PHENOMENOLOGY AS ANOTHER TOOLBOX FOR NEUROSCIENTISTS?

Lars Schwabe and Olaf Blanke

"[I]t has become next to impossible for a single mind fully to command more than a small specialized portion of it. I can see no other escape from this dilemma [...] than that some of us should venture to embark on a synthesis of facts and theories, albeit with second-hand and incomplete knowledge of some of them – and at the risk of making fools of ourselves."

Erwin Schrödinger in "What is life?" (1944)

1. Introduction

In the preface to his book "What is life?", Erwin Schrödinger calls for scientific and scholarly "trespassing" despite exposing oneself to criticism with respect to possibly incomplete approaches to the question at hand, in his case of how physics and chemistry may account for the complexity of life (Schrödinger, 1944). His book has become an inspiration for many researchers from a variety of academic backgrounds, including biologist Francis Crick. Understanding self-consciousness and how it relates to the brain is certainly a project of similar complexity and in need of trespassing, due to the multidisciplinarity in cognitive science.

The book "The phenomenological mind" by philosophers Shaun Gallagher and Dan Zahavi (2008) is an introduction into the phenomenological philosophy of mind, which is an important and timely topic and believed to have the potential of making significant contributions to the interdisciplinary study of the conscious mind and consciousness. Phenomenology, according to the layman's understanding, refers to how perception and cognition "feel from the inside" with introspection being the primary method. This focus on subjectivity and the first-person perspective seems at odds with the third-person perspective adopted by the natural sciences. As a consequence, many researchers may not consider such phenomenological approaches and favor apparently well-defined approaches such as quantifying behavior or brain activations during perceptual and cognitive tasks.

Philosophical phenomenology, however, refers to a philosophical tradition originating in the works of Husserl, Heidegger, Merleau-Ponty, Sartre, and others. Gallagher and Zahavi argue that within this tradition, methodological and conceptual tools have been developed and successfully applied. Hence, there is a gap between the layman's understanding of phenomenology and the rich tradition of philosophical phenomenology. With their book, Gallagher and Zahavi aim at closing this gap by informing the reader about the developed methodological toolboxes of philosophical phenomenology. The authors also discuss recent and possible future applications to current topics in cognitive neuroscience such as, for example, in chapters on self, consciousness, embodiment, and motor awareness.

Here we ask, in which ways Gallagher and Zahavi's presentation of the phenomenological approach is of value to current cognitive neuroscience. In particular, we adopt the neuroscientific perspectives of cognitive neuroscience and computational neuroscience and ask: Could the phenomenological approach be of practical or epistemological value for the work done in any of these fields? We focus our discussion on the topics of the embodied mind, the neuroscientific investigation of the self, and the proposed methodologies.

2. Cognitive neurology and neuroscience

2.1 Körper and Leib

How the body "appears in experience" and how the body "structures our experience" is a prominent topic in philosophical phenomenology, and is starting to receive some interest in the cognitive neurosciences. Gallagher and Zahavi indicate that perceptual and cognitive processes are strongly influenced by bodily constraints such as posture and action capabilities as well as its relation to gravity: "Phenomenology [...] seeks to understand to what extent our experience of the world, [...] self [...] and other objects and people are formed by and influenced by our embodiment (p.136)." Gallagher and Zahavi introduce Edmund Husserl's notion of two different kinds of body representations, describing his "Körper" as "objective" body and his "Leib" as "subjective" body. These are defined as "two different ways that we can understand and experience the body" stating that "Leib" captures "the body understood as an embodied

first-person perspective" (p.137) characterized essentially by automatic and prereflective processes, whereas "Körper" focuses on the body as perceived "from an observer's point of view" characterized by cognitive and reflective processes. Maurice Merleau-Ponty is discussed as having developed these notions further by his detailed analyses of the special role of the "Leib" (or *corps vécu*) during perception and action. The chapter on embodiment is a very readable introduction into several unresolved issues for students, researchers, and philosophers alike interested in body- and selfrepresentations and its relevance for consciousness. The chapter leads to important questions such as: What is it like to have an embodied first-person perspective? How is such an embodied first-person perspective achieved that Gallagher and Zahavi describe as non-centered, non-perspectival, and as "a view from nowhere"? What might its mechanisms be? The authors seem to favour proprioceptive brain mechanisms.

2.2 A proprioceptive view from nowhere?

What is the basis of the embodied first-person perspective characterizing the Leib? Gallagher and Zahavi speculate about a spatial reference frame of the body that is tightly linked to sensori-motor body representations, especially proprioception. They argue that this reference frame is non-ego-centered and non-perspectival (to have no origin or centered perspective) and that it is grounded in position sense or proprioception. They write that aspects related to "Körper" have been studied more commonly and are characterized by ego-centric body representations in an egocentric reference frame. These latter views from somewhere are characterized by experience from a perspectival origin and are assumed to arise at a later stage with a perceiver as the experiential zero-point.

Does proprioception encode a non-perspectival reference frame? José Bermúdez (1998) seems to agree with Gallagher and Zahavi arguing for a fundamental difference between spatial reference frames based on proprioception as compared to those based on exteroceptive perception such as vision and audition (p.143). But what do the authors exactly refer to when mentioning proprioception or position sense? What are the sense organs, the preferred cues, and the neural pathways involved? Does proprioception really lack perspectivalness? We would argue that non-perspectivalness in position sense may apply for upper limb proprioception, but non-perspectivalness is probably

less strongly present for lower limb proprioception due to its role in directing and orienting the body via the legs' role in body support. This is even more the case for neck proprioception that orients and directs head and eyes and should in our opinion not be considered aperspectival. Accordingly, proprioception should not be considered as a unitary system, but as a single sensory system with multiple body part-specific subsystems. We would argue that lower limb and especially neck proprioception are computing a perspectival proprioceptive body representation, whereas arm proprioception has a different function. This would lead us to postulate that embodiment of the first-person perspective as based on proprioceptive input is likely to differ for these different body parts (head-trunk, arms, legs). We speculate that especially these perspectival proprioceptive cues may turn out to be crucial mechanisms for the embodied first-person perspective of the Leib.

2.3 Multisensory and sensorimotor origins of the embodied perspective

Are proprioceptive frames of reference (as proposed by Gallagher and Zahavi or as proposed by us in the preceding paragraph) the only origin of what Merleau-Ponty describes as "phenomenally experienced spatiality" (p.143) or the embodied first-person perspective? We do not think so. We think that several other non-prorioceptive sensory systems also contribute crucially to the embodied first-person perspective. Next to contributions from the motor system that shares several aspects with proprioception, neurophysiological research has revealed that it is also important to distinguish the contributions from tactile cues from the plantar sole (Roll et al., 2002) or from vestibular translational and rotational cues (Day and Fitzpatrick, 2005). Signals within both systems are processed automatically and pre-reflectively, are continuously present and mostly in the background of human experience, just as proprioception. "I am [generally] not conscious of my body [defined by proprioceptive, foot sole tactile, and vestibular cues] as an intentional object. [...] I am it" (p.143). Tactile and vestibular cues in addition to proprioceptive and motor cues are likely to contribute fundamentally

¹ These differential roles of proprioception on experience and behavior can also be demonstrated experimentally. Muscle vibratory stimulation of spindle afferents at the neck, but not at the upper limbs, may lead to illusory own head and trunk movements, tilts of the visual world, or shifts of spatial reference frames (Lackner and Levine, 1979).

to the embodied first-person perspective and need to be integrated in a more finegrained manner into philosophical and neurobiological models.

2.4 One or several experiential zero-points?

Gallagher and Zahavi describe the embodied mind as possessing one experiential origin or one zero-point constituting a single spatial reference "point in relation to which every object is oriented". Is human experience always characterized by a single first-person perspective? Recent data on so-called autoscopic experiences suggest that human experience may also be characterized by the absence of a single zero-point or embodied perspective, but by at least two simultaneous or rapidly alternating embodied perspectives. This suggest that the ego-centric reference frame, or the view from somewhere, the perspectival origin of human experience may not be as unitary as normally experienced. Recent neurological data suggest that this might be due to the different multisensory mechanisms involved in body representation. Thus, neurological patients with heautoscopy may claim to experience to perceive from two spatially distinct first-person perspectives ((Blanke et al., 2004); patient 2). Sometimes these patients report to "be split in two parts or selfs" or to feel as if "I were two persons" (Pearson and Dewhurst, 1954) or as a "split personality" (Lunn, 1970; for further discussion see (Blanke and Mohr, 2005)). Other patients may describe a auditory firstperson perspective that is spatially distinct from a simultaneous first-person visual perspective. As Gallagher and Zahavi endorse the heuristical importance of clinical case studies, what do these observations tell us about the mechanisms of (the) embodied first-person perspective(s)? How can these experiences be accounted for and integrated into phenomenological philosophy? Similarly, the experiential origin or the indexical "here" is not only characterized by the experience of a perspective that is directed towards the world, but also by an experienced location of the self. Self-location refers to experiencing the self to be localized in one's body and at a certain position in extrapersonal space. Interestingly, the experienced location of the perspective that is directed towards the world can be dissociated spatially from experienced self-location. In a recent neurological study, for example, the authors reported that - after electrical brain stimulation - the patient's self-location was systematically experienced at a location that was spatially distinct from his visual first-person perspective (De Ridder et al., 2007).

Collectively, these data suggest that human experience (in these clinical cases at least) may be characterized by multiple simultaneous first-person perspectives and selflocations that are grounded in multisensory and sensorimotor brain mechanisms. These findings have recently been employed in experiments in healthy subjects using virtual reality (Ehrsson, 2007; Lenggenhager et al., 2007) suggesting that a similarly complex experience can be uncovered experimentally. The dialog between neuroscience and phenomenological approaches on "how the body shapes the mind" should certainly be a two-way route. More philosophically informed neuroscientific work is needed to describe and account for the mechanisms leading to the embodied first-person perspective and especially the question how so-called aperspectival mechanisms lead to our global and centered perspective of the subject. We have proposed here that proprioceptive, vestibular, tactile, and motor perspectival cues related to head and trunk representation are crucial. Phenomenological analyses of multiple simultaneous firstperson perspectives may be one interesting avenue to pursue as they reflect limits of body and self representation. For example, training and performing phenomenologically informed interviews and studies in neurological patients with heautoscopy could be rewarding. However, trained phenomenologists could also use virtual reality techniques in order to evoke similar experiences, which could then be analyzed using phenomenological methods.

3. Computational neuroscience

3.1 Current practice

We believe that the field of computational neuroscience is well suited to mediate between philosophical and empirical approaches. Therefore, we ask in which ways ideas from the phenomenological tradition and ideas proposed by Ghallager and Zahavi could be integrated and put to work. Unfortuantely, the field of computational neuroscience is still a rather young discipline with almost as many different conceptual and methodological approaches as there are computational neuroscience labs. It mainly lacks a broadly accepted basis comparable to, for example, Newton's laws of motion or

the Navier-Stokes equations, upon which subsequent work can build. In part, the lack of such a basis is due to the complexity of the systems investigated. At least for the near future, however, one may have to live with a multitude of different approaches, and one may have to select the proper level of description to match the problem at hand without a rigorous derivation from underlying constituent dynamics. This is a methodological issue, which may or may not be overcome in the future.

A more conceptual issue is to link neuronal activations and human experience. How do computational neuroscientists approach human experience? Do they have, according to phenomelological philosophy, a proper notion of human experience, or are they just interested in neural mechanisms? Here, it is instructive to have a look at the current practice in the field. Based on a taxonomy suggested in a widely used textbook (Abbott and Dayan, 2000), one can distinguish three kinds of models used in the field: descriptive models, mechanistic models, and computational models. Descriptive models are black-box models, which account for the input-output transformation performed by a particular neuronal system. For example, the time-averaged rate of action potentials emitted by retinal ganglion cells as a function of the spatial light pattern used for their stimulation can be well described by the weighted difference of two Gaussian functions. Such a descriptive model, however, always involves an assumption regarding the format of the neuronal code, and it abstracts form the underlying mechanisms. Mechanistic models are intended to account for the underlying mechanisms. Finally, computational models are supposed to make explicit the functional role of particular neuronal systems, often by applying concepts developed in other fields like pattern recognition, information theory, control, decision or game theory.

The field of computational neuroscience should be expected to contribute explicit formal links between different levels of description. It certainly can link neuronal activations to motor responses, and this is not only a valuable, but also an important task often underestimated. It is, however, obvious to almost every researcher in the field, that a simple identification of an experience from a first-person perspective with activations in, say, sensory areas of the brain is not a satisfying explanation of how human experience is linked to neuronal activations. These shortcomings of such an identification of experience and the first-person perspective with neuronal activations remain even if we think of them as being spatio-temporal patterns of electrical activity

distributed over brain-wide networks having or not having the property of being oscillatory, containing synchronous firing patterns of action potentials, involving subthreshold neuronal activity, etc. To the best of our knowledge, beyond such rather non-satisfying identifications, no other ideas have been postulated so far. In which ways can the phenomenological approach contribute to clarify the link between first-person experience and neuronal activation? Would a phenomenologist favor one out of the three kinds of modeling approaches? Does a phenomenological approach suggest a particular way of thinking about their mutual relation?

3.2 Recent findings

The investigation of conscious experience and embodiment are at most very minor topics in the field, but some topics could be of relevance. Here, we consider Bayesian processing of sensory information and models of sensory-motor processing.

Bayesian processing

Bayesian processing is a computational paradigm often used as an analogy to the processing of sensory information in the brain. It is rooted in logic (Cox, 1961) and statistics (Jaynes, 2003), and it is appealing because it formalizes two important aspects of biological information processing. First, information is always considered as being inherently uncertain. Second, the Bayesian approach shows how to optimally combine new sensory information with previously acquired information, the so-called prior beliefs, in order to arrive at the so-called posterior beliefs. Hence, this approach is well suited to account for the information processing in sensory areas as well as sensorymotor processing. Since it is a computational approach, it does not make strong predictions about measureable neuronal activations.

How can such a framework be applied to the embodied first-person perspective? Recently, we proposed that the vestibular component of so-called out-of-body experiences, which involves the illusion of flying and an elevated first-person perspective despite the fact that the physical body is stationary, is compatible with a Bayesian approach to vestibular information processing (Schwabe and Blanke, 2008). In particular, we proposed that a Bayesian integration of the sensory vestibular signals received in the supine position and a false prior belief leads to the illusory experience of the elevated first-person perspective and self-location that are reported during out-of-

body experiences in the sense that the posterior belief corresponds to these illusory vestibular sensations of elevation. In other words, this particular work and probably many previous studies have (implicitly) identified the posterior belief within the Bayesian framework with first-person experience. Such identification is not a naïve identification of first-person experiences with neuronal activations, because the Bayesian posterior belief is a mathematical object. Here, phenomenological reflections on this identification would be of great value.

Sensory-motor processing and the sense of agency

Gallagher and Zahavi propose that sensory-motor processing is of relevance for understanding conscious experience and embodiment, because it deals explicitly with controlling movements based on sensory and motor information. Using paradigms developed within control theory, computational models of sensory-motor processing have been developed and tested experimentally. One of the key findings is that in some carefully designed behavioral experiments human subjects behave as if they make use of so-called forward models (Wolpert et al., 1995), i. e. statistical models predicting the sensory inputs caused by the motor outputs and the subsequent limb and body movements. For example, according to these models, once the command to move an arm is sent out, the forward models are already predicting the future sensory inputs to be received if the arm actually moved. The corresponding neuronal processes are believed to occur almost automatically in the sense of not demanding cognitive efforts (i.e. "pre-reflectively") and are probably relevant for the embodied first-person perspective.

When discussing agency, Gallagher and Zahavi point out further aspects and discuss them in the context of brain-imaging experiments. Another important mechanism may be the identification of a vanishing prediction error with the sense of agency. Gallagher and Zahavi ask (p. 163): "should we think of the pre-reflective sense of agency as belonging to the realm of motor control and body movements, or as belonging to the realm of intentional action"? While the reviewed studies in chapter 8 (according to Gallagher and Zahavi) associate the sense of agency either with bodily movements (Tsakiris-Haggard), their goals (Farrer-Frith) or higher-level reflective thoughts (Graham-Stephens), the authors argue that the sense of agency depends on the integration of all three aspects.

The authors' account of the sense of agency may or may not be superior to the other three more focused proposals. However, the discussion seemed too short to us and was difficult to follow. We believe, however, that all reviewed studies and agency accounts lack more precise formulations in terms of mathematical models. Given that the – due to its tight link to bodily movement – rather low-level Tsakiris-Haggard explanation has also been proposed only informally, a first step would ideally be the mathematization of these proposals (see also (Schwabe and Blanke, 2007)). We believe that a more complete explanation of the sense of agency should be given in terms of more quantitative models of sensory-motor processing. They may, however, involve at least multiple time-scales in order to account for long-term goals (Kilner et al., 2007) and recurrent loops to account for predictions and prediction errors. In particular, such models could incorporate intentional feedback in order to sidestep Ghallager and Zahavi's objection that "the sense of agency is not reducible to awareness of bodily movement or sensory feedback of bodily movement" (p. 165).

3.3 Added value of the phenomenological approach

Neuro- and front-loading phenomenology

In this section we ask as to whether and how other concepts from philosophical phenomenology as introduced by Gallagher and Zahavi have a practical or epistemological value for computational neuroscience and could be integrated. In particular, we consider some aspects of the methodologies reviewed in their Chapter 2 as well as the different notions of embodiment considered in Chapter 7. Throughout their book, the authors emphasize that philosophical phenomenology starts with experience. Experiences are analyzed from the first-person perspective using the phenomenological method. The authors' review of Husserl's method of epoché and the phenomenological reduction is brief, but it very well serves the purpose of convincing experimentalists and theoreticians, who are not experts in the exegesis of phenomenological texts or not even aware of the original writings of Husserl and other phenomenologists, that an analysis of experience involves much more than just introspection.

The authors move on to suggesting several ways of how phenomenology could contribute to the neurosciences. In particular, they suggest that first-person reports about

their experience during perceptual tasks shall be correlated on a trial-by-trial basis to the neuronal activations measured with neuroimaging methods ("neurophenomenology"). However, even if the subjects are very well trained in the phenomenological method, the authors did not indicate in enough detail, in which manner such an approach is distinct from current approaches investigating the neurobiology of self-consciousness (Laureys, 2005). Under the term "front-loading phenomenology" they also consider experimental tests of hypotheses obtained using phenomenological analysis. For example, phenomenological analysis reveals that the sense of ownership and the sense of agency can be dissociated as the same body movement can be experienced as being caused externally (for example when being moved passively) or caused by oneself. Again, however, it is not clear in which manner such an approach is truly distinct from scientific methodology seeking to identify the neuronal correlates of experience, self, or first-person perspective. In both cases, experimenters correlate subjects' reports with measured brain activity. Would the authors argue that the phenomenologically trained subject activates distinct or different brain regions when performing experimental paradigms involving the sense of ownership and agency as compared to naïve subjects? Does she report items not available to the untrained? How does front-loading phenomenological method differ from current approaches on agency and ownership currently employed in the cognitive neurosciences? For example, how would Gallagher and Zahavi classify the rubber hand illusion experiments (Botvinick and Cohen, 1998)? In our opinion, both introduced approaches of the naturalization of phenomenology do not yet demonstrate unique features of the phenomenological approach, which would make it attractive for experimentalists to consider them.

Husserl, Helmholtz, and the first person perspective

What about Ghallagher and Zahavi's proposal about neurophenomenology and the first-person perspective of the "Leib"? In Chapter 2, they suggest that a mathematization of phenomenological first-person descriptions together with a mathematization of experimental third-person descriptions could lead to a formal theory of how first- and third-person descriptions are related. How can such an approach be linked to computational neuroscience?

We have emphasized that computational neuroscience is still a rather young field. However, the taxonomy of descriptive, mechanistic and computational models can be used in order to organize the variety of different approaches. How does the suggested integration of experimental science and phenomenology via mathematization relate to this taxonomy? Gallagher and Zahavi identify dynamical systems theory as the main mathematical approach currently applied by theoreticians. They correctly criticize it as being too narrow in the sense of neglecting the subjective dimension of perception and point to a need for a different kind of mathematization to account for the first-person perspective. It would be important to know the authors' position regarding the way these mathematizations should be linked to each other. For example, are they thinking of isomorphisms, in the sense of an "implementation" of perceptual mechanisms and laws, which could be revealed by phenomenological analyses, to be executed on a neuronal wetware described from a third-person perspective like the three levels of description – problem-, algorithm- and implementation-oriented – introduced by David Marr (1982) in the context of vision?

The Bayesian approach makes heavy use of probabilistic descriptions, and interpreted in a certain way, it can be viewed as a modern form of an automatic, unconscious, and pre-reflective perceptual inference about the state of the world as proposed by Hermann von Helmholtz. What would Edmund Husserl think of such a probabilistic mathematical description? Which notion of possibility would be adequate for Bayesian perceptual inference? An important question remains: how much insight into experience and the embodied first-person perspective can be gained in the cognitive sciences by relying on phenomenological analysis of *unconscious* processes, given that phenomenological analysis starts with *conscious* experience? What are the limits of phenomenological analysis of *unconscious* and normally *attenuated* and *transparent* processes?

Lars Schwabe and Olaf Blanke

Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland olaf.blanke@epfl.ch

References

- Abbott, L.F. and Dayan, P. (2000), *Theoretical Neuroscience*, Cambridge MA: MIT Press.
- Bermudez, J.L. (1998), The paradox of self-consciousness, Cambridge MA: MIT Press.
- Blanke, O. and Mohr, C. (2005), 'Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin Implications for neurocognitive mechanisms of corporeal awareness and self-consciousness', *Brain Research Reviews* 50, 184-199.
- Blanke, O., Landis, T., Spinelli, L. and Seeck, M. (2004), 'Out-of-body experience and autoscopy of neurological origin', *Brain* 127, 243-258.
- Botvinick, M. and Cohen, J. (1998), 'Rubber hands 'feel' touch that eyes see', *Nature* 391, 756.
- Cox, R.T. (1961), *The Algebra of Probable Inference*; Baltimore: The Johns Hopkins Press.
- Day, B.L. and Fitzpatrick, R.C. (2005), 'Virtual head rotation reveals a process of route reconstruction from human vestibular signals', *Journal of Physiology* 567, 591-597.
- De Ridder, D., Van Laere, K., Dupont, P., Menovsky, T. and Van de Heyning, P. (2007), 'Visualizing out-of-body experience in the brain', *New England Journal of Medicine*, 1829-1833.
- Ehrsson, H.H. (2007), 'The experimental induction of out-of-body experiences', *Science* 317, 1048.
- Gallagher, S. and Zahavi, D. (2008) *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*, London: Routledge.

- Jaynes, E.T. (2003), *Probability Theory: The Logic of Science*, Cambridge MA: Cambridge University Press.
- Kilner, J.M., Friston, K.J. and Frith, C.D. (2007), 'Predictive coding: an account of the mirror neuron system', *Cognitive Processing* 8, 159-166.
- Lackner, J.R. and Levine, M.S. (1979), 'Changes in apparent body orientation and sensory localization induced by vibration of postural muscles: vibratory myesthetic illusions', *Aviation*, *space*, *and environmental medicine*, 346-354.
- Laureys, S. (2005), *The boundaries of consciousness. Neurobiology and neuropathology*, Amsterdam: Elsevier.
- Lenggenhager, B., Tadi, T., Metzinger, T. and Blanke, O. (2007), 'Video ergo sum: manipulating bodily self-consciousness', *Science* 317, 1096-1099.
- Marr, D. (1982), Vision. New York: Freeman and Company.
- Pearson, J. and Dewhurst, K. (1954), 'Two cases of heautoscopic phenomena following organic lesions', *Encephale* 43, 166-172.
- Roll, R., Kavounoudias, A. and Roll, J.P. (2002), 'Cutaneous afferents from human plantar sole contribute to body posture awareness', *Neuroreport* 13, 1957-1361.
- Schrödinger, E. (1944), What is Life? The Physical Aspect of the Living Cell, Cambridge MA: Cambridge University Press.
- Schwabe, L. and Blanke, O. (2007), 'Cognitive neuroscience of ownership and agency', Consciousness & Cognition 16, 661-666.

- Schwabe, L. and Blanke, O. (2008) 'Out-of-body experiences: false climbs in a supine position?', In: *International Joint Conference on Neural Networks IJCNN'2008*, Hong Kong.
- Wolpert, D.M., Ghahramani, Z. and Jordan, M.I. (1995), 'An internal model for sensorimotor integration', *Science* 269, 1880-1882.