

Univerzita Karlova

Filozofická fakulta

Ústav filosofie a religionistiky

Filosofie

Teze disertační práce

Naturalizing the Unity of Consciousness: can neuroscience explain
a fundamental feature of subjectivity?

Naturalizace jednoty vědomí: mohou neurovědy vysvětlit zásadní
rys subjektivity?

Autor: Mgr. Ing. Martin Vraný

Školitel: Prof. RNDr. Jaroslav Peregrin, CSc.

Rok: 2017

1 INTRODUCTION

The aim of the dissertation is to analyze the concept of the unity of consciousness as an explanandum for natural sciences and assess how good an explanation do leading neuroscientific theories of consciousness provide. The motivation behind this project is the idea that it is the unity which poses the greatest challenge for the scientific quest for consciousness. I put this as a point of contrast to the influential view raised by that the hard problem of consciousness consists in explaining the quality of experience.

1.1 Cartesian materialism debunked I argue that the reason why some theories of consciousness lead to what Dennett calls Cartesian materialism is precisely because they fail to address the problem of the unity of consciousness. If we had a good understanding of the unity of consciousness and its place in nature, we could more easily avoid the tendency to devise accounts of consciousness that are homuncular in disguise. I recount Dennett's analysis of why we tend to fall for Cartesian materialism and focus on the intuitive idea that there are facts of the matter of how things seem to us independent of what we *think* about how they seem to us. I argue that while Dennett's rejection of this intuition is correct, his analysis of the reasons that foster this intuition is incomplete. Specifically, he misses an important feature of self-reflective consciousness, namely self-reference without identification, as Shoemaker put it: we know. non-inferentially, the identity of the reflective subject and the subject of the object thought. I also introduce the concept of transcendental self-consciousness and argue that Dennett underestimated the importance of this concept the consequence of which is that his explanation of consciousness feels to be incomplete.

1.2 Homuncular subject and the unity of consciousness In this part I analyze what explanatory role does the concept of subject play in both folk-psychology and cognitive science. The point of this analysis is that arguing that many of the phenomena that motivate using the concept of subject could be explained if we had a good account of the unity of consciousness. The advantage of speaking about the unity instead of the subject is that the former does not lead so easily to the homuncular picture of the mind.

1.3 Overview of the following work A shorter version of this summary.

2 CLARIFICATION OF THE CONCEPT OF THE UNITY OF CONSCIOUSNESS

There are three dimensions along which one can make distinctions regarding the unity of consciousness: 1) temporal - whether the unity is considered as the unity at a time or over time, 2) qualitative - whether the unity concerns access or phenomenal consciousness, and 3) structural - what are the elements that form the unity.

2.1 Synchronic and diachronic unity of consciousness A distinction is made between the unity of conscious contents at a time (synchronic unity) and over time (diachronic unity of successive states of consciousness). Diachronic unity is important for questions related to personhood and can be thought of as a matter of *representing* and ascribing the succession of conscious states to a single empirical subject. However, even momentary consciousness arguably extends over a short interval of time (in the sense that Husserl introduced using his concepts of retention and protention) and continuity of contents in this sense cannot be secured by a representation of the continuity.

2.2 The unity of access and phenomenal consciousness Following a distinction introduced by Block (1995), it should be specified whether the unity pursued is that of access or phenomenal consciousness. Here I recount Block's definition and touch on the question whether access and phenomenal consciousness are empirically independent or only conceptually independent, as Chalmers (1997) suggests.

2.2.1 Access consciousness: integrated representation I specify what the unity of access consciousness means: the integration of contents in joint control of behavior. I defend this interpretation against that of Bayne & Chalmers (2003) who, in their definition, refer to the subject. Such reference is to be avoided when we study the unity, for explaining the unity in terms of a subject accessing various contents would confer the unity trivially. Consequently, I argue that avoiding reference to the subject in characterization of the unity of consciousness renders the unity problem non-trivial. I also argue that studying the unity at the level of *access* consciousness, rather than phenomenal consciousness, is to be preferred because discussion of how phenomenal properties of mental states are integrated into the phenomenology of the total conscious state presupposes that the phenomenal properties are identified in virtue of content anyway.

2.2.2 Phenomenal consciousness: one experience or many? Despite the criticism in the previous section, I try to spell out what the unity of phenomenal consciousness might mean. Following Brook & Raymond (2017), I distinguish two alternative views: the experiential parts view, according to which the phenomenology of total conscious state is composed from phenomenal properties of constitutive states; and non-experiential parts view that states there is only one global state of which we are phenomenally conscious. I argue that neither view makes a good sense for the problem of the unity of consciousness. The experiential parts view faces the problem described in the previous section, namely how phenomenal properties of the constituents are identified. The non-experiential parts view renders the unity problem trivial (there is only *one* experience at a time).

2.3. The unity of conscious state The structural dimension invites four distinct aspect of the unity: the objectual unity, spatial unity, subject unity and integration unity. The objectual unity refers to ascribing various recognized features to a single object. The spatial unity is a matter of location objects in space. Subject unity refers to the sense in which we say that consciousness is unified because there is a single subject having multiple representations at a time. Integration unity refers to the idea that our consciousness at a time is composed of a multitude of representations that are coherent and integrated into one global state.

2.4. The unity analyzed I state that in the rest of the dissertation I will focus on the synchronic unity of access consciousness. In the structural dimension, I pursue both the subject and integration unity because these are two sides of the same coin and they jointly present the greatest challenge for science.

3. NORMATIVE AND OBJECTIVE ASPECT OF THE UNITY OF CONSCIOUSNESS

The unity is an essential property of consciousness. This claim can be interpreted in two ways. The first, normative way is that we cannot make sense of a consciousness that would not be unified. When we judge the extent of one's consciousness, we apply the norm of rationality of one's behavior. What we think somebody is conscious of is thus constrained by what we can accept as rational behavior given the circumstances. The second, objective way is that consciousness could not serve its evolutionary function (integrating contents to allow for context-sensitive action) if it were not unified - the unity is thus a constitutive feature of consciousness.

I recognize that the specification of the objective sense of the unity refers to contents and that these, in turn, cannot be directly identified with brain states (I endorse multiple realizability, or generally externalism about content). Despite this, I hold that the unity is an empirical problem to be addressed by science. I outline the defense of this claim that is the subject matter of the rest of this chapter.

3.1. Function(s) of consciousness A naturalistic account of consciousness can be convincing only under the assumption that consciousness has some adaptive function. Scientific description of a process can count as an explanation of consciousness only if we have its functional

specification and the empirical research shows the process to fulfill that function. I recount the functional specification offered by Baars (1988).

3.2. Representation in the brain Using the content/vehicle distinction, it is correct to say that brain states are vehicles of conscious contents, not contents as such. What justifies talking about representation in the brain then? That the concept of neural representation should be justified follows from its heavy use in current neuroscience: all neuroscientific theories of consciousness rely on the concept of neural representation, in some sense. Reviewing work by Smolensky (1995), Ramsey et al. (1990), I present arguments against representation in connectionist networks, show how it extends to neural representation, and then reject them.

3.2.1 Neural correlates I discuss in greater details typical research methodologies in cognitive neuroscience that yield conclusions about neural representations. The point is to show that neuroscientific methods allow for establishing more than just neural correlates - contrary to popular belief in philosophy. For this sake, I briefly outline basic neuroscientific concepts: tuning function, feature space, sparse coding, fitting of computational models, coincidence detectors etc.

3.2.2 Neural representations Even though neuroscience yields more than just correlates, the concept of representation is used to freely. I thus specify conditions that a complex neural correlate must meet to count as representation. These are 1) modular, detachable use, and 2) that an action that is causally mediated by the neural correlate cannot explained in a more parsimonious way than as a result of transformation of representations. I then concede that neural representation is always underdetermined with respect to the experimental manipulation that was used to find it, but argue that it is not a substantial problem for empirical research.

3.2.3 Neural metarepresentation? I argue that there are no reasons to assume that what counts as metarepresentation at the content level would exhibit some specific feature at the vehicle level, i.e. at the neural level. As a consequence, metarepresentation at the vehicle level should be thought of as a higher-level representation of lower-level features in the model of neural cascade.

3.3 Do we need neuroscience to explain the unity of consciousness? An argument is formalized against the idea that neuroscience can explain the unity of consciousness because the latter concerns content and the former studies vehicles of contents. I outline two steps to reject the argument: 1) neuroscience can offer explanation of types (contents) as well, not just of tokens, 2) regarding consciousness, science is not limited to *causal* explanations only.

3.3.1 Hurley and the need for an objective account of the unity I present the just-more-content argument by Hurley (1998). The argument shows that the unity of consciousness cannot be accounted for solely at the content level - an account at the vehicle level is necessary and this is a task for neuroscience.

3.3.2. Science and type-explanatory accounts of the mental I argue that theoretical neuroscience, evolutionary and cognitive psychology do offer type-explanatory accounts of the mental, not just causal, token-explanatory accounts. Natural selection is the key concept that allows science to refer to contents.

3.3.3 Content and type-token distinction I describe what kind conditions an experimental procedure must meet to be relevant for the question of what neural processes constitutes the unity of consciousness. The crucial issue is how to establish that an integrated representation of [A+B+C] occurred, in contrast to joint occurrence of independent (not unified) representation A, B, C.

3.3.4 Self-consciousness and content-vehicle distinction I discuss how to conceive of self-consciousness at the level of vehicles of conscious contents. At the content level, self-consciousness manifests itself as a relation between the reflective and the object thought. At the vehicle level, however, we should not conceptualize it as one state having access to the content of another state, for that would conflate the vehicle/content distinction. I recount arguments by Hurley (1998) to make these conceptual pitfalls clear. Arguments in favor of vehicle externalism (Hurley (1998)) and internalism (Clark (2009)) are presented, the conclusion being that while internalism is more likely, content is still an essential *relational* property of vehicles.

3.4 Summary Brief summary of the points made in this chapter.

4. KANT ON THE UNITY OF CONSCIOUSNESS

4.1 Preliminary remarks Unified consciousness, the synthetic unity of apperception, as a starting point of Kant's B deduction of his CPR. I present a summary of the Transcendental Deduction (TD) by Rosenberg (2005). I argue that while the project of scientific naturalization of the unity of consciousness is incompatible with tenets of Kant's transcendental idealism, there is a substantial line of interpretation according to which Kant's insights into transcendental psychology can be treated separately from it (kitcher1993kant, Brook (1997), Rosenberg (2005)). Kant's transcendental psychology can thus be used to specify the explanandum for cognitive (neuro)science.

4.2 The unitary subject of mental states I summarize Kant's main points about the unity of consciousness (the transcendental unity of apperception). In the unity of consciousness I conceive of myself collectively as a subject of many representations. The unity is a condition of possibility of empirical self-consciousness. The unity is not experienced, rather it is a form of experience. The unity allows for self-reference without identification.

4.3 Unitary experience of the world - transcendental apperception from a logical point of view I present an interpretation of Kant by Strawson (1966) according to which the multitude of representations at a time is unified in the idea of the objective world. Self-consciousness arises from the distinction between how things are independently of how we represent them. Any meaningful conception of experience must assume this distinction to be available for the conscious being, i.e. consciousness entails some minimal sense of self-consciousness.

4.4 Kant's syntheses - transcendental apperception from a psychological point of view According to Kitcher (1993), Kant held that consciousness is unified in virtue of three syntheses. However, Kant claims at B133 self-consciousness requires that I be conscious of the synthesis of mental contents. Alternative interpretations of this claim by Brook (1997) and Kitcher (1993) are briefly discussed.

4.4.1 The threefold synthesis I describe Kant's conception of the three syntheses that jointly constitute unified consciousness and interpret them in terms used in cognitive science. The syntheses are: synthesis of apprehension in intuition, synthesis of reproduction in imagination, synthesis of recognition in a concept.

4.4.2 The synthetic consciousness In order to specify what transcendental self-consciousness means, I address the differences between two ways of interpreting Kant's insights about the mind: the psychological reading of Kitcher and Brook vs the epistemological reading of Strawson or In Keller. The main conclusion is that the transcendental self should not be thought of as an object with properties (while empirical self can be, e.g. as an ego or person).

4.5 Two kinds of self-awareness I clarify the distinction between empirical and transcendental self-consciousness. The latter does not yield knowledge in the Kantian sense, it is best interpreted as that aspect of the mind's conceptual system thanks to which it can conceive of things as independent of its representations of them.

4.5.1 Brook's account of the transcendental unity of apperception To further clarify the distinction between empirical and transcendental self-consciousness, I discuss Brook's distinction between empirical self-awareness and apperceptive self-awareness.

4.5.2 Global representation I present Brook's interpretation according to which the Kantian unity of consciousness is to be conceived of as a global representation that integrates constituent representations in a specific way. This interpretation in terms of representation helps to clarify the unity as an explanandum for cognitive science, for reasons discussed in chapter 3.

5 SELF-REFERENCE AND SELF-AWARENESS

In this chapter I discuss recent philosophical accounts of self-awareness.

5.1 The unity of consciousness: a working definition I expand on the following definition of the unity as "synchronic* connectedness of representations in a global representation such that the representing subject is thereby transcendently aware of being the representing subject." The explanandum for scientific explanation of the unity consists of two areas: 1) integration or synthesis of representations, and 2) transcendental self-consciousness.

5.2 Shoemaker and Castaneda: the logic of 'I' Shoemaker (1968) and Castaneda (1966) analyzed the logic of the first-person pronoun and addressed the question whether 'I' is a referring expression. The distinction between subject and object use of 'I' is introduced - the subject use is philosophically more interesting and relevant for the unity problem. The subject use requires that I be able to refer to myself without having to recognize myself via by some properties - a feature called self-reference without identification. Thus by analyzing the logical conditions of the use of 'I', Shoemaker and Castaneda argue that transcendental self-awareness is presupposed by it. The explanandum of what we identified earlier as transcendental self-consciousness can thus be specified as explaining a crucial feature that the subject use of 'I' exhibits: immunity to error through misidentification.

5.3 Hurley's two-level interdependence model I present a model of consciousness by Hurley (1998) which builds in the idea that the distinction between action and perception does not correspond to the distinction between output and input to the brain. I explain Hurley's conception of the self using her concepts of dynamic singularity and perspectival self-consciousness. The self is, so to speak, the center of feedback loops between action and perception. The unity of a subject or agent could potentially be explained in terms of a third-personal description of a dynamic system (the organism and its interaction with the environment). To support this idea, Hurley describes neurological cases showing how what we do (output) or intend to do changes what we perceive without any mediation from the environment, how our actions depend on constant proprioceptive feedback etc. Embodiment and agency are crucial for something to have a perspective and therefore some minimal self-consciousness (i.e. awareness of what changes in inputs are self-caused, e.g. by movement). Hurley's account is an important step in naturalizing the unity of consciousness because it shows what it takes to be an agent using concepts familiar to cognitive science, without alluding to the notion of subject or any other notion typically used in the first-person descriptions.

5.4 Empirical self-consciousness Hurley's concept of perspectival self-consciousness does not cover the full-fledged reflective self-consciousness that we are capable of. I try to account for the capacity to reflect on one's content of consciousness from a naturalistic stance by describing a process of cognitive adaptation that possibly led to it. A prerequisite for this is to distinguish between intentional and cognitive access. Intentional access refers to the capacity to spontaneously act using the accessible information, given one's intention. Cognitive access, on the other hand, is a matter of explicitly recognizing my representations *as* representations. Reflective self-consciousness is a matter of cognitive access, not intentional access.

5.4.1 A short evolutionary story of the origin of self-consciousness I argue that the reason for developing cognitive access and hence reflective self-consciousness is the ability to track perceptual errors and imagine counterfactual situations to plan actions.

5.4.2 Self-evidence, cognitive access and redundancy I argue at length why the reflective state ought be considered as a higher-order representation of the reflected state, i.e. representation that does not contain the object state as its proper part but rather abstracts from its details and conceptualizes it.

5.4.3 Return of the perceptual model of self-reflection? I address a potential objection to the previous account of empirical self-awareness, namely that it relies on a perceptual model of self-reflection that I rejected earlier. The key point is to recognize that the higher-level reflective state cannot be identified with the subject and hence there is no sense in which something 'perceives' the lower-level state. Also, the account does not assume that the reflected state is fixed or independent of the reflective state, unlike the perceptual model of self-awareness.

6 UNITY OF CONSCIOUSNESS IN COGNITIVE NEUROSCIENCE

6.1. Global workspace theory I describe the global workspace theory as formulated by Baars (1988) and Dehaene & Naccache (2001). I discuss the key concepts of global workspace, hierarchy of contexts, sustained activation, and modularity. The global workspace architecture has been proposed as a solution to the conflicting demands that the brain needs to meet: make decisions (actions) fast and at the same time guided by as much of relevant information as possible. I argue that the concept of relevance or informativeness plays a crucial role in explaining consciousness but is not itself sufficiently explained in neural terms. For in the end it is relevance that is referred to when describing what guides the competition of specialist modules for access to the global workspace.

6.1.1 The access unity according to the global workspace model The GW theory holds that the unity of access consciousness is a matter of a set of neural representations forming a pattern (loop) of sustained activation: for a short period of time, the content of consciousness would correspond to the representations that mutually support each other (i.e. the activation that is considered to be the vehicle of the represented content). The GW theory assumes that each brain instantiates *one* global workspace because that is the optimal architecture for controlling behavior of *one* organism.

6.1.2 The subject unity according to the global workspace model I present an account of the subject unity congenial to the GW theory. I first present a strong but illustrative account of the subject unity and then relax some the assumption to make the account more plausible but also more complicated.

Empirical self-awareness is a matter of explicit representation of properties of the self via modules subserving introspection (a self-related content enters the GW). Transcendental self-awareness, on the other hand, refers to the capacity to form a higher-order thought about current contents of consciousness. Self-reference without identification could then be explained in terms of a higher-order representation being sustained by the activity of the systems in the

global workspace in virtue of representing their content. From this model it follows that we are self-aware only if redescribing the content of consciousness is somehow informative for the current global workspace, for example if we need to monitor and check our inferences.

6.1.3 Objections to the subject-unity account I discuss few potential objections to previous account: 1) that it conflates the empirical and transcendental consciousness, 2) that the assumption of a monitoring module implies the possibility of single dissociation between consciousness and self-consciousness, 3) that a common language of thought needs to be assumed, 4) that the monitoring module is effectively a bottleneck of consciousness. Discussing these potential objections leads to a richer account of the subject unity according to the GW.

6.2 Predictive coding

6.2.1 Main principles Drawing on Clark (2013); Friston (2010); Penny (2012); Ma et al. (2013), I explain basic tenets of the predictive coding theory: the brain as a Bayesian inference machine; predicting sensory inputs via hierarchy of generative models of the world; prediction, prediction error and precision weighting; free energy formulation; uncertainty and long-term minimization of surprisal.

6.2.2 Attention as precision-weighting and related problems Predictive coding is primarily a theory of cognition, not consciousness. I discuss the main PC interpretation of consciousness by Hohwy (2012) which identifies its content with the hypothesis that is guided by the current precise prediction error. I argue that if attention is defined as increasing precision of the models processing the attended-to channel of information, then precision weighting is supposed to play two conflicting roles: marking the reliability of some feature detectors and attentional amplification. The consequence of this problem is that the PC theory fails to make a convincing account of mental action.

6.2.3 Bodily self-awareness The PC theory provides a detailed and convincing account of bodily self-awareness (Apps & Tsakiris (2014); Seth (2013) : it is a result of matching predictions of proprioceptive and interoceptive signals. The account can be extended to cover Hurley's concept of perspectival self-consciousness as well. The important point is the sense of bodily self-awareness is subserved by many models, each contributing to a specific sense in which an object is interpreted as being part of myself.]

6.2.4 Transcendental self-consciousness I outline a possible account of transcendental self-consciousness in the PC theory. I argue that the idea of the objective world, which was recognized as the ultimate principle of the unity of consciousness in chapter 4, corresponds to the PC idea that the brain minimizes prediction error thanks to having an accurate model of the causal structure of the world. The reason why the brain instantiates only one integrated representation of the world is because competing hypotheses would increase the long-term prediction error. I argue, however, that the architecture achieving this cannot be hierarchical all the way and that the PC theory does not address the question of how the generative models of the most abstract features of the world are organized. I speculate that the use of language as a thinking tool could possibly obviate the hierarchy problem.

6.2.5 Subject unity of consciousness I try to extend the account of bodily self-awareness to explain self-reflection in general and self-reference without identification in particular. I first discuss and criticize a model by Taylor (2012). I argue that self-reflection could be understood as an even in which the articulation of one's mental contents actually matches (produces low prediction error) the contents of consciousness. I draw on works of Karmiloff-Smith (1992); Sellars (1956); Gopnik (2009) to argue that the capacity to articulate the contents of one's consciousness directly develops only later in ontogeny when the child first learns to attribute

mental contents to others by observing their behavior and then learns to attribute them to herself. The reason why language is essential to explicit self-awareness is not only that it allows us to articulate our mental contents, but more importantly that its use fosters formation of the key concept of a subject endorsing a statement.

6.2.6 Predictive coding and Kantian echoes Kant's idea that the world conforms to our knowledge because we cannot but experience it through the concepts and forms of intuition we have is echoed in the idea that the brain represents the world in virtue of modelling the possible causes of its sensory activation. We are experientially open only to those patterns in the world that we can pick up using our senses.

6.3. Integrated information theory and the dynamic core I describe the basic tenets of the IIT as formulated by Tononi & Edelman (1998) and Tononi (2004): differentiation and integration, main complex and dynamic core, and effective information. Technical details are further described in Appendix B. I argue that the IIT gives us clearer understanding of integration and relevance. More specifically, it specifies what it means for a pattern of activation to be part of the global workspace or main complex. I emphasize that the concept of information used in the IIT (Shannon's concept) cannot be interpreted as saying anything about personal-level contents. Nonetheless, to defend the relevance of the concept of information for consciousness science, I point to empirical studies showing how information-theoretic measures relate to conscious behavior.

6.3.1 Contrasts and comparison of the IIT with the GW and PC theories I argue that the IIT specifies the pattern of sustained activation that the GW theory holds as constitutive of consciousness. The IIT also specifies relevance as effective information. Unlike the GW theory, the IIT does not imply that neural activation is necessary for conscious representation. The difference between the IIT and the theory is shown on explanation of consciousness of absence of a stimulus.

6.3.2 The subject unity and the dynamic core I use the IIT concepts to support and elaborate the account of the subject unity presented earlier. First, the IIT implies that a self-reflective thought occurs only if it is informative to the rest of the system, e.g. in case of planning or decision making. Second, the IIT supports the idea that the self-reflective thought is higher-level, abstract representation of the object state. Finally, the IIT implies that a system where there are more areas subserving self-reflection will integrate more information than a system with only one such area.

6.4. Summary Key points of the chapter are that the review theories: 1) are compatible and complement each other; 2) discuss consciousness mostly at the computational level of description with little specification and supporting empirical evidence at the level of neural implementation; 3) try to account mainly for the integration unity - a fitting account of self-consciousness and the subject unity requires a considerable extension of the theories; 4) the predictive coding theory is the most promising as it covers many aspects of the unity recognized earlier in the dissertation.

CONCLUSION

I wrap up the dissertation by presenting the previously discussed points from a wider perspective. Consciousness is a challenging concept for scientific explanation only insofar it entails self-consciousness. The key point of naturalizing the unity is recognizing that 1) consciousness is unified in virtue of a cognitive structure aimed at providing us with a coherent model of the world; 2) since we are embodied agents, the coherent picture of the world is one that enables successful action in the world, hence the picture is *egocentric*. Egocentric representations are the foundation of perspectival self-consciousness. The predictive coding provides currently the

most detailed explanation of how these egocentric representations are realized. However, the explanation is still incomplete and the most significant gap is in explaining the architecture of the information flow among the most abstract generative models that cannot be hierarchical.

In extending the account of perspectival self-consciousness to empirical self-awareness, the key idea is that the higher-level self-reflective thought become conscious if it accurately matches current contents of consciousness and if the higher-level, abstract representation of the contents is informative for the rest of the cognitive system. I speculated that language plays a crucial role in this process - not only because it enables us to articulate the content but also because its use fosters formation of the concept of subject endorsing a statement (and hence a subject of thought).

REFERENCES

- Apps, M. A. & Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neuroscience & Biobehavioral Reviews*, 41, 85–97.
- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Bayne, T. J. & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The Unity of Consciousness*. Oxford University Press.
- Block, N. (1995). On a confusion about a function of consciousness. *Brain and Behavioral Sciences*, 18(2), 227–247.
- Brook, A. (1997). *Kant and the Mind*. Cambridge University Press.
- Brook, A. & Raymont, P. (2017). The unity of consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2017 edition.
- Castaneda, H.-N. (1966). 'he': A study in the logic of self-consciousness. *Ratio*, 8(December), 130–57.
- Chalmers, D. J. (1997). Availability: The cognitive basis of experience? In N. Block, O. J. Flanagan, & G. Guzeldere (Eds.), *The Nature of Consciousness* (pp. 148–149). MIT Press.
- Clark, A. (2009). Spreading the joy? why the machinery of consciousness is (probably) still in the head. *Mind*, 118(472), 963–993.
- Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03), 181–204.
- Dehaene, S. & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79(1), 1–37.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. Random House.
- Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, 3.
- Hurley, S. L. (1998). *Consciousness in Action*. Harvard University Press.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Keller, P. (2001). *Kant and the Demands of Self-consciousness*. Cambridge University Press.
- Kitcher, P. (1993). *Kant's transcendental psychology*. Oxford university press.
- Ma, W. J., Körding, K., & Goldreich, D. (2013). Bayesian modeling of perception.
- Penny, W. (2012). Bayesian models of brain and behaviour. *ISRN Biomathematics*, 2012.
- Ramsey, W., Stich, S., & Garau, J. (1990). Connectionism, eliminativism, and the future of folk psychology. In *Philosophy, Mind, and Cognitive Inquiry* (pp. 117–144). Springer.
- Rosenberg, J. (2005). *Accessing Kant: A Relaxed Introduction to the Critique of Pure Reason*. OUP Oxford.

- Sellars, W. S. (1956). Empiricism and the philosophy of mind. *Minnesota Studies in the Philosophy of Science*, 1, 253–329.
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in cognitive sciences*, 17(11), 565–573.
- Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy*, 65(October), 555–67.
- Smolensky, P. (1995). On the projectable predicates of connectionist psychology: A case for belief. In C. Macdonald & G. F. Macdonald (Eds.), *Connectionism: Debates on Psychological Explanation*. Blackwell.
- Strawson, P. F. (1966). *The Bounds of Sense: An Essays on Kant's Critique of Pure Reason*. Methuen.
- Taylor, J. (2012). The problem of 'i': A new approach. *Journal of Consciousness Studies*, 19(11-12), 233–264.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC neuroscience*, 5, 42.
- Tononi, G. & Edelman, G. M. (1998). Consciousness and Complexity. *Science*, 282(5395), 1846–1851.

ODBORNÉ AKTIVITY

STUDIJNÍ POBYTY

2012-2013 UC Berkeley, supervisor: John Searle (Fulbrightovo stipendium, 12 měsíců)

2015-2016 University of Edinburgh, supervisor: Andy Clark (Anglo-Czech Educational Fund, 6 měsíců)

ÚČAST NA KONFERENCÍCH S PŘÍSPĚVKEM

2011 Res Per Nomen 3, Reims, Francie; příspěvek *Language and the unity of consciousness*

2014 Zimní kognitivní škola, Malá Skála; příspěvek *Metodologická úskalí neurovědeckého výzkumu vědomí*

2014 Hradecké filosofické dny, Hradec Králové, ČR; příspěvek *Snesitelná automaticita bytí*

2016 Non-linear Dynamics in Work and Organizational Psychology: To Non-linear Modelling ... and Beyond, Barcelona, Španělsko; příspěvek *A Dimensionality-Reduction Algorithm for Accurate Visualization of Personal Relationships Within a Team*

VŠ VÝUKA

LS 2014 *Philosophy and psychology of consciousness*, FF UK

2010-2011 vedení semináře ke kurzu AFS100591 *Wittgenstein* (přednášející V. Kolman, PhD.)

2010-2011 *Úvod do filosofie*, Katedra marketingové a mediální komunikace, FSV UK

+ dílčí přednášky na FF UHK (LMS), FF UPOL (Jednota filozofická), FF UK (Katedra psychologie; Společnost pro kognitivní vědu)

PUBLIKACE

obor filosofie

- Vraný, M. (2013). Empirická výzva filosofickému pojetí mysli. *Organon F. medzinárodný časopis pre analytickú filozofiu*, 1(20), 50-71.
- Vraný, M. (2012). Language and the Unity of Consciousness, In P. Frath, V. Bourdier, E. Hilgert, K. Bréhaut, J. Dunphy-Blomfield (Eds.), *Res per nomen III : La référence, la conscience et le sujet énonciateur / Reference, Consciousness and the Speaking Subject*, Editions et Presses Universitaires de Reims

empirické studie

- Džúrová, D., Spilková, J., Vraný, M. (2016). Substance misuse and its risk perception in European teenagers. *Children's Geographies*, 14(2), 203-216.
- Vraný M., Bahbouh, R. (2014). Relating team size and communication patterns in work teams. In R. Bahbouh, G. Lasker (Eds.). *Sociodiagnostics and sociomapping: Volume III*. (pp. 29-33). Canada: The International Institute for Advanced Studies in Systems Research and Cybernetics.

- Vraný M., Bahbouh, R. (2014). The relation of communication and sympathy in a work team. In R. Bahbouh, G. Lasker (Eds.). *Sociodiagnostics and sociomapping: Volume III.* (pp 41-45). Canada: The International Institute for Advanced Studies in Systems Research and Cybernetics.
- Sněhotová, J., Vraný, M., Bahbouh, R. (2014). Evaluation of quantitative and qualitative aspects of family communications and interpersonal relations. In R. Bahbouh, G. Lasker (Eds.). *Sociodiagnostics and sociomapping: Volume III.* (pp 1-8). Canada: The International Institute for Advanced Studies in Systems Research and Cybernetics.

PŘEKLADY

Překlady vybraných článků z AJ do ČJ ve sbornících *Syntetické apriori* (H. Janoušek, V. Kolmans eds.) a *Vědomí bez redukce* (J. Mihálik, J. Hill eds.).