

## ENACTIVISM AS A NEW VERSION OF EXTERNALISM AND THE PROBLEMS OF SELF-KNOWLEDGE

Department of Cognitive Science and Epistemology, Institute of Philosophy, Nikolaus Copernicus University Torun Poland

*The main topic of this article concerns the question, whether the self-knowledge can be still authoritative from the enactivistic point of view. The problem rests on two assumptions: 1. The definition of self-knowledge claims that a subject has a direct given knowledge about an intentional contents of his attitudes. 2. The content of subject's attitudes is determined by external factors, which could be unknown to the subject. It means that the subject has the limited access to the content or its determinants understood as satisfaction conditions of his mental states. On this basis of the obvious conflict between the two theses the first-person-authority can be questioned.*

### Introduction

In this paper I want to consider the problem of the enactive cognition in reference to self-knowledge, which could be formulated after M. Wilson and Ch. Peacocke as it follows [1, p. 625; 2, p. 203]:

1. The main thesis of enactivism claims that organisms including human subjects organize themselves by interacting with their environment.

2. The definition of self-knowledge claims that a subject has the knowledge about the intentional contents of his attitudes without first checking their environmental relations.

3. To understand cognition we must study the situation and the situated cognizer together as a single, unified system.

The followers of the theories of embodied mind and enactivism would say that the content of subject's mental states is deeply rooted in the body's interactions with the environment because the whole cognition is. These theories characterises a focus on the neurobiological basis of the mind and models of the Self, which is reduced to its physical basis. According to these theories a subject is a system to which the self is necessary for proper functioning, and consciousness allows the system to control and to monitor these functions. The self has therefore any special ontological status. As Thomas Metzinger writes: "Subjective experience has not been developed in pursuing the old philosophical ideal of self-knowledge, but it has been evolutionarily successful, because it has enabled a more flexible form of action control." [3, p. 175].

Such evolutionary position will therefore refer to the "needs" of the organism. Self-knowledge is required by human (called "system") because it is useful to him. In the same way the problem of access to the mental content will be solved. Organisms (systems) in some cases need not know the broad context, which would influence the content of their thought. Metzinger assumes active externalism, which claims that "the content-fixing properties in the environment are active properties within a sensorimotor loop realized in the very present" [3, p. 115].

From that reasoning the following question arises: How to connect the first-person authority with the enactivism, which is *de facto* a new version of the externalism about content, if the judgments constituting a subject's knowledge about his own mental states have a different epistemic status than the judgments constituting knowledge about the external world? Could the self-knowledge be still authoritative from the enactivistic point of view?

### The problem of the mental content determination

The problem of this kind of determination concerns the nature of the content of mental states. The main question, which arises here sounds: How the content of mental states is determined? Is it determined by external or internal factors? It was said in the upper made quotation, that such factors are external to the subject. So how does he gain a knowledge about this content?

This kind of knowledge, called *de se*, could be defined as a subject's knowledge about his own mental states. This knowledge is infallible, direct and authoritative, which means that a subject has a privileged access to his own mental states. But this knowledge is possible only when there is a content constituting the propositional attitudes. In other words, the subject has certain beliefs only if he stands in some relation to the environment. It is a metaphysical statement about the intentional content of subject's beliefs. Mental states like beliefs, desires and wishes have a propositional content that is expressed in the form of "that-clauses". Such "that-clause" expressing the content of the belief is epistemically different from the first-person clause (i.e. "I believe..."). While the subject cannot be mistaken by the second-order belief (i.e. "I believe that I believe"), he might doubt content of his first-order belief (e.g. "I believe that at night every cat is black.") because the content of the first-order belief is not self-verifying. The subject has to stay in relation to his environment and this environment is to what he refers when he expresses his first-order beliefs about something. The problem is, that the relation between a subject and an environment is not complete because the environment is external to him. That means, that the content of subjects attitudes can be in part determined by factors, which are unknown to the subject.

On the other hand the statement: "I believe" is self-verifying, because a subject cannot be wrong about it. If he has a belief, than he knows, that he has it. This statement is Cartesian like. Descartes's paradigm for the self-knowledge was cogito. According to Descartes, "I think" leaves no doubts, that if I think, that I really do it. It leaves therefore no place for scepticism about me, as a subject, who make statements about his propositional attitudes. These sorts of judgements constitute knowledge in virtue of being made and having the content they do.

From the internalistic point of view the problem of the content determination is related to the former paradigm in philosophy of mind namely the computational theory of mind. In this approach the

content determination is described in terms of function (hence functional role semantic) because concepts, as a part of the content in propositional form, fall under a function. Using a concept in reference to an object means satisfaction of the function expressed by sentences.

Since concepts are the basis constituents of the content it is noncoherent to claim that there exists something like nonconceptual representational content, because representation in such a theory is a sentence in the language of thoughts, where some parts of the sentence are concepts (hence they are parts of the representation as well), which are symbols.

Hence on the assumption, that representations are functions, does not exist something like a separate representation of property  $F(x)$ , and a separate representation of object  $a$ , because a representation of property is every time the representation of object  $F(a)$ , because that needs a satisfaction of function (this point of view is connected with Frege's thesis [4]).

On the field of the computational theory of mind does not make a sense to consider, whether for example perceptual experience has a nonconceptual representational content, what is the subject of the dispute between McDowell [5], who is against the thesis, and Peacocke, who is agree with this thesis. Both of them assume the representational character of content. But they assume a different character of concepts. For Peacocke concepts are abstract objects and can be sometimes a part of the content and sometimes not. But if somebody regards concepts as symbols (like Fodor for example [6]), then the case is simple, because every single process of thinking (regardless, whether it is conscious or unconscious) consists of symbols.

2. Externalist's points of view about the possession of the concept of belief

What could be such factors, which determine the content? To answer the question first we have to make two distinction:

- i. Between synchronic and diachronic externalism.
- ii. Between social and physical externalism.

According to the (i.) synchronic externalism holds that the content of propositional attitudes is determined by an current environment of the subject and his disposition to respond to it. This externalism take into consideration only the actual situation of the subject, without reference to his causal story in the past.

In contrast to this, diachronic externalism holds, that the causal story, that means, all facts in the past, which have had an influence on the subject, together with an environment, are important determinants of the content of subject's propositional attitudes.

According to (ii.) social externalism holds, that content of thoughts is determined in part by the social environment of a subject, especially by how others in our linguistic communities use words. This "others" could be experts, who establish and better know the scientific names of such objects like for example trees. This version of social externalism propagate Hilary Putnam and also Tylor Burge although their versions of social externalism are different.

In opposition to the social externalism the physical externalism says that contents of subject's thoughts is

determined in part by our relation to our physical, non-social environment. This kind of externalism was propagated for example by Donald Davidson. His basic statement claims, that the objects of our beliefs are causes of this beliefs [7, p. 151]. The main question, despite of the versions of externalism is: How is it possible, that a subject is able to know his own mental state, with a particular content, without having an idea, what constitutes this content?

This is the place, where two aspects of self-knowledge meet together: the metaphysical aspect concerning content, and epistemological aspect which concerns the subject's capacity of attributing mental states to himself.

The epistemic aspect of self-knowledge is interesting only in the situation where a subject attributes propositional attitudes with a particular content to himself. That is, only in the case of the clause: "I believe that..." but not "He (she) believes that..." It is a special capacity of human beings (for sure, animals – who knows?) to refer to himself, precisely, to his inner state. This capacity is called self-reference, or reflective or reflective consciousness, or simple reflection. According to Davidson the act preceding the action understanding is simply the observation of an agent's behavior. On this basis an interpreter ascribes to an agent intentions, beliefs, desires and other propositional attitudes.

Davidson's approach has important consequences for the first-person authority, because it seems to be obvious, that the thinker does not need to watch himself, that means, to watch his own behavior, to know, that he has a belief. Davidson's argument, why it should be in that way, proceeds as follows [9, p. 395-396]:

1. To have beliefs is to be an agent.
2. To be an agent is to be capable of acting.
3. To be capable of acting one must be able to have intentions.
4. One's beliefs and desires must be coordinated in the right way to provide rationalizations of one's potential actions.

Davidson claimed that a person can have neither beliefs nor concepts without having the concept BELIEF, and this means that the person must have some beliefs about beliefs, at least the most crucial belief that beliefs can be right or wrong.

...[A] person cannot just believe that he or she is seeing a cat: in order to believe this, one must know what a cat is, what seeing is, and above all, one must recognize the possibility, however remote, that one may be wrong. [9; 8]

So one cannot be an agent without having the concepts of belief and other related propositional attitudes [8, p. 396]. Davidson calls the way of interpreting agent's behaviour the epistemic triangle: To understand the behavior of an agent the interpreter has to have a hypothesis about his intention and then check this hypothesis with respect to the external conditions of the world. In this way, he can verify or falsify his interpretation. If it is wrong, then he has to change it and form another hypothesis. To do this, the interpreter needs to possess the mechanism of thinking and rationalizing. In order to treat the subject's behavior as rational, one cannot quit the mentalist terminology,

because it refers to subject's beliefs and intentions, which form a coherent, logical, integral whole.

This standpoint meets a critic presented by E. Lepore and K. Ludwig [8]. According to them the recognizing whether somebody has a belief or not, happens by characterisation of internal state of a person, so by his intention. This recognizing and rationalizing could have a form of the condition: "If I intend to do A, then that is a result of this intention to do so." [8, p. 396]

This different opinion Lepore's and Ludwig's derives from the internalistic standpoint. For supporter of such point of view, the satisfaction conditions of attitudes must be represented in their content, hence internal. The way to recognizing that one has a concept of belief leads not like by Davidson from action to the concept of the action, but from concept of intention to action. But finally, also for them, one has to have concept of belief to have a concept of intention. So the conclusion is the same:

One cannot be an agent without having the concepts of belief and other related propositional attitudes. [8, p. 396]

Externalistic point of view presents Davidson. According to him one can recognize, that somebody has a belief, by an external observation of the behaviour of this person. And this claim has important consequences for the first-person authority, because it seems to be obvious, that the subject does not need to watch himself, that means, to watch his own behaviour, to know, that he has a belief.

Davidson's starting point for investigation about self-knowledge is the problem of epistemic asymmetry in the case of ascribing current mental states by a subject to himself and by other persons to the subject. In other words the problem concerns the asymmetry between first and third person perspective. He claims, that this is the same kind of the asymmetry, which happens between speaker's and interpreter's knowledge about word's meaning.

The main argument against that explanation says, that Davidson confuse the explanandum and explanans. Namely he will give an answer for the question of asymmetry between first and third person perspective with help of the asymmetry, which happens between speaker's and interpreter's knowledge about word's meaning. So the explanans and explanandum are the same, namely it is still asymmetry. But the asymmetry does not explain why the speaker always have a better epistemic position in knowing his own mind. The thesis about asymmetry is just a first step to the explanation, but not the explanation itself, as a premise is not a conclusion.

If Davidson explains the asymmetry with the help of his theory of meaning, then must be said, that according to Davidson the meaning is dispositional. So it is to identify for an interpreter on the basis of speaker's behaviour. This thesis has two consequences: First, that the knowledge about meaning has to be inferential. It seems to be strange in the case of self-knowledge, because we have direct, noninferential access to our mental states. But still it is not so, that we do not know our thoughts. We know for sure the mode of our mental states, like believe or desire. What we could not know concerns the true-value condition of the content.

Second, that in the case of this thesis consciousness does not matter, but for explaining mental states it should. This is the third, psychological aspect of self-knowledge. Believing, reasoning, etc. are the mind's activities. Thinking is a mental process, and is thus psychological. But because of this aspect our reasoning, and also rationalizing can fail.

Despite of it Davidson's externalism could be called "naturalised" though there is a difference between Quine's and Davidson's concept of naturalisation. Quine uses it in reference to stimulus meaning, because he propagates the proximal theory of meaning, where the meaning has internal character. Davidson advocates the distal theory of meaning, because of the external character of meaning and reference to external causes, which make the causal history of speaker. The following thought-experiment became the canonical example of physical externalism:

Suppose, that a thunder strikes a tree in a swamp and Davidson stands nearby. By an accident his body is reduced to the molecules and from different molecules comes in to being Davidson's physical Doppelgänger. The Doppelgänger moves like Davidson's, recognizes Davidson's friends, calls the house of Davidson "home". It seems to be no difference between him and Davidson. But there is a difference. The Doppelgänger "Swampman" can not recognize Davidson's friends, he can not remember the house, because he has different causal history. He was not born, grew up, he never meet the friends before. So he did not learn the context of using words in the same sense as Davidson did.

This thought-experiment have been criticised by Ludwig and Lepore. It seems to be as follows:

If we accept the radical interpretation theory in identification of mental states of speaker, then we need also an omniscient interpreter, which has to have also a knowledge about causal history of speaker and not only about actual context. Otherwise the speaker could have the true beliefs but the interpreter could not find him in agreement with oneself. So there is a methodological inconsistency in this thought-experiment. It seems to be inconsistent also because Davidson claims the synchronic externalism. This experiment shows, that on mental states of the Swampman has had the influence also his past.

Swampman is abnormal in two ways: his ontogeny is wrong and his phylogeny is wrong. It was Swampman's ontogeny, his personal intellectual development (the lack of it), his prior involvement with the world (his lack of it), his (missing) wider embryology as it were, that exercised Davidson. What exercised me, more especially, was the peculiar phylogenesis (the lack of it) of your newly created double, the fact that Doppleyou has no supporting evolutionary history [10, p. 93].

So the externalism in Millikan's account appeals not only to causal history of the subject but also to his evolutionary story. According to her having beliefs is a biological function, hence burdened biological history, which is the main determinant of beliefs.

But coming back to the problem of the incomplete relation between a subject and an environment, Millikan does not give a solution of it. She does not accept any kind of intermediaries in the relation

between subject and object, which are making this relation complete.

The standard reference relation has the form  $R(x,o)$ , where “x” is a subject and “o” an object in the world. But everybody knows the Fregean cases of mistaking Morning Star with Evening Star, where the both names refer to the same object. One can not know, that Morning Star is Evening Star. This person think, that the object, which he see mornings on the sky is different from the object visible evenings. But in fact it is the same planet, namely Venus. Frege solves the problem with the help of intermediaries called sense. Sense is the mode of presentation of the object, so one can have two different senses of the object, because sense determines his reference. So the relation has in fact the form  $R(x,s,o)$ , where “s” stands for sense.

So once again according to Frege, names (*Eigennamen*) have sense (*Sinn*) and reference (*Bedeutung*). The sense of a name is a concept which content is the way an object is given to us, or its mode of presentation (*Art des Gegebenseins*). The reference of a name is a real object. The architecture of Frege’s semantics on the level of the name is as follows:

	<i>Name</i>
<i>Sense</i>	Concept (mode of presentation)
Reference	Object in the world

Millikan does not accept this solution. She claims that Frege has confused the intentional contents of a representation, with attributes of the vehicle of representation. She says, that [s]omething like mingling of vehicle with content motivates Frege’s conception of *Sinn* or mode of presentation [11, p. 499].

The corn of her critic is what Garreth Evans called “the datum sense fallacy” or “the homunculus fallacy”:

[...] when one attempts to explain what is involved in a subject’s being related to objects in the external world, by appealing to the existence of an inner situation which recapitulates the essential features of the original situation to be explained... by introducing a relation between the subject and inner objects of essentially the same kind of relation existing between the subject and outer objects [12, p. 397].

This error of course should be avoided. Despite of Millikan’s further considerations, this quotation leads to important question, which is relevant to our problem of externalism and self-knowledge: “How will the inner eye then perceive the inner picture? In the same way that the outer eye does?” [11, p. 440] And henceforth: “How it must move the thinking system in order to represent itself?” [11, p. 456] This questions are important, because, between a subject and his mental states should be no intermediary. The relation of self-reference should be direct, otherwise there is no first-person authority. The subject has to have direct, privileged access to his mental states. The question can be formulate now in this way: “How the subject “sees” his beliefs?” Or more professional: “What is for a subject to posses a concept?”

**The solution of externalist problem about content and the first-person-authority**

Epistemological aspect of self-knowledge is associated mainly with the question of subject’s access

to the content of his mental states. On the other hand the issue of first person authority refers to the question of justification of person’s own mental states with certain content. That means that at the very beginning, before the question was asked, it was assumed that the subject has the privileged access to his own mental states. He does not need to make any inferences and to refer to some external relations, to know what propositional attitude he is in. Justification of such knowledge is thus independent of the influence of environment on the subject. In general, the problem of reconciling self-knowledge with externalism is associated with an acquiring the content of mental states.

It was said, that a person can have neither beliefs nor concepts without having the concept BELIEF. So now it seems to be important to investigate the connection between the content of mental states and concepts, which are constituents of the content. This connection has to have implications for self-knowledge. I think, that without an answer to the question, “What is it for a subject to posses a concept?” we cannot answer the question, “How does a subject acquire self-knowledge?”

This assumption is supported by reference to Christopher Peacocke’s theory, which can be divided into two parts:

- A theory of concepts, which answers the general question, “What is it for a subject to posses a concept?” Peacocke considers this question in *A Study of Concepts* [13].
- A theory of knowledge, which tries to reconcile the externality of content with first person authority understood as the infallible and incorrigible knowledge about our mental states. This is the main subject of Peacocke’s *Being Known* [2].

These two theories together show a possible way of acquiring self-knowledge. Peacocke thinks, that the constituents of propositional content are concepts that are individuated trough their possession conditions. These possession conditions together with determination theory tell how a given concept’s semantic value is fixed, and guarantee the rational sensitivity of a subject. This sensitivity is required for somebody who attributes propositional attitudes with a particular content to himself or another. In other words, rational sensitivity is sensitivity to the satisfaction of the possession condition for the concepts in the content attributed.

Self-knowledge is a special case of knowledge, yet it still rests on possession conditions for concepts describing our mental states. To say, “I believe that I believe,” we have to have the concept of belief. But it is not enough to state that making these judgments itself makes them true. There is still a question: Why does making judgments about our own mental states itself make them true? The answer is given in the last thesis:

Our distinctive ways of coming to knowledgeable self-ascribed beliefs are correct methods because of the nature of belief. [15, p. 102]

If self-knowledge is a special case of knowledge, then to have self-knowledge means:

To have reasons for a belief, which ensure its truth, is a consequence of the nature of the concepts it contains.

This is the best epistemic position to have the knowledge about our own mental states, but this is only a special case of some more general phenomenon [13, p. 158].

Therefore if the subject is rational, then is he able to make transitions between contents of first-level beliefs to the content of second-level beliefs. It is allowed by the very nature of such a content, because the content holds the true-value, from the basic content. This explanation leads to conclusion about some hidden inference in first person knowledge. But this solutions allows also to integrate the externalism about content with first person authority.

## References

1. M., Wilson, 2001 Six views on embodied cognition, in: "Psychonomic Bulletin & Review", 2002, 9 (4), pp. 625-636.
2. Ch. Peacocke, Being Known, Oxford 1999, Clarendon Press.
3. T. Metzinger, Being No One. The Self-Model Theory of Subjectivity, Cambridge, Massachusetts, London 2004, MIT.
4. G. Frege, Über Sinn und Bedeutung. Günther Patzig (ed.), Funktion, Begriff, Bedeutung. Göttingen 1994.
5. J. McDowell, Mind and World, Cambridge 1994, Harvard University Press.
6. J.A. Fodor, Fodor J.A., Concepts. Where Cognitive Science Went Wrong?, Oxford 1998, Clarendon Press.
7. D. Davidson, A Coherence Theory of Truth, [in:] Subjective, Intersubjective, Objective, New York, 2001, Clarendon Press.
8. E. Lepore, K. Ludwig, Donald Davidson. Meaning, Truth, Language and Reality, Oxford 2005, Clarendon Press.
9. Donald Davidson, The Emergence of Thought. Erkenntnis, 51(1999), 7-17.
10. R. G. Millikan, Language, Thought and Other Biological Categories, Cambridge MA 1984, MIT Press.
11. Millikan R.G., 1991, Perceptual Content and Fregean Myth, Mind, vol. C.4.. October.
12. G. Evans, Molyneux's Question, [w:] Collected Papers, Oxford 1985, Clarendon Press.
13. Ch. Peacocke, A Study of Concepts, Cambridge 1992, MIT Press.
14. A. Pacholik-Żuromska, Poznając siebie, czyli konceptualne podstawy samowiedzy, Toruń 2011, Wydawnictwo Naukowe UMK.
15. Ch. Peacocke, Conscious, Attitudes, Attention and Self-Knowledge, Privileged Access. Philosophical Accounts of Self-Knowledge, ed. Gertler B., Ashgate 2003.

А. Пахолік – Журомська

**ПРОБЛЕМИ САМОПІЗНАННЯ : ІНАКТИВІЗМ ЯК НОВА ВЕРСІЯ ЕКСТЕРНАЛІЗМУ**

Основна тема статті стосується питання : чи може самопізнання все ще бути авторитетним з точки зору інактивізму ? Дана проблема ґрунтується на двох припущеннях : 1 ) визначення самопізнання свідчить, що суб'єкт отримує безпосереднє знання про його інтенціональнoм змісті установок; 2 ) зміст установок суб'єкта визначено зовнішніми факторами , які можуть бути невідомими для суб'єкта. Це означає , що суб'єкт має обмежений доступ до даного утримання чи його детермінанти розуміються як умови задоволення його ментальних станів . На основі цього очевидного конфлікту між двома тезами авторитет першої особи може бути поставлений під сумнів.

А. Пахолик-Журомская

**ПРОБЛЕМЫ САМОПОЗНАНИЯ: ИНАКТИВИЗМ КАК НОВАЯ ВЕРСИЯ ЭКСТЕРНАЛИЗМА**

Основная тема статьи касается вопроса: может ли самопознание все еще быть авторитетным с точки зрения инактивизма? Данная проблема основывается на двух предположениях: 1) определение самопознания гласит, что субъект получает непосредственное знание о его интенциональном содержании установок; 2) содержание установок субъекта определено внешними факторами, которые могут быть неизвестными для субъекта. Это значит, что субъект имеет ограниченный доступ к данному содержанию или его детерминанты понимаются как условия удовлетворения его ментальных состояний. На основе этого очевидного конфликта между двумя тезисами авторитет первого лица может быть поставлен под сомнение.