Prof. DSc. Sergey Gerdzhikov

REVIEW

On the distration of

Plamen Nikolaev Chergarov

for acquiring the educational and scientific degree "Doctor"
by professional field 2.3. Philosophy (Theory of Knowledge)
On the topic:

THE EPISTEMOLOGICAL EXTERNALISM
IN MENTAL MODELS

Formal data of the dissertation

Dissertation has a volume of 213 pages and consists of an introduction, five chapters and a conclusion.

Bibliography has 269 titles in English and Bulgarian, and an impressive part of them are neurobiology.

The author has two published articles on the topic of the dissertation.

The topic of the dissertation is the mental models analyzed in the light of externalism.

My overall impression of Plamen Chermarov's dissertation is extremely positive. It is a highly informed and creative philosophical text. This is what I need to protect with a brief overview of the content with some notes.

Prerequisite and basic concept

Monism is assumed: brain and mentality, neural structures and mental patterns are isomorphic, which here means a whole, even the same: "mental models can interact with computer models and any other models, but remain distinguishable from them because they are neural models" (Dissertation, p. 78). This is a dominant position in modern science. My agreement is corrected by the important note that a whole does not mean the same thing: the mental is an experience, a phenomenon inaccessible to observation, and neurons are observable structures, and this fundamental difference is irrevocable.

The dissertation follows the project of naturalizing epistemology and its associated strong dependence on natural sciences such as neurobiology. Such is the dominant contemporary philosophical project.

Mental models, a major topic and concept in the dissertation, are defined by the author as "the integrated representations of the world that allow contextualization of information that comes from the world and action that is based on that information" (Autoreferat). The mental model is representation. It is composed of various perceptual and\or semantic elements. It creates a map of the world that allows interaction with the world itself." (Thesis,p. 84)

Thesis:

"The main thesis of the dissertation is that external influences are the main factor in building a model for the world. This is justified by the assumption that "mental" is equivalent to "neuronal". The text contains arguments and empirical data that support this thesis." (Thesis, p. 189)

Thus formulated, the thesis is not new. Here comes a specific thesis, one of the main ones in the work: "Mental models are influenced and shaped according to this transactional activity with the world, and in particular external influences are a key part through which we can understand the constitutive and functional nature of mental models. Fourth, the relationship between the mental model and the neural structure is isomorphic." (ibid.)

Tasks:

The first task is to reconstruct briefly the debate between internalism and externalism. The second is to analyze the concept of "mental model". This task includes the conceptual analysis of the concept of "model" and why it can be assumed that the mental model is identical to a neural model. The third is related to demonstrating the influence of external factors on the expression of the model and how it functions. The fourth is to show that there is an intellectual virtue tied to mental patterns that is of the externalist type.

Methods (techniques):

I would say the techniques used in the dissertation are conceptual analysis, scientific analysis or "interdisciplinary approach", the attraction of facts and concepts from cultural anthropology, neuroscience, psychology, etc., and analogy.

Historical overview:

The positions of internalism and externalism and their sub-positions and aspects on the nature and value of knowledge and cognitive virtues are presented concisely. The debate is also traced, with four important topics, such as the endless regression of justification. I do not find quoted the work of Anna Ivanova Believing Rationally, 2020, which studies in detail the debate between internalism and externalism and offers a solution in the direction of activity. I do not see here two new Bulgarian analyses of the debate: The monograph of Anna Ivanova: *Believing Rationally* (2020) and Madlen Elchinova: *From the Inside Out*, 2022.

Mental model - the concept:

Looking at mental models in a purely externalist framework can be considered a major contribution of dissertation, according to Plamen Chergarov. After Chalmers and Clark, the author assumes that mental patterns are amplified or weakened by cultural narratives and technology. The very concept of a mental model first arose in Crake's work

(1943), but was developed as a theory by Johnson-Layard (1983). One of the definitions given for a mental model is (Jones, R. et al, 2011, p.1): "Mental models are personal, internal representations of the outside world that people use to interact with the world around them. They are created by individuals based on their unique life experiences, perceptions and understandings of the world. Mental models are used to think and make decisions, and can be the basis of individual behavior. They provide the mechanism by which new information is filtered and stored." (Thesis, p. 82).

Similar concepts such as that of Hawkins and Kurzweil: Jeff Hawkins, 2004: "The brain uses a huge memory where it builds a model of reality. In this model, he predicts what will happen. That's intelligence" (Jeff Hawkins, *On Intelligence*, 2004). Ray Kurzweil 2012: "The main element (in the neocortex) is the *pattern* recognizer, and it is the fundamental component of the neocortex." (Kurzweil, *How to Create a* Mind, 2015 p. 48) and that's what generative AI models are built on.

By the way, the successes of generative AI (language machines) are *a confirmation of the great concept of models* in cognition and language.

The mental model creates order in the flow of information. This also makes the scientific model or scientific inquiry, but also some non-scientific forms. The world can only be represented within a mental model, which means that much of it is ignored." (Autoreferat, Dissertation 84 et seq.) (Damasio is quoted.)

Note: This can only be done if and to the extent that agent S understands and accepts models, narratives and technologies, if they have become internal. This concept of pattern is relevant to the concept of "pattern" or "pattern" of Hawkins and Kurzweil, as well as generative models of AI. A similar concept of relativity and relations-matrices is present in my study "Scientific Research as a Relational Ordering" (Gerdjikov 2017)

The dissertation contains numerous references to experiments on humans and animals and builds conclusions on them. "The laws of perception are such that the organism will give the physical environment a certain form (Gestalt) or organizational quality (Gestalt-Qualität) (Robinson (1995). (Thesis, p. 86). An example is vision and maps in the visual cortex.

Crucial argument for externalism: "Unconscious information processing is supportive of the externalist project. The theory of the adaptive unconscious is an appropriate addition to the understanding of mental patterns, and the reasons for rejecting internalism are growing" (Dissertation, pp. 94 et seq.)

In general, *Plamen Chergarov's work on experimental data and generalizations in neurobiology has the argumentative power and this is for me a real achievement of the dissertation*. Only few philosophical theses here are so well informed and analytical about natural sciences.

Cultural influence is naturally noted and examined in spheres and examples, but it is even underestimated: our language itself is the backbone of the culture in which we

are immersed. Everything we know and can do is dependent on and influencing culture. "Technology shapes our models for the world" (Autoreferat). This is an obvious truth.

"The reliability of mental models is constituted in a manner external to the mental model itself (P1). When different mental models lead to the same result, that result is more reliable (P2). If different experts make arguments in defense of the same thesis, the probability that this thesis is true is more likely than if only experts from one field defended it." (Autoreferat). On this point in the text I find external information to science.

Plamen Chergarov justifies with scientific experiments that mental models such as maps of the world, isomorphic to neural networks, are mostly managed *automatically, unconsciously and without access for introspection*.

Important is the thesis of some neurobiologists about the self-organization of neural models (e.g. Soren Ventegodt: "we propose a new model for brain functions, i.e. information-driven self-organization of neural models, where information is provided by the abstract integrity of an organism's biophysical system (often referred to as the true self, or 'soul') (*The Scientific World Journal* 8:621-42)).

Culture and mental models

The premise is that "culture is the transmission of ideas and symbols."

The fact is that "the cultural narratives and technologies we use either enhance or lower the epistemic value of our mental models. Because these things are external to S, they fall under the explanatory framework of epistemic externalism." Individual judgments are formed by group judgments. It turns out that 75% of people give the wrong answer when there are dummies in the group who unanimously give a different answer than the right one.

Everything written in this part is true, but not enough. Note: The natural local language as a matrix of culture and the cognitive patterns embedded in it in the form of grammar and vocabulary have not been analyzed. There are semantic notes on language: the influence of messages on mental patterns. This is how mass delusions and prejudices are formed, such as racism and then anti-racism. The main Western narratives such as Christianity and the logos are also mentioned. The topic is analyzed in detail and is a strong argument in explaining mental models. "Repeating a particular story in the community creates a map of the world that interprets information only in certain categories." (Thesis, p. 113)

Technology and Reading

The brain is looking for new information, and this determines the Internet and new digital technologies as the main sources of mental models. The abundance of information is a challenge to attention and to the connectivity of cognitive maps. The quality of memory and intelligence decreases in the saturated virtual environment. Genes and environment complement each other in the formation of the mental patterns of the world. It is mentioned vaguely that language and thinking interact. In the topic of reading,

I can add that we read holistically, and a text written in misplaced letters is still read successfully – an argument in favor of models. "Our intelligence depends on the preliminary 'cognitive patterns' (here these are mental models) that we learn over time. These schemes are complex concepts that arrange scattered information and give it new associative relationships" (Sweller) (p. 142)

A number of experiments on cognitive processes such as using software have shown how the brain does better with less help and substitution from digital technology, and the explanation is that it builds stronger cognitive circuits. These technologies weaken slow and higher processing and deeper emotional states such as empathy. The metaphors of technology influence our basic patterns of the world and its elements.

Reliabilsm and mental models

"Reliabilism is a system that seeks naturalism to give the necessary epistemic status to beliefs so that they can be accepted as knowledge" (p. 161). The cognitive reliability of mental models, according to the author, depends on alternative mental models with the same result and on the consistency of the results over time – reliable yielding of the result. An example is the multitude of templates in the chess game. Chess is also an example of the complementarity of natural and artificial intelligence such as best solution compared to the two separately: "The application of model A, which is evaluative, to model B, which is computational, creates a better ability to play chess than A and B separately" (p. 167).

Mental models are interdependent and applied in other realms on the basis of "abstract thinking." "Mathematical models are used outside mathematics, historical models outside history, and physical models outside physics."

Here is an important conclusion: "The problem with the appropriate mental model is this: sometimes we fit the world to our model and not the model to the world. This is an epistemic vice, so far as one can be derived. That is why diversity or breadth of mental patterns is such an important part of cognition. Because they can reflect the world accurately enough to talk about knowledge. The criterion for this is external to the separate model" (p. 176). Ethics and the theory of knowledge are associated through values and virtues.

In a number of cases, Plamen Chergarov formulates symbolically logical theses, arguments and models, which is another plus in the dissertation.

There are many more publications on neurobiology that are usable on the subject, but it is understandable that they are not included due to volume limitations and requirements for a doctoral dissertation.

Conclusion

"Consideration of mental models in a purely externalist framework can be taken as the main contribution of the dissertation. At the time of writing the dissertation, there

is not a single title that deals with internalism or externalism in relation to mental models." (I can't say that for sure).

The author concludes that the objectives of the study have been met:

"The debate between internalism and externalism was outlined and the notion of a mental model was introduced into it. A virtue that is not an internalist is successfully demonstrated. The result of the study was realized thanks to the methods that, due to the many relevant fields of research, are interdisciplinary." (Autoreferat)

Requested contributions:

„… arguments of a conceptual nature for the acceptance of the explanatory framework of externalism in epistemic virtues, justifying, justifying factors, and knowledge.

A solution to the infinite regression problem is proposed, which does not undermine the positions of externalism.

A conceptual analysis of the concept of "mental model" has been made.

The dissertation constructs the argument that culture, which is external to the knower, has a huge influence on mental models and therefore their explanation must be in an externalist framework.

The influence of technology on mental models with the same result as in the previous point is considered.

Aretic Realiabilism has been shown to be appropriate in explaining epistemic virtue as an aspect of mental patterns." (Autoreferat)

These claimed contributions are acceptable and adequately formulated.

My conclusion is:

Plamen Chermarov's dissertation is a highly informative, scientifically based, philosophically grounded work and responds educationally, philosophically and scientifically to the degree of **Doctor of Philosophy**.

That is why I most strongly suggest that the PhD student receive this scientific degree for his Thesis.

Prof. DSc. Sergey Gerdzhikov

9.05.2024