

Martin Vraný

## Abstract

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 argue in the Introduction 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.

In chapter 2 I analyze various aspects in which consciousness is thought to be unified and conclude that two such aspects are particularly challenging for naturalizing the unity and that they cannot be treated separately. They are 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.

Chapter 3 describes main conceptual and methodological issues faced by naturalistic accounts of the unity. I distinguish between a normative and an objective aspect of the unity and show that explaining the latter is a domain of cognitive neuroscience. I also defend the concept of neural representation without which naturalization of the unity is inconceivable.

In chapter 4 I turn to Kant's transcendental psychology in order to analyze the unity further. Kant provided a very detailed and insightful account of the unity under the term transcendental unity of apperception and drew an important distinction between empirical and transcendental self-consciousness. Kant's arguments put the unity of consciousness into the broader context of cognition, situatedness and agency.

In chapter 5 I present philosophical accounts that help to specify the unity as an explanandum for natural sciences. First, I review Shoemaker's and Castaneda's work on the logic of self-reference. Second, I present Hurley's two-level interdependence model of consciousness and her concept of perspectival self-consciousness. Her account not only preserves the conceptual richness of the unity as presented earlier but is congenial to many theories in cognitive science. Finally, I present an account of empirical self-consciousness that puts together the threads explored in previous chapters. The main point is to argue that the self-reflective state should be understood as a representational transformation of the object state, not as a higher-order state that contains the object state as its proper part.

In chapter 6 I finally review selected neuroscientific theories of consciousness: the neural global workspace theory, the predictive coding theory, and the information integration theory. I specify what the unity amounts to in these theories and assess how good their explanation is. Overall, I conclude that the predictive coding theory offers the best, albeit incomplete, explanation.

**Keywords:** unity of consciousness, neuroscience, naturalization, philosophy of mind