https://doi.org/10.18778/1689-4286.38.01

HYBRIS nr 38 (2017)
ISSN: 1689-4286

**PAWEŁ GRABARCZYK**
**DAWID MISZTAL**
UNIVERSITY OF ŁÓDŹ

# INTRODUCTION: TRENDS IN CONTEMPORARY POLISH PHILOSOPHY OF MIND

The landscape of current philosophy of mind in Poland is varied and reflects most of the contemporary international trends in the subdiscipline. Its health can be easily seen by the amount of institutional backing it gets (various cognitive studies courses across the country) and the existence of specialized journals (for example *Avant* and *Studia z kognitywistyki i filozofii* umysłu). For this reason, we decide to focus mostly on one specific group of contemporary trends visible in the subdiscipline: the recent resurgence of various forms of internalism and the critical reception of this resurgence. Let us briefly outline the phenomenon in question. There are no doubts that at the end of XX century both: philosophy of mind as well as philosophy of language made a very distinct turn towards externalism. Even though classic arguments of Hilary Putnam (1975), Saul Kripke (1972), and Tyler Burge (1979) concerned linguistic (as opposed to mental)[1] content the upshot of the discussions they spawned made a great impact on philosophy of mind. One of the most visible results of this externalist tendency is the idea of embedded and extended cognition. According to the former, cognitive content is the result of an interplay between the cognitive agent and its environment. The relations between the agent and its environment are understood to be so crucial that the identity conditions of mental contents is oftentimes construed as dependent on the surroundings of the agent (similarly to how the environment was determinant for linguistic content in Putnam/Kripke's theory). The latter idea (the notion

[1] See Grabarczyk 2016 for a review of different understandings of „linguistic content" and „mental content".

of extended cognition) points out to the fact that not only the mental content but also the operations performed on this content can sometimes be delegated to external factors. What is even more important, the notion of mental content (or mental representations) started to be automatically understood in externalist terms as for several authors content has to be understood as relating to some external object, property, or event and the idea of mental content devoid of external targets seems to be almost self-contradictory (Kriegel 2008). It is in many ways fascinating to see how quickly externalism changed from a new and radical approach to the dominant perspective.

As is often the case, this dominant position resulted in a void which could then inspire a new wave of more sophisticated takes on internalism – new reasons to "turn inside". Of course, the theories characterized as such do not have to self-identify as "internalistic" or refer to internalism, but they do retain the main aspect of it: they create a space in which notions important for cognition can be defined without the appeal to external environment of the cognitive agent. Let us list some of such approaches that are relevant for the papers included in this volume and characterize them briefly.

The first notion that internalistic philosophy of mind often appeals to is the notion of computation. The reason for it is that it seems to be possible (at least in principle) to construe computation as a purely internal set of operations that is devoid of any external targets (it is, of course, still perfectly possible for elements of computations to refer to internal states of the computing machine). If it is possible to specify computations regardless of their target or application (in other words, if the identity criteria of computations do not demand us to refer to external objects), and if the notion of computation is relevant for cognition (which is, of course, a contentious claim in and of itself), then there still is some hope left for internalism in philosophy of mind. One specific subset of computations that (according to some authors) is especially relevant for cognition is the inferential subset of computations performed by the system (this inferentialist approach to cognition is especially evident in the classic example of Ned Block's theory of narrow mental content (Block 1987)).

Another important notion that creates space for contemporary internalism is the idea of structural representations championed by Cummins (1989) and revived lately in Ramsey (2007). In short, the idea

of structural representations boils down to the fact that some internal structures of the cognitive agent relate to their targets due to the fact that their structure is homomorphic to the target's structure. On the face of it, this idea is fully externalist (as the structure in question is specified by appeal to the target) but the trick is that it allows for a fully internalistic reading, because one of the characteristic aspects of structural representations is that they can be processed by the system off-line. For this reason, it is possible to imagine a situation in which a given system entertains and modifies a given representation to a point in which it loses the connection with its target (but still has some cognitive value).

The list of new internalistic ideas and trends in contemporary philosophy of mind wouldn't be complete if we did not mention the theory of predictive coding and the interrelated understanding of minds as anticipatory mechanisms (Hohwy 2013, Clark 2016). Contrary to traditional views on mind, which saw it as passive receiver and categorizer of stimuli, these theories interpret the mind as constantly constructing the reality. According to this view, the reality we live in is more similar to a simulation or conscious hallucination than to reality (understood in an old-fashioned sense). We could say that we literally live in a bubble and use the connections with our surroundings as "reality checks" – signals which help us correct and modify our predictions. Minds do not interact with reality, they live in self-constructed models of it. It is not hard to see that this radical idea (which has been in many respects anticipated by Metzinger 2009) gives hope to internalists as it relegates the role of cognitive system's environment to that of a pragmatic "checkpoint" needed only to steer our cognition in the right direction (but not to shape it).

***

Papers collected in this volume relate (directly or indirectly) to these "inward" trends of modern philosophy of mind. In a paper entitled "The false dichotomy between causal realization and semantic computation", Marcin Miłkowski shows that mechanistic understanding of computation does not prevent us from semantic considerations. As he points out (following Bechtel 2009), "computational modeling is not just about 'turning inside'. It requires looking up, down, and around". Miłkowski does not prevent computations to be devoid of meaning – on

the contrary, he admits that there definitely are examples of computation which are not semantic (in other words, the ideas of semantic content and mechanistic computation are logically independent). Still, it should be pointed out that being logically independent does not mean that there are no significant relations between mechanistic computations and semantic content. To the contrary – if present, semantic aspects constitute constraints on computation. In this sense Miłkowski shows that computational theories of cognition (specifically mechanistic ones) are in fact agnostic when it comes to the difference between internalistic and externalistic interpretations of cognition. One of the advantages of this paper is that it clearly differentiates between the social and environmental factors that could influence internal computation of a cognitive system (a difference that is well known but, sadly, often conflated). Miłkowski shows this on a very convincing example proposed by Shagrir (2006) in which the internal states of a machine can be interpreted as a conjunction or as a disjunction, depending on the social practices that surround it.

A similar line of argumentation can be found in Paweł Gładziejewski's paper "Just how conservative is conservative Predictive Processing?". Gładziejewski looks at the theory of predictive coding and shows that, contrary to what may seem to be the case at the first glance, this theory does not have to clash with the ideas of 4E cognition (embedded, embodied, extended, and enactive). Similarly to what Miłkowski does for mechanistic computationalism, Gładziejewski argues that the theory of predictive coding can be seen as agnostic in the sense that it is possible to interpret it as compatible with externalism. This idea is novel, since, as Gładziejewski points out, predictive coding "was initially construed in a manner that dovetails with traditional approaches in cognitive science, i.e. ones that see cognition as matter of inferential, exclusively intracranial processes involving richly structured representational states" (he calls it "conservative" or "radical" reading). One of the reasons why this conclusion is possible is that, as Gładziejewski points out, the notion of "inference" used in the theory of predictive coding is very liberal and differs from strict understanding proposed by Friston (2013). Specifically, the inferences proponents of predictive coding talk about should be truth preserving (which obviously ties the cognitive system with its surroundings). In addition to this, Gładziejewski shows that the methods the theory of predictive coding

uses to delineate internal and external processes do not suffice for such a demarcation. Last but not least, what the paper explains is that the type of representationalism that predictive coding appeals to has necessary ties to environment and to the surrounding social practices.

To complement the new forms of internalism (and their critical adoption), it is also good to look back at the original externalist's argumentation and reevaluate it from contemporary point of view. This task is taken by Witold Hensel in a paper entitled "Watered Down Essences and Evasive Speech Communities. Two Objections to Putnam's Twin Earth Argument". Hensel analyzes the seminal Twin-Earth thought experiment and shows that it rests on two necessary assumptions which are very hard to accept in the light of contemporary science. The first assumption is that objects referred to by a given natural-kind name contain common micro-structures (microessentialism). As pointed out by Hensel, this assumption is not corroborated by contemporary science (neither biology nor chemistry). The second, less obvious assumption Putnam makes is that it is possible to delineate different communities (and thus the intended reference of the terms they use). The problem can be presented as follows: Putnam helped us realize that all natural-kind terms have a hidden indexical component that ties them to a given environment. For example – the term "water" used by inhabitants of Earth was always used as referring to the microstructure of a specific liquid found on Earth. But why should we treat linguistic communities of Earth and Twin-Earth as separate? It is not obvious why should the boundary be put in this particular place, but this ability seems to be presumed in Putnam's argumentation.

An interesting illustrations of the tension between external and internal perspective can also be seen in the papers of other authors. Katarzyna Kobos discusses the situations when perception occurs in the absence of sensory stimuli. Can we say we actually perceive anything in such circumstances? Can silence be said to be heard or darkness to be seen? What is the role of the brain (if any) in forming of sensory experience? To what extent the latter is dependent on external input? In her attempt to answer these questions, Kobos meticulously analyzes two models of perceptual response to the absence of sensory impingement. Consequently, she turns to embodied predictionism as it seems to be more theoretically satisfying and more promising in terms of its explanatory power.

Marek Prokopski brings our attention to the problem of other minds. The author is mainly interested in the conceptual formulation of said problem (as opposed to ontological and epistemological formulations) which poses the question of possibility of universal mental concepts describing emotional states or inner experience. In other words, can we – asks Prokopski – justifiably use the mental concept of – say – pain, based on personal experience of pain, not only in the first person but also in third person cases? The challenge here, according to the author, is "to develop plausible positive account of mental concepts", since the negative one would lead to the disputable conclusion that we have two different mental dictionaries: a first-personal and a third-personal.

As it is often the case with opposing theoretical proposals, however incompatible they may seem, they may be inspirational for searching possible ways to reconcile them. Przemysław Nowakowski's interesting attempt to integrate computational and embodied approach to cognition can be read precisely in this context. However, to achieve his goal the author adopts an interalist rather than externalist perspective on the evolution of cognition. He assumes that internal complexity of organism is at least equally important in evolutionary shaping of cognitive processes as external, environmental factors. On this basis, Nowakowski presents his own approach to embodied cognition which he dubs E-codes' approach (E-codes being "Efficient, robust and body-specific processing"). And what he hopes to obtain by means of this approach is to create an opportunity for developing conceptualizations that would do justice not only to the embodiment thesis but to empirical data as well. Although, as he cautiously remarks, that would require "more comparative meta-analysis and computational modeling than psychological experiments".

The debate between internalism and externalism is continued in the next two chapters. In the first one by Krystyna Bielecka, this opposition is thoroughly examined in the context of the problem of intentionality, and the focus is on the semantic internalism as a potential solution to this problem. Analyzing the notion of narrow content (which basically means a content limited to its functional role within the cognitive system) in its radical interpretations, the author presents detailed critique of the aforementioned stance. In her opinion, semantic internalism deprives the content of any other than formal (i.e. syntactic)

properties, and thus it renders ascribing truth to representations (or any other semantic property for that matter) impossible.

The anti-internalistic tone of Bielecka's text is seemingly further reinforced in the last section of our book in which Maria Matuszkiewicz offers her exhaustive discussion of Robert Stalnaker's work entitled *Our Knowledge of the Internal World* (2008). The chapter identifies and elaborates the central issues of Stalnaker's argument such as our epistemic relation to our experience, the relation between experience and knowledge, or the relation between objective knowledge and the knowledge we can have only from a certain perspective. But Matuszkiewicz not only fully exposes Stalnaker's version of externalism, pointing additionally to its affinity with other philosophical positions (with contextualism, for example). She also notices that Stalnaker's externalism, being rather a methodological perspective than metaphysical view, is not altogether so anti-internalistic as it may seem at first glance.

## REFERENCES

Bechtel, W. 2009. Looking Down, Around, and up: Mechanistic Explanation in Psychology. *Philosophical Psychology* 22 (5): 543–64. doi:10.1080/09515080903238948.

Block, N. (1987), Advertisement for a Semantics for Psychology. *Midwest studies in philosophy* 10.1, p. 615-678.

Burge, T., 1979. Individualism and the Mental, in French, Uehling, and Wettstein (eds.) *Midwest Studies in Philosophy*, IV, Minneapolis: University of Minnesota Press, pp. 73–121.

Clark, A. (2016b). *Surfing Uncertainty. Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Cummins R., (1989), *Meaning and Mental Representation*, Cambridge, MA, MIT Press.

Friston, K. (2013). Life as we know it. *Journal of The Royal Society Interface*, 10, 20130475–20130475.

Grabarczyk, P., 2016, How Meaning Became "Narrow Content", *Studies in Logic, Grammar and Rhetoric*, vol 46, issue 1.

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Kriegel, U. (2008). Real Narrow Content. *Mind and Language*, (23), 305–328.

Kripke, S., 1972. *Naming and Necessity*, Oxford: Blackwell.

Metzinger, T. 2009, *The Ego Tunnel*, New York: Basic Books

Putnam, H., 1975. The Meaning of Meaning, *Philosophical Papers, Vol. II: Mind, Language, and Reality*, Cambridge: Cambridge University Press.

Ramsey W.M., (2007), *Representations Reconsidered*, Cambridge, MA, Cambridge University Press.

Shagrir, O., 2006. Why We View the Brain as a Computer. *Synthese* 153 (3): 393–416. doi:10.1007/s11229-006-9099-8.