms does indeed go beyond a mere stimulus-response mechanism (which could warrant Stephane's claim that they employ mental representations), then it seems that we should be inclined to speak of intelligent behavior. But if that is right, and if there is also a strong correlation between self-consciousness and intelligence, then it would seem that we would not only have to ascribe consciousness, but also self-consciousness to Level 2 organisms. Alternatively (and this strikes me as being more plausible), we might hold on to the distinction between consciousness on the one hand, and self-consciousness on the other. But then the connection between "intelligence" and self-consciousness might not be so strong, after all.

This need not necessarily pose an insurmountable problem to Stephane's position (although he himself contends that his CP hypothesis implies the two are co-dependent). One obvious solution would be to dissociate concept possession from intentional behavior. In other words, he could argue that intentional behavior does not require conceptual representations. This would allow him to hold on to the claim that Level 2 organisms, although capable of (nonconceptual) mental representation, do not possess concepts, and hence cannot be self-conscious. Level 3 organisms, on the other hand, not only display intentional behavior, but are also capable of conceptual representation. Perhaps this is indeed what he has in mind when he says that Level 3 organisms have the ability to "ascribe meaning" to a mental representation, or to "conceive" of their environment (which arguably goes beyond the ability to display intentional behavior), but as it is these notions remain somewhat elusive. (And, as we have seen, appeal to the notion of intelligence does not really help either, as long as this notion itself is not clearly defined.) So I think that in order to make his point, Stephane needs to say more about how he understands the notion of a mental representation, and about the difference between the behavioral (and representational) abilities of Level 2 and Level 3 organisms. In particular, he should distinguish more clearly between nonconceptual and conceptual representations.

The CP-Hypothesis

This bring us to the second half of his paper, in which he goes on to defend his central claim that all concepts bear a relation to the self-concept, and hence that every concept possession entails self-consciousness. In other words, the self-concept is taken to be the primary concept on the basis of which every other concept is formed. (This means that any demonstration of conceptual abilities in animals would automatically entail a demonstration of self-consciousness, thus providing a criterion for empirical investigations – but again, it would be very helpful to know what kind of behavior in particular Stephane would admit as providing sufficient evidence for concept possession.)

(3) The Analogy to Perception

Stephane presents three lines of thought in support of his thesis. The first line of thought appeals to the notion of ecological perception. Since all perception is perception from an egocentric perspective (pace Gibson), everything that is perceived is perceived relative to the perceiving subject. Thus, perception necessarily contains self-specifying information. Stephane takes this to show that "nothing can be perceived without the perceiver co-perceiving itself".

But the fact that perception necessarily involves self-specifying information does not imply that therefore every perception is co-perception of a self. Because the self-specifying information can be thought to be implicit in perception, it is not clear that it enters the explicit content of perception at all. And if perception were indeed to entail self-perception, so much for the worse for Stephane's position. For it is unclear why we should then not say that perception entails self-consciousness. After all, if I perceive the computer in front of me, it seems most natural to say that I am thereby conscious of the computer. So if my perception entails co-perception of myself, then should we not say that I am conscious of myself insofar as I perceive myself? Stephane thinks that the fact that perceptual awareness can be nonconceptual excludes this possibility, for he takes self-consciousness to be conceptual. But as this is a claim that needs to be established in the first place, he cannot use it in his argument here.

In fairness, however, the perceptual case is just supposed to provide an analogy; it is not taken to be a genuine argument. Thus, let us move to the "real" argument.

(4) The Web of Concepts

The central argument of the paper is supposed to show that every concept is necessarily related to the

concept of an AGENT. And since agency was taken to be the central characteristic of the self-concept, it is thereby supposed to show that every concept is necessarily related to the self-concept. Stephane's example is that of the concept BLADE. Grasping the concept BLADE, he argues, entails grasping the concept CUT, and this in turn entails grasping the concept of AGENCY, as cutting is an action performed by an agent. And grasp of the concept AGENCY (and thus AGENT) implies the ability to apply the concept not just to others, but also to oneself, and hence entails the ability to think of oneself as-an-agent, which is a form of self-consciousness. Now, while this might be plausible for the concept BLADE (and other physical concepts like it), it seems far less plausible for other concepts. Stephane anticipates this objection and argues that this is indeed supposed to work for any concept; his additional example is that of the concept BEAUTY. According to Stephane, BEAUTY needs to be apprehended, apprehension is an action (which requires an agent), and hence the concept of BEAUTY entails the concept of agency. But is the apprehension of beauty really a type of intentional action? Is it not rather something that "happens" to you, as in when you are struck by the beauty of a piece of music you are listening to, or by the beautiful scenery in front of you? In what sense do we have to conceive of the apprehension of beauty as a form of intentional action? (Not to mention the fact that is seems quite possible to have a concept of beauty that does not necessarily entail apprehension.) Or take another example. What about the concept of an abstract object, or of possible worlds, or some other philosophical concept? Do these imply the concept of action? But in what sense? Do they require someone who actually thinks of these concepts? Even if they did, it is not obvious that thoughts are intentional actions (at least there is substantial philosophical disagreement as to whether thoughts are best thought of as actions, or whether they entail any sense of agency). What about perceptual concepts? They might arguably imply the notion of a perceiver, but that would at best give you a concept of yourself as a perceiver, not as an agent, since it seems implausible to think of perceptions as actions. Likewise emotions – the concept of fear, for instance, might at best provide you with the concept of yourself as the bearer of that affective state, but not as an agent, for being in fear is not something that you "do". Thus, the claim that every concept requires the prior possession of the concept of an agent seems unconvincing.

Let me conclude by saying that I very much enjoyed reading and thinking about Stephane's paper. I now look forward to the conference and to what I expect to be a very interesting discussion.