The minutes of the dissertation defence

Student:	Martin Vraný
Datum narození:	16. 7. 1983
ID studia:	315032
Studijní program:	Filozofie
Studijní obor:	Filozofie
Název práce:	„Naturalizing the Unity of Consciousness: can neuroscience explain a fundamental feature of subjectivity?“
Jazyk práce:	angličtina
Jazyk obhajoby:	angličtina
Školitel:	prof. RNDr. Jaroslav Peregrin, CSc.
Oponenti:	Mgr. Tomáš Marvan, Ph.D. Dr. David Ward
Datum obhajoby:	23. 3. 2018
Místo obhajoby:	FF UK, nám. Jana Palacha 2, místnost č. 224
Termín:	řádný
Předseda komise:	prof. RNDr. Ladislav Kvasz, DSc., Dr.
Persons present:	prof. RNDr. Ladislav Kvasz, DSc., Dr. doc. Vojtěch Kolman, Ph.D. doc. James Hill, Ph.D. Jan Palkoska, Ph.D. Dr. David Ward doc. Tomáš Marvan, Ph.D. Bc. Filip Rut Mgr. Andrej Virdzek Mgr. Jiří Tříska

12:00 The chair of the board prof. Ladislav Kvasz started the defence, introduced himself, members of the board present and opponents present. Then he introduced the student to the persons present.
12:02 The supervisor introduced briefly to the persons present his evaluation of the student's study and dissertation, given in his written report. He stated also the following:
"The dissertation is more of a work in the philosophy of mind on border with the cognitive sciences, than my main field of specialization, that is, the philosophy of language. There are important parts though, which are directly connected to the problems of language. As far as I can judge, Vraný wrote an excellent dissertation and he manages to bring important contributions to the field. I judge this dissertation to be of the required quality for the title of Ph. D., which I recommend to be granted to Martin Vraný."

12:03 The student presented the propositions of his dissertation to the persons present. He said in particular:

"The concept of the unity of consciousness is historically an important problem for philosophy, yet it is not elaborated as a central explanandum in the natural sciences when coping with the problem of consciousness. I focused on this concept, also with a goal of assessing how successfully can the contemporary leading neuroscientific theories of consciousness explain it and if they fall short of this requirement to try to account for those aspects of the unity which the leading theories cannot explain. These are the main questions which I tried to elaborate in the presented dissertation. I believe that the unity of consciousness is actually the central problem of the whole scientific quest for consciousness, because the theories that cannot account for it or ignore it are either philosophically naive or lead to the so-called Cartesian materialism. Another reason is that the discourse is still dominated by the research programme of David Chalmers – with his formulation of the hard problem. There are more senses in which we could say that consciousness is unified, but I am interested mainly in two – the unity of conscious contents at a time and the unity in the sense of a single subject having conscious contents and being able to reflect on them. I claim that unity is an essential property of consciousness. If we hold that the adaptive function of consciousness is to allow for context-sensitive, flexible action, then the capacity for integration of information from various domains is essential for consciousness. But the explanation of information integration is an empirical question and in that sense some may view the problem as not philosophical. I have worked with Kant's exposition of the unity in *Critique of Pure Reason* as a comprehensive account of consciousness which provides us with actual requirements on the contemporary leading neuroscientific theories. Those proposed theories that can fit these requirements are then judged to be better than the others. After an analysis of the question of reference of first personal pronoun "I" and whether it can be analysed without a recourse to indexicals I specify more requirements for the neuroscientific explanation of the unity in question. With these requirements in mind, I review the three leading neuroscienitific theories of consciousness. The global neuronal workspace theory, the predictive coding theory and the information integration theory. These have some similarities, but in the light of the question of unity, which is the central one, I conclude that the predictive coding theory offers the most convincing account of consciousness. It accommodates a philosophically robust account of representation. It provides a unified account of cognition and action and it provides a compelling

account of bodily self-awereness and perspectival self-consciousness. Since all of the main weaknesses that could be held against the predictive coding theory are shared by the other two theories I conclude that it is the most viable alternative present."

12:21 The opponent Dr. David Ward presented the main points of his report and the conclusion, stating that he recommends the submitted dissertation for defence. He put these questions to the student:

"There are three main points, that I would like to hear about, first of all, how does the argumentative structure of the dissertation as a whole hang together?, second, in chapter 3, you claim that the practical unity is either necessary or sufficient for the content unity, but can you elaborate on why shall we think so, and third, what is the role of the cultural and linguistic scaffolding in explaining the unity of consciousness? Isn't the normative unity of consciousness built into our practices, for example, in the language game of giving and asking for reasons? There are also a few specific points where I would like to hear more, but I don't insist on it. Does an explanation of consciousness need to have naturalistic credentials (f.e. biological adapatations)? In chapter 5 you say Hurley doesn't give us a complete explanation of unity, in what particular sense does she lack a positive theory? How does information integration theory fit in with the story of metarepresentation which you obtained from Kant and Hurley? I don't expect all of the questions to be answered here, it also may be, that they will be answered when other points are elucidated and some of them were already answered in the overview. I believe this to be an extremely impressive and important dissertation, which raises a lot of interesting questions."

12:32 The opponent doc. Tomáš Marvan Ph. D. presented the main points of his report and the conclusion, stating that he recommends the submitted dissertation for defence. He put these questions to the student:

"I reiterate the point that the dissertation is not completely clear in the most technical details. The questions that I want to raise are these: a) Must the naturalistic account of consciousness be based on the assumption, that consciousness has an adaptive functions? I believe this is a question already asked by dr. Ward, and a case can be made that consciousness is explainable as a spandrel. It might turn out that consciousness has no adaptive functions of the kind mentioned in your work, b) I want to ask about the notion of Cartesian Materialism, I never really understood the criticisms by authors who look for naturalistic explanation of consciousness, the attacked idea is supposed to be, that

there is a place in the brain where the consciousness as a whole occurs, but most of empirical theories of consciousness use this idea, even those who criticise it. What is the problem with such a picture?, c) How can the two verions of the predictive coding theory be put into one comprehensive story? On one hand it says, that the higher-level areas try to predict the activation at lower level areas, on the other the higher level areas are supposed to model the causes of the lower-level areas. I find the latter to be much more ambitious and in need of further explanation. All in all, I fully recommend that the title Ph.D be granted to Martin Vraný on the basis of this dissertation thesis, which is a valuable addition to the field of consciousness studies."

12:40 The student reacted to the supervisor's comment: There was no need for a specific reaction because supervisor's comment was wholly positive and didn't raise any questions that were not asked by the opponents.

12:40 The student reacted to the opponents' comments and their questions:

Martin Vraný: I will start with the question of the adaptive function of consciousness, as it was raised by both opponents and can be viewed as most pressing. The question is contentious. If I hold the position that the project of reduction can be granted without consciousness being viewed as a functional adaptation, then what about the open question argument? Anything that the actual neuroscientist will discover can be tackled by the philosopher with the open question argument. This could be used by people such as Chalmers who would ask: Why does X give rise to consciousness, where X stands for a neural or, in general, a physical process. Without the functional description of consciousness the project of finding a naturalistic account is futile. Even the paper that doc. Marvan quotes, starts with the sceptical problem, that is demonstrated by the open question. The claim that consciousness is an adaptive function has the value that it provides us to start the project in a naturalistic sense when it makes sense to ask about what gives rise to consciousness. So the question that remains is how to define consciousness, which has as its essential property unity, in the evolutionary sense, in the sense that makes sense of it as having the adaptive value. Conceptual possibility of zombies becomes rather unimportant, because there is a naturalistic explanation of the causes of consciousness, even though it may be claimed, that information integration is not sufficient, but only neccessary for consciousness.

David Ward: It (consciousness) must be explained in a naturalistic way, but it is not necessarily an adaptive function, you can give a naturalistic story of how the system ticks the boxes for what is called a unity of consciousness. But does the confluence of the parts have an adaptive function?

What does have the function, the parts or the whole system? Why commit yourself to the view which needs more explanation, but doesn't neccessarily grant better result?

Martin Vraný: Consciousness is not something that appears only due to a contingent interplay of various functions which are evolutionary independent of each other. At the lowest level consciousness starts with guidance of the behavior, which better be context-sensitive. On the higher level, it makes possible for the subject to distinguish betwen how things are and how they seem to be. The integration of information is central here and needs to be explained, if we are to understand the adaptive value. These structures that constitute consciousness are not independent of each other, not from the viewpoint of its function, there they are explainable as a whole.

Regarding the question on predictive coding theory and the probem of causes vs. predictions: One of the claims has a statistical meaning where, if you feed the model some data, the output is a prediction. It is a prediction in the statistical sense and it has nothing to do with anticipation in the ordinary sense. In some cases the brain tries to predict its future states, but these are two different things. I can demonstrate the claim made by proponents of predictive coding theory on a simple model. Imagine a particularly simple environment. Simple square, simple circle, edged square, and something in the shape of a star. We can have a pixelated representation of this. Now our question is: How many neurons do you need to code ten bits of information (in these example ten bits can represent the field)? This is an abstraction that the model does of the features that appear at the retina. These models learn to reconstruct, or classify the inputs on the layer. Using the information that is abstracted in the second layer, I can have a representation that uses fewer logical nodes, I can see a square a circle etc., but with more efficient coding. But this kind of mechanism does not really uncover the true causes in the world, it is an abstraction. The brain, or any neural network, will arrive at a simplification of the information given, and it will adjust to the number of neurons it has.

Tomáš Marvan: But this is still the statistical story, it needs to be combined with the anticipatory story.

Martin Vraný: This anticipation is not the classical anticipation we experience, it happens in miliseconds, it is under the threshold of what we can experience as happening.

Regarding the question about Cartesian Materialism - doc. Marvan claims, that global neuronal workspace theory says, that we can precisely locate, when and where in the brain something becomes conscious. I still can interpret the global neuronal workspace theory, as not neccessarily

claming this to be the case, or rather that the locating does not need to happen at a precise place in the brain. Information becomes conscious by the top-down amplification in the brain. Whether something is part of the workspace is, then, a property of the whole brain. Dennett would endorse the claim that consciousness is produced by brain, he is only against specifying a particular part of the brain. Microtiming the arising of consciousness is also not possible, but then consciousness does not arise completely instantaneously, the feedback and feed forward reaction happen in extended time. There is then the minimal bound of time where one becomes aware and there is the maximal bound of making a verbal report, but in the time inbetween I don't think it is possible to locate when exactly consciousness has arisen.

13:04 The opponents (and the supervisor) commented on the performance of the student.

David Ward: I am satisfied with the answers.

Tomáš Marvan: I am also satisfied with the answers given.

13:05 The chair of the board opened the discussion.

These persons spoke in the following discussion: Prof. Jaroslav Peregrin, David Ward, doc. James Hill

Jaroslav Peregrin: My question is more general, when you use a naturalistic approach to this problem my suspicion would be that the question that Kant asked evaporates, but you seem to be saying that the Kantian question can be naturalized? Doesn't it then become a different question?

Martin Vraný: I use Kant as someone who specified the transcendental conditions of consciousness. Kant's transcendental psychology could be understood in the naturalistic perspective, in cognitive scientific terms and the question of unity of consciousness is part of this approach as I understand it.

Jaroslav Peregrin: So you agree that the question can be naturalized?

Martin Vraný: As long as Kant's conditions are put in terms that can be satisfied by a scientific theory, yes. But, I must conclude that during the process of writing this thesis, the question evaporated in a different way.

David Ward: Why do you think the question becomes irrelevant?

Martin Vraný: In the beginning of my research, I thought that the unity has some phenomenal quality, that can be reflected from the first-person point of view. But this is rather the form of consciousness, it is not something that can be experienced as a transcendental property. All that

needs to be done then is that the problem of integration of information has to be solved, but I don't see this as a strictly speaking philosophical problem.

James Hill: Do you need language to be conscious, self-conscious? I am interested in the relation between language and consciousness. Are you committed to the idea that there cannot be a unity of consciousness without language, and aren't you excluding other living beings from being conscious?

Martin Vraný: The short answer is no. There is a good sense of talking about neural representation without the verbalization. Kant and Hurley have shown the fact that we can specify the notion of consciousness, where the contents are first unified without a linguistic influence, but I still find the influence of language, and cultural scaffolding strong. I don't think though, that the dictum "I think" must be represented via some linguistic articulation. "I think" refers to being aware of our representations as representations, and that is manifested by an organism that can recognize something as a perceptual error. This capacity is what brings some rudimentary kind of self-consciousness, this is a viable interpretation of "I think" without language. Therefore, I think that linguistic scaffolding is not fundamental enough.

13:16 The chair of the board ended the defence itself and the board initiated a private meeting on classification of the dissertation defence.

The chair of the board announced to the student and persons present the result of the defence: The board voted using ballots), number of the members of the board 5 – number of the members of the board present 4 – positive votes 4, negative votes 0. The dissertation defence was classified pass.

The chair of the board on behalf of the board stated that the dissertation is unsatisfactory in no ways.

The board determined that the dissertation should be revised/supplemented as follows: X

Recorder: Mgr. Andrej Virdzek

Name and signature of the chair of the board:

Name and signature of another member of the board: