

Dieter Wandschneider

Farewell to Chalmers' Zombie.
The 'Principle Self-Preservation' as the Basis of 'Sense'¹

Abstract

My argument is that Chalmers' Zombie fiction and his rigid-designator-argument going back on Kripke comes down to a *petitio principii*. Rather, at the core it appears to be more related to the essential 'privacy' of the phenomenal internal perspective. In return for Chalmers I argue that the 'principle self-preservation' of living organisms necessarily implies *subjectivity* and the emergence of *sense*. The comparison with a robot proves instructive. The mode of 'mere physical' being is transcended if, in the form of *phenomenal* perception, sense appears on the stage of higher animals – a transition explained here as an *emergence phenomenon* based on the systemic co-operation of perception, evaluation and action ('perc-val-act system'). Some fundamental considerations are added: Those consequences implied by the *principle self-preservation* reveal the *natural-biological* origin of the organism – primarily seeming a more insignificant circumstance – as a momentous *fundamental difference* (end-in-itself-character, subjectivity, constitution of sense) compared to technical artefacts (robot). And the *emergentist* approach indicates the – maybe paradoxical – possibility of a *dualism* of physical and psychical phenomena in an overall physical system, that is *not dualistic* at the same time.

1. Introduction

Behavioral decisions, such as avoiding a disturbing obstacle, are everyday mental processes that realize *senseful relationships*. Brain research also ultimately aims at senseful relationships, trying to develop them from coincidences of behavior and brain activity (or verbal reports), for example, if the intake of food goes hand in hand with the activation of the reward center. But what it finds empirically are always *physical* facts and never senseful mental states. Joseph Levine (1993) has pointed to this discrepancy between physically demonstrable brain activities and the correlating senseful mental as an *explanatory gap* of brain research. David Chalmers calls it "the hard problem", in contrast to "easy problems" of neuronal-functional relationships such as learning, memory, figure segregation etc. (Chalmers 2007, 59). In my opinion that will not be the last word, because, as I have explained elsewhere, this criticism is based on a deficient concept of the 'physical' (Wandschneider 2015 and 2016a).

The diverse organizational forms of *life* testify to the possibilities that are actually realized in the physical. The living organism is not only the marvel of a system that functions perfectly in a technical sense, but – which will be explained in more detail below – due to the *principle self-preservation* the basis of *sense* and thus of something that is no more only a primitive 'bare physical'. This senseful component that goes beyond the merely physical becomes more clearly recognizable in the contrasting juxtaposition with the *robot* (in today's sense), which – despite all technical perfection – is characterized by manifest *sense dementia*. The *principle self-preservation as the basis of 'selfness' and 'sense'*, and thus as a transgression of the primitive bare physical, is the central theme of this study.

Before that, however, I have to deal with a mighty objection that might well concern an undertaking like this: I think of the doctrine developed by *David Chalmers* in *The Conscious Mind*, according to which the request, phenomenal mentality (and thus something like 'sense') to understand from physical must fail for (*modal*) *logical reasons*. I would like to show that this view is not tenable but boils down to a *petitio principii* and that Chalmers' basic argument for an existence of

¹ English translation of: Dieter Wandschneider - *Abschied von Chalmers' Zombie. Das 'Prinzip Selbsterhaltung' als Basis von 'Sinn'*, in: Zeitschrift für philosophische Forschung 72 (2018), 246–262.

the mental independent of all physical, which goes back to Saul Kripke, rather concerns the undeniable *privacy* of the – nonetheless physically based mental inner perspective.

2. Chalmers' Zombie-Doctrine

Chalmers argues with the *logical possibility* of a 'zombie world': "We can consider the logical possibility of a *zombie world*: a world physically identical to ours, but in which there are no conscious experiences at all. In such a world, everybody is a zombie. [...] So let us consider my zombie twin. This creature is molecule for molecule identical to me, and identical in all low-level properties postulated by a completed physics, but he lacks conscious experience entirely" (Chalmers 1996b, 94, also 71 and others). In this thought experiment, the zombie world is assumed to be an exact physical replica of our world, in which physically and biologically everything can be explained as in our world, including such abilities as are characteristic of higher animals and humans: learning, memory, perception, behavioral control, attention, categorization, language skills, etc. (16 ff). All this, according to Chalmers, can be derived from the physical facts assumed to be identical in both worlds – with one single exception: "There will be no phenomenal feel" (95), i.e. phenomenal states would not exist in the zombie world.

Why not? In contrast to "almost everything" else, Chalmers argues (1996b, 45), the phenomenal-mental is *not logically* supervenient on the physical, but possesses "*mere* natural supervenience". It is therefore indeed nomological derivable, i.e. on the basis of contingent empirical natural laws, but not in a purely logical way (37 f). For this reason, so Chalmers, it is to be *conceivable*, and thus also *logically possible*, that my zombie twin, who physically and even (in a neuronal-functional sense) psychologically exactly resembles me in the zombie world, does not possess phenomenal consciousness, unlike me. Although it *arises* from the physical, it does *not follow logically* from it (125).

What is unclear here is first of all why 'almost everything', explicitly also biological properties (e.g. 35, 74), should *logically* supervene on "physical properties", finally therefore on the laws of nature, but *this kind of logical* supervenience is denied to the phenomenal, since it is *only nomological*, i.e. – likewise! – supervenient on natural laws. In fact, there is a fundamental ambiguity in Chalmers on this point: "I am not suggesting that high-level facts and laws are entailed by microphysical *laws* [...] I am making the much weaker claim that high-level facts are entailed by all the microphysical *facts* (perhaps along with microphysical laws)" (71), whereby the microphysical 'facts', despite this less expressive characterization – should be completely different in other possible worlds and thus have *contingent* status. What Chalmers thus is presumptively about is simply the strict differentiation of physical structures on the one hand, which (in principle) can throughout physically be explained (in this respect 'logical' supervenience) and phenomenal-mental forms on the other hand for which this, following Chalmers, should not apply (insofar 'only natural' supervenience) – but both on the merely contingent basis of 'microphysical facts'.

This argument of the *contingency* of the physical is what underlies the zombie thought experiment. The zombie world is presupposed as exactly the *same world as ours* – i.e. including all the laws of nature – however without the forms of phenomenal experience. That such a world, so Chalmers, is *contradiction-free conceivable*, is intended to illustrate the *logical possibility* that the phenomenal is essentially different from the physical, since it is logically independent of the latter. Chalmers puts forward five arguments in favour of this:

(1) A zombie world is conceivable without contradiction. Of course, this is more of an assurance than an argument. For this, Chalmers can only put his own subjective certainty into the field: "I can discern no contradiction in the description" (Chalmers 1996b, 96, 99). And the hint that the otherness of my zombie twin could be based on "*nonstandard realizations of ... functional organization*" (97) contradicts the basic assumption that it should be physically exactly the same to me.

(2) The possibility of a so-called *inverted spectrum* is also conceivable (Chalmers 1996b, 99 ff), i.e. where I see 'red' my zombie twin sees 'blue', although we both use the same name for it (about 'red'). What Chalmers asserts here is basically the inaccessibility of the 1st-person perspective from the outside, I will come back to this later.

(3) This also applies to the fact of *epistemic asymmetry* (Chalmers 1996b, 101 ff) that consciousness is only given for myself, i.e. in the first person perspective – while learning, memory, perception etc. very well are accessible and researchable from the outside.

(4) The thought experiments of Frank C. Jackson ('Mary') and Thomas Nagel ('What is it like to be a bat?') mentioned by Chalmers also go in the direction of the 1st-person perspective (Chalmers 1996b, 103 f): Mary, who grew up in a pure black and white environment, has never experienced any colour qualities, although, as a top expert, she *knows* everything about visual perception. But *knowledge* can neither convey nor replace her color *experience*. And analogously, the researcher who knows everything about bats cannot know what it *feels* like to be a bat.

(5) Chalmers finally asserts (Chalmers 1996b, 104 ff) that the analysis of neuronal functions (= 3rd person perspective) can explain memory, learning etc., but not phenomenal experience (= 1st-person-Perspective). One reason for this is to be the vagueness of what should actually be determined (e.g. "does a mouse have beliefs?" (105)).

In my view, these arguments do not do what they are conceived to do: (1) and (5) have only an assuring character with regard to what actually was to be proven (the impossibility to derive the phenomenal from the physical). (2), (3) and (4) are based on the inaccessibility of the 1st-person perspective. This seems to be at all the central issue that is ultimately always at stake in this book. Chalmers himself admits that the *intuition* on which his view is based has its origin in this (Chalmers 1996b, 110).

As I would like to show, this is in fact also the key to the following argumentation, which goes back to considerations by Saul Kripke and is intended to theoretically secure Chalmers' zombie fiction – extending beyond its mere *conceivability*. A decisive role is played by the concept of the *rigid designator*, which denotes the same thing in all possible worlds (Kripke 59). Kripke also develops the idea of a *two-dimensional semantics* in which a distinction must be made between a *primary intension* and a *secondary intension* – I take over the terms used by Chalmers here and thus formulate the decisive argument at once (Chalmers 1996b, 146 f):

For example, the primary intension of water is 'wateriness' (as I call it that for simplicity's sake), the secondary intension of water is 'H₂O', namely in all possible worlds. If in one of the possible worlds there would be a form of wateriness that is not H₂O but XYZ, then the term 'water' would be misused here; it might be a similarly watery seeming liquid like gasoline.

A strictly analogous argument for the phenomenal state of *pain* and the neural *pain nerve stimulation* would take the following form: Primary intension of pain is '*pain sensation*', secondary intension of pain is '*pain nerve stimulation*', namely in all possible worlds. If in one of the possible worlds there would be a form of pain sensation that is not pain nerve stimulation but UVW, then the term 'pain sensation' would be misused here; it might be perhaps a tickle.

But according to Kripke or Chalmers, in this case it cannot be argued in this way, because 'feeling of pain' as a *rigid designator* means the same thing in all possible worlds, namely a painful condition. So if in one of the possible worlds a form of pain sensation exists which is not pain nerve irritation but UVW, then the term *pain sensation* is not misused here, because 'pain sensation' is a rigid designator. A feeling of pain is a feeling of pain in all possible worlds, *regardless of* whether a pain nerve stimulation or UVW is realized. For this reason, the feeling of pain *cannot be identical* with pain nerve stimulation (in Kripke's conception) or *logically supervene* on it (in Chalmers' conception).

Now, 'wateriness' too is a rigid designator, i.e. it describes something watery in all possible worlds. This raises the question why there is a misuse of the word when water is called 'water' in the XYZ world (since it is actually gasoline) – while pain, regardless of how the associated pain sensation is instantiated in a possible world (as pain nerve irritation or as an UVW process), is still aptly called 'pain'.

The answer to that, I think, is now clear: Water, interpreted as watery stuff, is *constitutively* linked to H₂O, so that gasoline is wrong watery stuff. Pain, on the other hand, interpreted as a feeling of pain, is to stand for itself, i.e. the connection with pain nerve irritation is *from the outset as not constitutive defined*, so that in other possible worlds completely different instantiations can occur without affecting the character of the feeling of pain. Other instantiations are thus classified as merely contingent side effects, in other words: Not the fact that 'feeling of pain' is a rigid designator

justifies its completely different character compared to 'watery' stuff – 'watery' too is a rigid Designator -, but because 'feeling of pain' is *set up from the start as something that is principally not constitutively linked to the physical*. So, what the argument was intended to show – the logical possibility of zombies, in the sense of the absolute heterogeneity of mental and physical states – has been *included in the premise from the outset* – obviously a *petitio principii*. The rigid designator argument, which at first appears to be promising, and which was to provide the theoretical basis for the presumed logical possibility of zombies, is not conclusive. Maybe the zombie friends will answer: 'But the zombie fiction remains *conceivable* and therefore *logically possible*'. But the attackable mere conceivability was to be replaced by the rigid designator argument, so that the renewed appeal to the mere conceivability is *circular*.

Some statements of Chalmers himself underline these considerations: "The secondary intension and the primary intension of 'pain' coincide" (Chalmers 1996b, 149): naturally, because the pain phenomenon is seen as standing for itself, i.e. as *not constitutively* being bound to the physical, while the wateriness of water, for example, is linked to H₂O with the consequence that the primary and secondary intension of 'water' *differ*. Chalmers himself states that "Kripke's apparatus of rigid designation and the like is not central" (149), which is consistent with the developed argumentation. What he, on the other hand, takes over positively from Kripke is the idea of the "logical possibility of dissociating physical states from the associated phenomenal states" (147) – so indeed precisely a "dissociation" of the physical and the phenomenal-mental as it is being *presupposed* in Chalmers' fiction of a zombie world.

Altogether, it seems to me, Chalmers' central arguments, as explained, are on the one hand not conclusive, on the other hand reducible to the undeniable fact that the phenomenal-mental 1st-person perspective is inaccessible from the outside, i.e. cannot be taken in principle by neuroscience which is arguing from the 3rd person perspective. Ultimately, this intuition also seems to underlie the claim used in the rigid-designator argument the sensation of pain to be body-independent. In short, the inaccessibility of the 1st-person perspective 'from the outside': this is probably the real core argument which is more or less hidden under intricate modal-logical considerations in this irritating book.

In the following I would like to develop an argument to *reconstruct* – not to discuss away – that 1st-person horizon:² The point is to show that the 'physical', *as far as it is organized as a living being*, develops *subject character* and thus an *own dimension*, in which *sense* emerges, which in higher animals occurs in the form of *phenomenal* perception. In conclusion some basic considerations about this project follow.

3. The Organismic 'Care Structure'

In my opinion, the concept of *care structure*, which Heidegger refers to human existence (Heidegger 1960, e.g. 196), offers a generally apt characterization of organismic existence: The basic condition of all living beings is *self-preservation* and therewith the concern for one's own existence. Self-preservation: This is not the way of existence of a boulder, for example, that may easily be preserved for thousands of years – or not. Whatever happens to the boulder, it is in any case *indifferent* to that. For the living organism, on the other hand, there is no eternal life either, but it *never is indifferent* to that what happens to it, on the contrary: In its being it is permanently concerned with nothing else but its own being. (We are not talking here about the very special situation of the spirit-being 'human being').

This *existential concern* is the basic character of all living beings. *Everything* in it is pervaded by that: Originating from evolutionary selection, it is – as it were *by definition, i.e. logically necessarily* – endowed with precisely those qualities and abilities that enable it (under the given environmental conditions) to maintain itself in existence. The *principle self-preservation*, to which all organisms are subject, means that they are *essentially concerned about their own being*.

² Of course not in showing concrete neuronal structures and functions which (in this point I agree with Chalmers) in no case can be *identified* with the phenomenal. – For the correlated brain physiological functions – see the revealing explanations in Donald 2001 and Damasio 2010.

For the individual, the principle self-preservation means the permanent monitoring and securing of his essential conditions of existence. For this purpose, it requires a *valuation system* that controls all external and internal changes with regard to their system compatibility and initiates appropriate reactions – a fundamental function of organismic systems. (In higher animals, *memory* plays an important role in this process: Valuations of earlier perceptions and actions are stored in it and can be recalled in similar situations. When we speak of 'valuation' in the following, memory is therefore always included).

But where does this elementary capacity of all living things – the secret of life, as it were – come from, by which it is distinguished from inanimate nature in a fundamental sense? The answer to this seemingly difficult question arises, I think, from a fundamentally 'Darwinian' argument: Only organisms capable of such control and assurance of the essential conditions of their existence were able to survive and reproduce.

Essential for the organism in the sense of the principle self-preservation is therefore the constant reference of all life processes to the conditions of its existence, i.e. to that which *constitutes its essence*, or in short its 'self'. With the principle self-preservation a self is constituted, and the organism thus as a 'self-ish' being, i.e. *reflexively* related to its *self*, and thereby a *subject*³. In this sense, organisms, even at the lowest levels of organization, are subjects, and as such permanently reflexively related to *themselves* by the *concern* for their own being.

But can self-preservation not also be implemented in *robots*? Mind you: Robots in today's sense – not to be confused with Chalmers' zombie twin, which also exists without consciousness, but is biologically perfectly equipped – while the "more-or-less serial virtual machines" that Daniel Dennett has in mind with regard to human consciousness are more similar to robots (Dennett 1991, e.g. 218 ff)⁴. In any case, the robot can be programmed to perform 'self-tests', such as checking the charge of its battery and looking for a socket when that needs to be charged. Similarly, other 'self-preservation actions' can be implemented – behaviors that are determined by the robot's *programming*. But isn't evolution, to which organismic self-preservation is owed, also a kind of 'programming', here in the form of natural selection? So what would be the difference to robots? Is the organism not also a robot? There is no doubt that more comprehensive 'self-maintenance programs' can be realized for robots than just looking for a power outlet when the battery charge drops. The self-tests can be extended and appropriate safety options can be programmed. A *principled limit* of such 'self-preservation programming' is not recognizable. Does the robot's lack of self-preservation thus ultimately depend on the more or less complex programming of its functions?

One of *Kant's* great insights is undoubtedly his interpretation of the organism as an *end in itself*, and that is to say as a purpose-means configuration in the form of *inner expediency* (Kant KU § 63 ff): All organs here are means to the end of the preservation of the organism, and this is conversely means to the end of the preservation of the organs, so that in the organism "*everything is an end and mutually also means*" (Kant KU § 66). Inner expediency thus means that the only end of all organismic achievements is the existence of the organism itself, i.e. its self-preservation. In contrast, the *robot* is conceived as a means to achieve a purpose defined by the programmer. The robot is thus a *means to an end*, but it is in any case not an end in itself, and its actions are accordingly not directed towards self-preservation – even if its behavioral routines provide for self-tests and appropriate measures to secure the device.

The decisive point is: When the battery charge drops, looking for a socket is by no means a purpose of the robot itself, but of the *programmer* to maintain the functionality of the device. Similarly the 'autonomous', *self-driving car* is programmed to prevent accidents, i.e. to ensure the integrity of the passengers, other road users and all the vehicles, including the self-driving car itself. But that is not the purpose of the self-driving car itself, but that of its designer. The self-driving car is just as little an end in itself as the self-charging robot. Both are merely a means to an end alien to them, set by the programmer or designer.

Thus it is also clear that the *natural* origin of the organism – which may at first appear to be a rather insignificant circumstance – actually means a *principled difference* compared to the technical

³ For a comprehensive analysis of the concept of subjectivity see Wetzel 2001, 8 ff.

⁴ For Dennett's assessment of the zombie fiction see Dennett 2007, 117 ff.

construct. Certain, evolution too can be understood as a kind of programming of living beings, and in this respect the comparison with robots is obvious. But what has *emerged from the selectionist fire* is essentially – quasi *logically necessary* (as already noted above) – an *end in itself*, i.e. a being on its own, which has been given the task of surviving in its habitat, with precisely those characteristics and abilities that were given to it by selection. To preserve its existence against possible threats: That is the essential existential orientation that results from its structure as an end in itself.

It is this *totalization of existential orientation* that distinguishes the organism from the robot. If, for example, the robot's 'clothing' is made of metal or plastic, which it does not take care of and does not have to take care of, since it is sufficiently stable dimensioned by the designer, the skin enclosing the living body is subject to permanent self-control in the sense of the principle self-preservation and is in a state of permanent regeneration and self-repair. The robot is a means to an end. This also applies, so to speak, to the organism – but in such a way that this is not an external end to it, but that this end is itself: a different formulation for the fact that it is a *self*, a *subject*. In short, this 'self-relatedness' is *essentially* due to its *natural* origin. For what has naturally come into being has come into being selectionistically, and what selection produces are, quasi *by definition*, beings destined for self-preservation. This very 'self-purpose' inherent in them, their *concern* for their own being, is indeed to be understood as the dowry of their *natural* origin. On the other hand, the robot is not able to shed its non-natural, technical origin. It owes its existence and function to the constructor's purposes pursued with it, which are *external to its existence* and do not establish an *existential care structure*.

4. The emergence of sense

If a subject pursues a purpose with his activity, then the activity has a *sense* for the subject. For the organism, food intake has sense in terms of its existential purpose of maintaining vital functions. Switching to another behaviour, such as attack or flight, has sense in a prey or danger situation. In this way sense results from the purpose of an activity. Purposes constitute sense for the subject pursuing purposes relevant for it. Thus sense can be taken as *subject relevance*.

This also makes it clear *for whom* sense is constituted thereby: The actions of the *robot* realize the purposes specified by its designer. Therefore they make sense for the designer, not for the robot itself. The situation is different in the case of the *organism*: All organismic processes are subject to the principle self-preservation and are thus directed towards the realisation of purposes that are purposes of the organism itself. The organism is an end in itself, and this means that the organismic processes make sense *for it itself*. In short: With the *determination* of the organism, as defined by the general purpose of self-preservation, a *sense dimension* to be assigned to it is also constituted.

Incidentally, in relation to the organism *determination* is probably a more appropriate term than 'purpose', because the use of 'purpose' is connoted with mental intent and planning and is thus inadequate for the organism in general. In the following, therefore, the more general term *determination* will be used instead of 'purpose', while this will be reserved for the human domain.

The determination of organismic subjects is self-preservation. Through this cardinal determination of all life, *sense* is constituted for the organismic subject. Of course, this is no longer a *physical fact*⁵ that could be empirically determined and measured, but a characteristic that can only be grasped by the subject himself: the *subject relevance* of his own life process, which is realized in the cooperation of perception, valuation and behavior. Elsewhere I have briefly referred to this trinitas of perception, valuation and behavioral action as the *perc-val-act system*.⁶ It forms the organismic framework for realizing sense, and *in this way it can be studied how something ontologically completely different – sense – can arise on a physical basis*.

The key term for understanding these relationships is the concept of *emergence*. I have developed that point in detail elsewhere (Wandschneider 2015) and only briefly outline the basic idea here: Emergence is a *systemic* phenomenon: A system has – fundamentally seen – properties and

⁵ On the question of the specific mode of being of the psychic see Wandschneider 2016 a; 2016 b.

⁶ In Wandschneider 2015 as *percept-act-system*, in Wandschneider 2016b more precisely as *percept-val-act-system*.

regularities other than its subsystems by which it is constituted. Instructive is the example given by Konrad Lorenz of the electrical oscillation circuit, which owes the oscillation behavior to its components, coil and capacitor, which themselves do not have such oscillation properties (Lorenz 1973, 47 ff).

Thus by system formation new laws *emerge, system laws*. The *perc-val-act system*, which is constitutive for the organism, is physically composed in all its parts, but on the level of the *perc-val-act system* itself, different, purely subject-related properties emerge – subject relevance or sense – that can no longer be qualified as *physical*. They are indeed neurally realized, but what perceptually emerges on the *perc-val-act level for the subject himself are ideal sense-contents* (note: *not sensual contents!*) (Wandschneider 2016 a): precisely that which is relevant for their survival. Qua 'sense', the 'merely physical' is exceeded: as an *existential demand* on the subject, as a determination that motivates it to behave according to the principle self-preservation. The valuation of the sensory stimulus – in short: the *valuation response* – is a genuine *addition of the subject himself*. It is the valuation response that controls the activity of the subject, not the original sensory signal, because the subject *qua subject* does not act on the neural basis, but on the *perc-val-act level*, which is determined by the cooperation of perception, valuation and action, through which sense is constituted.

It is therefore clear that 'sense' is not something that would be visible or measurable to the empirically working researcher, but is *sense only for the organismic subject itself*, an exclusive format of the subject perspective, so to speak – in humans one would speak of the 1st-person perspective. For this reason, the biologist or brain researcher, who as such always takes the 3rd person perspective, will never find anything like 'sense' – which makes the notorious empirical skepticism about 'sense' as understandable as insubstantial. This subject-relatedness means that *I* cannot know 'what it is like to be a *bat*' (Nagel 1981) – of course, because I myself am not the bat subject.

In this context a clarification is necessary: The talk of 'sense' may seem strange in view of lower organisms, which are only able to react to very elementary stimuli such as light intensity or acidity. However, they too are subject to the principle self-preservation and value the stimuli received. They also exist as subjects and as such not on the purely neural, but on the *perc-val-act level*: Qua valuation the sensory stimuli have subject relevance and therefore sense for them as well – but, and this is the reason for the strangeness of the attribution of sense in this case, *they don't 'know' about it*. Neither the rose nor the earthworm is able to grasp this sense *as sense*. But everything that happens to them is valued by the valuation system in terms of the organismic general purpose of self-preservation and thus has indeed existential sense for them. How the lower organism reacts to stimuli is, according to the principle self-preservation, *objectively* sense-ful for it, but the process of sense-constitution happens implicitly, as it were, 'behind its back'. The existential sense is not explicitly grasped *as such*, but it is *objectively* given and thus indeed efficient in controlling the behavior.

The situation is completely different in the case of *higher animals*, which have *phenomenal perception* (including sentiences). Phenomenal perception is to be understood as a differentiated *perception scenario* that not only presents the individual a (species-specific) more or less structured view of the reality it encounters, but also contains 'qualia' such as color impressions, tactile, taste, olfactory, pain sensations etc. How neural, i.e. physical brain processes are capable to produce such phenomenal perceptions, is considered the *hard problem* of the philosophy of mind (Chalmers 1996a, 1996b). The *explanation gap* (Levine 1983, 1993), which is claimed in this regard, turns out, as I concede, to be more difficult than I have argued elsewhere (Wandschneider 2015). Nevertheless, I would like to refer to the considerations developed there, which come to the conclusion that the phenomenal perception, can – and, as I think, only in this way – be explained as an *emergence* phenomenon of the *perc-val-act system*, due to its essential relatedness to valuation and behavior.

The perception scenario presented by the phenomenal perception offers the individual, in addition to sensations of internal sensitivities, a (species-specific) detailed view of the external reality, e.g. the spatial relationships of the objects to one another ('before', 'between', 'behind', 'above', 'under'), together with the valuations assigned to the perceptions, e.g. the 'threat' of a bird of prey contour or the 'attractiveness' of a prey object. What is decisive in this context is that the behavior is thus no longer controlled in a purely reflective automatic manner through the respective valuation response as in lower animals, but it rather has the scenario of phenomenal perception in front of it and is predominantly oriented to *that*. In contrast to the primitive, purely valuation-reflective beha-

vioral control, behavior in higher animals is thus increasingly linked to and controlled by phenomenal perception. I would like to call this the *principle of phenomenally-instructed behavior*. The differentiated perception scenario, in which the valuations are integrated, allows for far more flexible behavior than is possible through genetically rigidly predetermined reflex reactions, which require a constant habitat, but are no longer effective in rapidly changing environments. This increase in behavioral efficiency makes the evolutionary tendency towards phenomenal perception in the course of the faunal development seem plausible.⁷

The *principle of phenomenally-instructed behavior* in higher organisms means that the behavior is based on a differentiated perception scenario, whereby the perceptual contents are *perceived as senseful* by the subject. For the mouse, for example, the bird of prey contour appearing in its environmental scenario immediately gains priority with regard to its behavioral decisions, with the result that the flight reaction appears to be the only sensible option in view of the numerous other behavioral options that the perception scenario also offers. *Sense* is thus given in a more concrete, more explicit form. The negative sense of the prey contour stamps it as an *aversive object* and is perceived in this form. And in sensations like 'lustful', 'disgusting', 'painful' etc., the associated positive or negative valuation itself is directly experienced as senseful.

In this respect, in the case of higher animals, whose behavior is controlled via the phenomenal perception scenario, one can speak of a *sense dimension*, the sense contents of which are *explicitly perceived as senseful*. And as the sense character that is pervading all organismic processes, in the form of *phenomenal perception* now becomes explicitly perceptually graspable, too, a *new subject dimension* has been established, in which processing procedures become possible that already have the character of *elementary mental operations*. An obstacle, for example, that blocks the path to a goal, is perceived as '*disturbing*'. At the same time, the perceptual scenario, which, as already mentioned, also presents the spatial relationships of the obstacle in its environment and with regard to the goal, also presents clues on how to remedy the disturbance, for example the possibility of skirting the obstacle.

Of course, the possibility to recognize an obstacle and to make the decision to avoid it is also available to the robot. But with regard to the considerations developed above, it is clear that the robot does not exist as an *end in itself, as a subject*, but follows an external purpose – the purpose given to it by the designer. So only for the designer do the robot's actions make *sense*. The algorithmic decision-making competence given to the robot so does not change anything in regard to its notorious *sense dementia*. He is just doing physical processes, *not mental operations*.

The proper *sentiences*, such as lust, pain, disgust, smells, etc., form a subclass of the phenomenal perceptions. They are tinted strongly subjectively and can thus be understood as *valuation perceptions* of the (perceptual) valuations carried out by the subject. I have dealt with this distinction between valuations of perceptions and valuation perceptions in detail several times (e.g. Wandschneider 2015; 2016b) and here I will only mention a few points that are of interest in view of the constitution of sense in the perception:

Sentiences are to be understood as an important *completion* of the phenomenal perception: If the contact with a flame would only trigger an avoidance reflex *without* sensation of pain, the principle of phenomenally-instructed behavior would be broken and the behavior-controlling sense of perception would be devalued. Another important function can be recognized, for example, in the sweet-sensation: In this, the physiological success, i.e. the increase in blood sugar, which occurs only with a delay, is, so to speak, symbolically anticipated and is thus immediately available for behavior control – for example, in order to be able to react immediately in competition with conspecifics for food. The sweet-sensation is something like the preliminary headline of the delayed onset of blood sugar increase and thus an essential, behaviorally effective *carrier of sense* at the level of higher animals, which is characterized by the principle of phenomenally-instructed behavior. The naked 'measured value' of the sugar concentration in the food would not be sufficient for this, because it would also have to be *valuated* by the organism, just like everything else that the organism encounters. But sugar measurement plus valuation: That's exactly what makes up the *sentience* of

⁷ For this in detail Wandschneider 2015

sweetness. In short, sentences are to be understood as a necessary completion of the phenomenal perception and the sense dimension constituted by it.

5. Some Fundamental Considerations

The analysis of Chalmers's arguments carried out above had shown, among other things, that his view of what makes the phenomenal-mental substantial seems to result in the basic privacy of the inner dimension. I have undertaken to reconstruct the basic structure of this privacy here according to the systemic phenomenon of *emergence* in view of the *principle self-preservation* that pervades all life.

Actually, it is obvious that self-preservation as a generative principle must lead (a) to *subject formation* and thus to *privacy* and (b) to the establishment of a *sense dimension* (because of the subject relevance of all life processes). This seems almost to be *logically necessary*, because, as explained, self-preservation always implies something like selfness and sense. In this respect, Chalmers' zombie twin as well as the creatures of the zombie world would have to develop selfness and sense phenomena. Yes, one could even come to the conclusion that this is to take place *in all possible worlds* in which the principle self-preservation can be implemented, regardless of *how* it is realized. Because self-preservation is self-preservation; what can preserve itself can preserve itself, whether as an organism made of flesh and blood or as a little green martian. And what is subject to the principle self-preservation, as explained, must then also develop subjectivity ('privacy') and sense ('survival sense').

It is noteworthy that the *natural-biological* origin of the organism – which might initially appear to be a rather contingent circumstance – makes a *momentous difference on principle* compared to inorganic objects and artificial-technical constructs (robots), namely in regard to the constitution of sense and thus (in higher animals) phenomenal-soul. Put in a formula (which I'm liking to repeat), this means that *soul and self-preservation essentially belong together*. Thereby it should not be seen as a burdening mortgage that using the principle self-preservation the 'secret of life', as it were, is taken into service. The possibility of self-organization of 'hypercycles' (Eigen 1976) and selectionist arguments, also for the development of a valuation system, are sufficient for a *principled* understanding today.

The second fundamental point concerns the phenomenon of *emergence*. Often suspected as a kind of magic formula, it is, viewed impartially, the simple systemic fact that cannot be disputed that the whole is more than the sum of its parts, that the properties of a system can therefore be completely different from those of its components. The system under consideration here is what I call the perc-val-act system, which, in the cooperation of perception, valuation and behavior, as a whole realizes what the *subject* actually constitutes. The fact that *sense* emerges in this system is, as I said, due to the principle self-preservation. But the fact that sense is no longer a '*mere physical*' fact (see above) is to be seen as an emergence phenomenon: as the certainly paradoxical possibility of realizing the non-physical in a system that is nevertheless physically constituted. I described this seemingly paradoxical configuration in a previous work as follows: "If it is true that the mental can be understood as an emergence phenomenon of a physical system – such as the brain – then the mental is indeed physically based, but as a *holistic phenomenon* of the system as a whole it is also a *novel type* of property compared to the properties of the subsystems, which as such have 'mere physical' character, in other words: Although we have to speak of a physical system as a whole, the mental, as a phenomenon of wholeness, is *typically different* from the physical properties of the subsystems which, on the other hand, are, as it were, the 'merely physical' (i.e. without mental properties). *In this respect*, there is a *dualism* of physical and mental phenomena, but this in an overall physical system that is at the same time *not dualistic*, in short: Emergence relationships lead beyond the elementary-physical to typical novel properties, without, however, exceeding the overall range of the physical" (Wandschneider 1999, 71 f).

Chalmers, flirting with dualism and panpsychism in search of a theory of the phenomenal-mental, could therefore, I think, be happy with emergentism. But he regards this theory to be one that only concerns the physical and remains attached to the physical (378 f), according to "the general principle that from structure and dynamics, one can infer only structure and dynamics" (Chal-

mers 2003, Ch. 7). But in the search for *psycho-physical laws*, which Chalmers would like to place alongside the laws of nature in order to take into account phenomenal-mental forms, emergence phenomena are expected to become important – as in the previous *sense* as an emergent property of the perc-val-act level was interpreted. But all of this is based on the laws of nature, as it were as a *logic* inherent in nature itself: That being natural thus always already contains the possibility of psychic is then no longer such a distant thought, that – in connection with the concept of emergence – could replace – rationalize – the rather strange hypothesis of panpsychism.

A question that comes to my mind at the end: What if the emergentist reconstruction of the phenomenal-mental should succeed in the context of 'our' physical world? Are there any other, *higher* reasons in order to nevertheless still hold onto a dualism?

References

- Bieri, Peter (ed. 1981) *Analytische Philosophie des Geistes*. Königstein/Ts. 1981
- Blackmore, Susan (ed. 2007) *Gespräche über Bewusstsein*. Frankfurt/M. 2007
- Chalmers, David (1996a) Das Rätsel des bewussten Erlebens, in: *Spektrum der Wissenschaft* 1996/2, 40–47
- Chalmers, David (1996b) *The Conscious Mind. In Search of a Fundamental Theory*. New York, Oxford 1996
- Chalmers, David J. (2003) Consciousness and its Place in Nature, in: Stich/Warfield (ed. 2003), 102–142
- Chalmers, David J. (2007) 'Ich bin mir bewusst: Er ist nur ein Zombie', in: Blackmore (ed. 2007) 56–74
- Damasio, Antonio (2011) *Selbst ist der Mensch. Körper, Geist und die Entstehung des menschlichen Bewusstseins*. München 2011
- Donald, Merlin (2001) *Triumph des Bewusstseins. Die Evolution des menschlichen Geistes*. Frankfurt/M. 2001
- Dennett, Daniel C. (1991) *Consciousness Explained*, New York, Boston, London 1991
- Dennett, Daniel C. (2007) Wir müssen unsere Intuitionen über das Bewusstsein aufgeben, in: Blackmore (ed. 2007), 115–132
- Eigen, Manfred (1976) Wie entsteht Information? Prinzipien der Selbstorganisation in der Biologie, in: *Berichte der Bunsen-Gesellschaft für Physikalische Chemie* 80 (1976), 1059–1081
- Hegel, Georg Wilhelm Friedrich: *Werke in 20 Bänden*, ed. Eva Moldenhauer und Karl Markus Michel. Frankfurt/M. 1969 ff (Zitierweise, Beispiel: '9.48 Zus.' verweist auf: *Werke* Bd. 9, S. 48, Zusatz)
- Heidegger, Martin (1960) *Sein und Zeit*. Tübingen ⁹1960
- Hösle, Vittorio/ Koslowski, Peter/ Schenk, Richard (ed. 1999) *Jahrbuch für Philosophie des Forschungsinstituts für Philosophie Hannover*, Bd. 10. Wien 1999
- Jackson, Frank C. (1986) What Mary Didn't Know, in: *Journal of Philosophy* 83 (1986), 291–295 (doi.org/10.2307%2F2026143)
- Kant (KU) *Kritik der Urteilskraft*. Zitiert nach der 3. Originalausgabe. Berlin 1799 (Philosophische Bibliothek, Berlin 1959)
- Kripke, Saul A. 1972, (hier deutsch 1981): *Name und Notwendigkeit*, Frankfurt/M. 1981
- Langthaler, Rudolf/ Hofer, Michael (ed. 2016) *Naturphilosophie. Traditionelle Themen und gegenwärtige Herausforderungen*. Wien 2016
- Levine, Joseph (1983) Materialism and Qualia: The Explanatory Gap, in: *Pacific Philosophical Quarterly* 64, p. 354–361
- Levine, Joseph (1993) On Leaving Out What It's Like, in: Davies/Humphreys (ed. 1993), 121–136
- Lorenz, Konrad (1973) *Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens*. München, Zürich 1973
- Nagel, Thomas (1981) Wie ist es, eine Fledermaus zu sein? in: Bieri (ed. 1981), 261–275
- Pauen, Michael (2005) *Grundprobleme einer Philosophie des Geistes. Eine Einführung*, Frankfurt/M. ⁴2005

- Popper, Karl Raimund/ Eccles, John Carew (1977) *The Self and Its Brain*, Berlin/ Heidelberg/ London/ New York 1977
- Stich, Stephen P./ Warfield, Ted A. (ed. 2003): *Blackwell Guide to the Philosophy of Mind*
- Wandschneider, Dieter (1999) Das Problem der Emergenz von Psychischem – im Anschluß an Hegels Theorie der Empfindung, in: Höhle/Koslowski/Schenk (ed. 1999), 69–95
- Wandschneider, Dieter (2015) Das 'eigentlich schwierige Problem' phänomenaler Wahrnehmung, in: *Zeitschrift für philosophische Forschung* 69, 550–568.
- Wandschneider, Dieter (2016a) Zur Seinsweise des Psychischen, in: *Zeitschrift für philosophische Forschung* 70 (2016), 28–46
- Wandschneider, Dieter (2016b) Phänomenale Wahrnehmung und die Idealisierungstendenz der Natur, in: Langthaler/ Hofer (ed. 2016), 50–71
- Wetzell, Manfred (2001) *Prinzip Subjektivität. Allgemeine Theorie. Erster Halbband.* Würzburg 2001