Decisions

Dr. rer. nat. Dr. phil. Anja Stemme

Philosophical Practise, Pradler Str. 74, A-6020 Innsbruck, Austria

anja_at_stemme_dot_eu

30. Dezember 2011

Abstract

Today, neuroscience claims to explain human behavior in detail and in general. Thus *our* decisions are considered to be taken on a neuronal level. In this essay a critical review of respective findings is provided which reveals that within Neuroscience a consistent and usable concept of a "decision" is completely missing but rather "ad-hoc" ideas guide experimental investigations with respect to decisions or the "free will", respectively. The consequences are highly contradictory empirical findings and neither the research area of machine consciousness, nor the attempts to model the neuronal correlate of consciousness or "synthetic phenomenology" are able to provide conceptual support and resolve severe inconsistencies. But "synthetic phenomenology", a research area aiming in the generation of artificial "phenomenal states" for machines, claims to have its roots within the philosophy of Edmund Husserl who, ironically, expressed severe criticism with respect to knowledge in science in general. In following this hint this essay tries to clarify our scientific knowledge with respect to brain and neurons in principal by investigating the relationship between *world*, *neuronal* and *phenomenal* states using the method of simple observation with the conceptual support of a computer analogy to understand these relations. This path finally (re)opens a surprising perspective with respect to decisive concepts and reveals the general limitations of current neuronal attempts to understand the brain and its relationship to *ourselves*: Obviously it is impossible to understand the brain without its user.

Or in other words: This essay demonstrates that the scientific consideration of brain and neurons leads necessarily to the recognition of something which formerly might have been called "soul".

Contents

1 Decisions

Once upon a time a father and his child walked along a street. Suddenly the child started to run towards the street obviously in order to cross it because it has seen some friends on the other side. "Stop it,", the father cried out aloud and very anxious, "there are a lot of cars coming down the road. Didn't you realize that?" Upon the cry the child ***decided*** *to turn its head towards the father and stop running.*

Soon et al. (2008) conducted a fMRI study where they tried to determine the outcome of a person's *decision* regarding which index finger to move out of her brain activity using pattern recognition techniques. Following their explanations it seems to be possible to decode at least an outstanding finger movement of a person from her brain activity already about 7 to 10 s before the movement actually occurs. This is a comparatively long period between the decision of a brain which finger to move and the actual movement itself. The timescale in the scene described above is obviously much shorter. Hence the child did not actually *decide* to turn it's head upon the father's cry but we face a simple case of reflex action where no conscious decision or any further brain preplanning of a movement was necessary or even possible. Contrary to the work of Soon et al. (2008) who asked the participants in their study to *freely* choose which index finger to move "when they felt the urge to do so", following and further improving the famous experiments of Benjamin Libet (Libet, 1985).

However, in both cases the activity of motor neurons was necessary. So *why* such a long period of brain preplanning if it is obviously *possible* to activate motor neurons as well on a much shorter time scale as illustrated in the child's example. Moreover, the participants in Soon et al. (2008) *reported* their decision to be formed within about one second before the movement. In the experiments conducted by Libet (1985)[1] the decisions, as reported by the subjects, occurred even about 200 ms prior to the movement. Hence the *reported* decision times under experimental conditions seem to be much more accurate in relationship to the actual movement and much less contradictory with respect to the child and its head movement, leaving even time for a conscious decision to actually turn the head following the fathers cry and stop running.

[1] Replicated by Haggard and Eimer (1999), see as well Haggard (2008) for an overview; Trevena and Miller (2010) for an alternate interpretation of the experiments.

So what actually is a "decision", from a neuronal point of view?

2 Decisive brain areas

Soon et al. (2008) argued that the brain activity allowing the prediction of the upcoming motor decision is associated with brain areas not explicitly responsible for motoric actions. Hence these areas, in particular the frontopolar cortex and the precuneus, an area within the parietal cortex, might represent a more general responsibility for various kinds of decisions, a cortical network for "high level controls" that "can begin to shape an upcoming decision long before it enters awareness" (Soon et al., 2008, p.3).

Goldberg et al. (2008) investigated the brain activity associated with decision making in a quite similar fMRI study. They measured the brain activity of participants under two conditions: a "determined" condition where the participants were asked to follow an external instruction to press either a left or right button, and a "free will" condition where the participants were required to follow a "voluntary decision" regarding the button to press. As the setups of the two experimental conditions were otherwise identically, both requiring to press a button with the index or middle finger, the authors aimed to draw conclusions regarding the brain areas incorporated in free will decisions.

Contrary to the work of Soon and colleagues the authors were not aiming to decode the concrete outcome of a decision but to analyze the brain activation patterns for "free will" compared to "determined" conditions. In the first case they detected a substantial level of activation in the right inferior parietal cortex and concluded that this cortex area might be part of an intrinsic cortical system dealing with "internally oriented and endogenous mental processes".

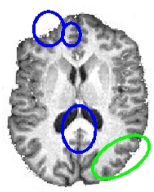

Abbildung 1: Sketch of brain areas identified in the study of Soon et al. (2008) (blue) and Goldberg et al. (2008) (green).

However, the cortical areas of major importance identified by Goldberg et al. (2008) differed from those identified by Soon et al. (2008) although both experimental setups dealt with simple "which-finger-to-move"-decisions in a comparable context: Goldberg and colleagues identified the right inferior parietal cortex to play a decisive role, and particularly *not* the precuneus, whereas Soon and colleagues concluded that the frontopolar cortex takes the decision which is stored in the precuneus until it enters awareness. Nevertheless, both studies claimed the responsibility of the identified cortical areas as part of a general network of neurons dealing with decision making, identifying even a "self-related role" for this area (Goldberg et al., 2008), respectively a cortical network for "high-level-controls" (Soon et al., 2008).

Ignoring first of all the aspect that the latter achieved a prediction accuracy of only about 60%, not substantially much above the chance level, the brains in the study of Soon and colleagues had much more time to form decision, awareness and action. If a decision takes such a long time is it possible that the brains in the study of Goldberg and colleagues did not actually *decide* anything?

What *is* a decision and *where* to find it?

3 Attention and the brain's philosophy

The father in the introductory example called upon the kids *attention*. In our every day life outside (but probably even inside) laboratories for consciousness research people are confronted with demands on their *attention*. We ask people to pay attention because we assume that they are able to somehow *freely* direct their attention. If the "unconscious driver"[2] passes a red traffic light and causes an accident he will receive a hard justice punishment because it is assumed that he would have been able to pay attention to traffic lights instead of dreaming about the last hours. It is assumed that people have a true choice here.

However, modern brain theorists started to question the accuracy of the justice system as following their opinion no such thing as a "free will" exists and hence no possibility to "freely" direct the attention. Thus besides the recently raised research fields of Neuroethics, Neurosocio-

[2]An example initially raised by Armstrong (1968), used f.e. by Dennett (1991) to illustrate the common misleading concept of consciousness and by Gamez (2008) as an example that a limited amount of behavior can be carried out unconsciously.

logy, Neurophilosophy etc. "Neurojurisprudence" seems to be highly required as well. The line is clear cut: Whereas at the beginning we seem to have somehow stumbled from brain research into central areas of philosophy by touching the question of free will (Libet, 1985) we are now sure that the only suitable place to look for central aspects of human experience is the brain, although we still lack usable concepts about the items we are talking about and their possible relationship with or dependence on neuronal activities.

The next item in this line is the question of the self and Goldberg et al. (2008) point in this respect to the work of Daniel Wegner (2003, 2004) to underline the obvious "tight link" between voluntary decisions and something commonly referred to as the self "since both appear to be so clearly linked to the same network of intrinsic processing."[3]

Thus the cortical areas identified to be activated in circumstances that require decisions seems to be the same as those responsible for the generation of our sense of self; *we find ourselves in our decisions*.

However, which cortical areas: The ones identified by Soon et al. (2008) or by Goldberg et al. (2008) or both?

The frontiers between neuroscience and philosophy have long been torn down and the only place left to look for a "self", "attention", the "free will" or "decisions" seems to be the brain.

Where to find these decisions?

4 How the brain makes decisions

Wang (2008) analyzes decision making processes on the level of recurrent neuronal circuits, hence on a level that does not care about attention, free will or self: From a reasonable scientific point of view neurons *themselves* do not attend, they spike; they have no will at all but generate action potentials following natural laws beyond dispute. Thus we might understand what we actually mean by a "decision" when we consider basic neural operations.

The models reviewed by Wang describe neuronal activities following the presentation of various stimuli (equals the explicit excitation of some neurons within the model) and generate responses (i.e. a certain pattern of neuronal activity within the model) that are similar to those

[3] Goldberg et al. (2008), p. 600

emitted by experimental participants usually with respect to response times and error rates.[4] The experimental tasks usually modeled concern tasks that require to discriminate between different perceptual stimuli, for example vibrotactile, visual or auditory and decide upon the correct response. In these experiments participants are usually confronted with different types or levels of stimuli and are to judge whether they are the same or different. In these tasks we find response times about a second and it is thus assumed that following the integration of sensory data (i.e. the processing of the presented stimuli within the brain) the decision is taken (fig. 2).

But: Is the ability to *discriminate* between different stimuli already identical to a concrete *decision*? If we consider the ability to discriminate between different stimuli already as a concrete decision we are forced to admit that the child in the above example took a decision as well: Following the integration of sensory data (the father's cry and the current visual scene) - which is thought to be the base of the decision making process - the child's brain *decided* to move the head.

But similar *decisions* were decode-able from brain activity already seven seconds before the actual movement. So why should brains need sometimes such a long delay between a decision and a movement (Soon et al., 2008) and in other cases (the child's brain; Goldberg et al., 2008; Wang, 2008) obviously not?

Is it possible that we are to differentiate between decisions regarding the discrimination of perceptual stimuli and perhaps somehow more arbitrary "which-finger-to-move" decisions? In this respect Wang himself provides some clarification by referring to the study of Soon:

"Moreover, a recent study showed that a free motor decision could be decoded from BOLD signals in the frontal and parietal cortex long (seconds) before it entered awareness (Soon et al., 2008)."[5]

and further concludes that

"the parietal and frontal cortices form a core brain system for accumulation of data and categorical choice formation in perceptual judgements."[6]

Thus at least Wang himself does not see different types of decisions in this respect.

[4] See f.e. Stemme and Deco (2008) for similar models.
[5] Wang (2008, p.219)
[6] Wang (2008, p.219)

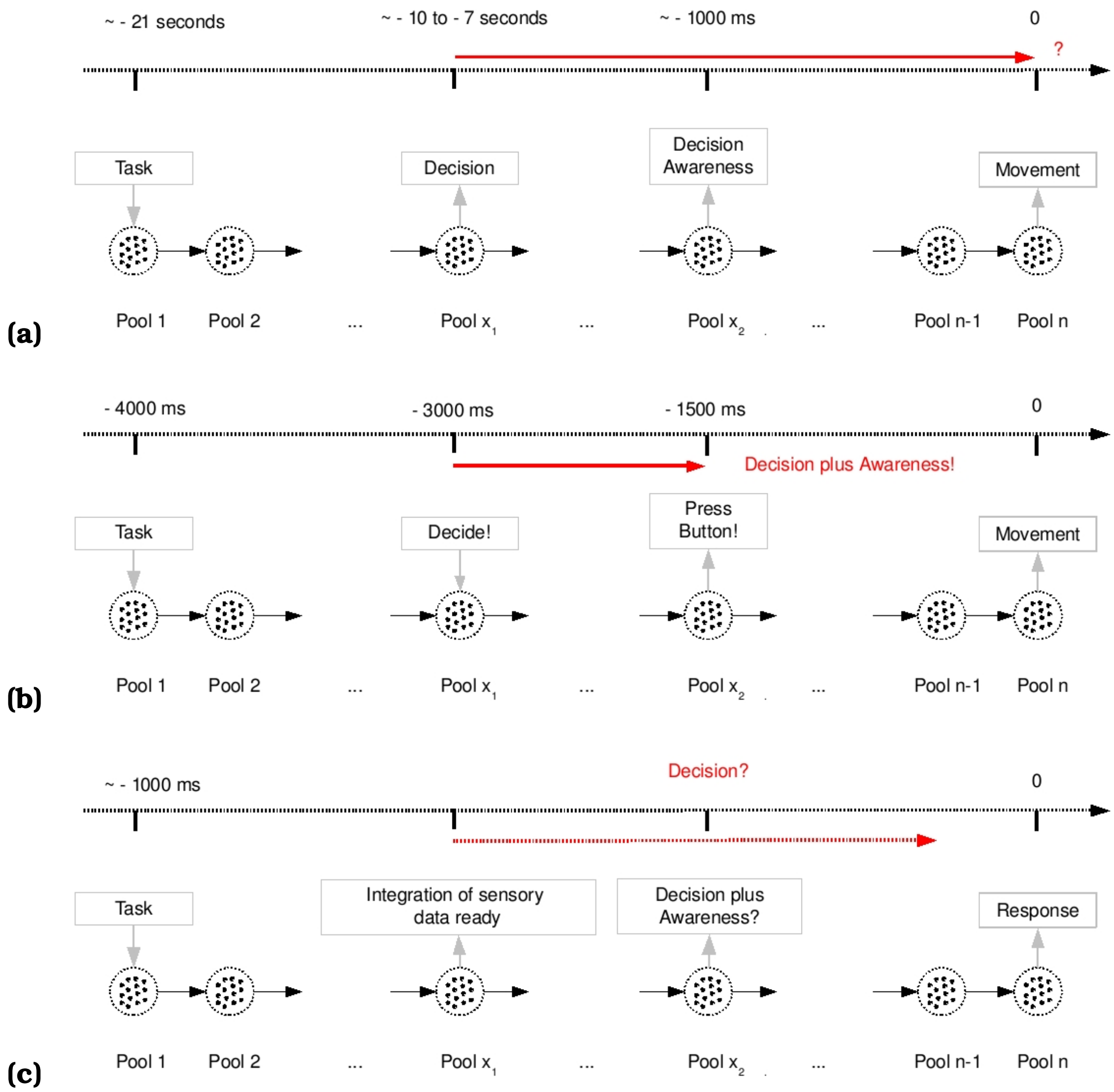

Abbildung 2: Comparison of the timing dynamics for decisions from a neuronal perspective. We assume that there are a range of neurons, organized in pools for the purpose of illustration, involved in different steps to form a decision. **(a)** In the study of Soon et al. (2008) the brains needed obviously about 10 seconds to find a decision (activity of possibly distributed neurons labeled x_1) and after further 6 to 9 seconds this decision was aware to the subject (x_2). **(b)** In the study of Goldberg et al. (2008) the participants were forced to decide within a range of 1500 ms and afterwards to press the button. Hence the awareness of the decision, if any occurred, must have been much faster in these experiments. **(c)** In the type of experiments considered for the modeling of "decision making processes" as reviewed by Wang (2008) decisions are taken even faster.

5 Conscious and unconscious decisions?

Following Wang (2008) we have to consider a decision to be somehow identical to the integration of sensory data, taking about a second including the corresponding motor response, whereas following Soon et al. (2008) the time span from a brain's decision to the subject's awareness of this decision *and* the brain's action turned out to be several seconds. The slight discrepancy with respect to the time span between a brain's decision and a brain's action might thus be related to the awareness of a brain's decision. Potentially, the participants in the perceptual discrimination experiments reviewed by Wang (2008) were not aware of their decisions.

Following Wang, and potentially as well Soon, we might therefore, for the beginning, define a decision to be equivalent to the integration of sensory data, at least from a brain's perspective, and separate this decision from its awareness by the subject. Hence it might be necessary to differentiate between the brain's decision (brain level, unconscious to the subject) and the subject's awareness of a brain's decision (subject level, conscious).

In the study of Soon the integration of "sensory data" turned out to be ready after about ten seconds (brain level), possibly due to the lack of clear sensory data to integrate, contrary to the cases Wang considered. Following the integration of sensory data the awareness of this decision occurred after another six to nine seconds (subject level) in the study of Soon. Whereas the subjects in the experiments Wang considered were obviously not aware of their decisions. The same holds for the subjects in the study of Goldberg et al. (2008, not enough time, see fig. 2) and of course for our child in the introductory example.

Following this first cautious brain-level definition of a decision and considering the timing dynamics observed in Soon's study we find us thus forced to state that all subjects considered by Goldberg and Wang as well as our child took somehow "unconscious decisions". Hence we are forced to assume that the conscious part of a decision is unnecessary or has at least only minor and non-contradictory proportions to the final motor response because otherwise the "decision making" models reviewed by Wang (2008) or Goldberg's investigations would lack an important component.

But if the subject level of a decision is not necessary for any operation why did the brains somehow "wait" for the subject's awareness (Soon et al., 2008) if they are able to initiate movements obviously immediately after their decision as the other studies indicate (Wang, 2008; Goldberg et al., 2008)?

Moreover, neurons operate on the scale of milliseconds: Where to find a seven-seconds-delay-loop in the brain, what happens during this time on the neuronal level? Further more, Soon detected a prediction accuracy of only 60% for the unconscious brain level of the decision which implies at least the possibility that the conscious part of a decision might differ from its unconscious counter part. Thus our first cautious brain-level definition of a decision did not take us any further with experimental results being still contradictory:

1. If we assume that a decision has to be defined on the level of neuronal activities (brain level) and has to be separated from its awareness (subject level) we have to assume unconscious decisions for the participants in the studies of Goldberg et al. (2008) and Wang (2008). However, if brains are able to initiate motor responses based on the unconscious part of a decision we fail to explain the delay between decision and action in Soon's study. Moreover, the low prediction accuracy for the unconscious brain level of the decision forces us to assume that subject level and brain level of a decision can differ which in turn implies that Wang and Goldberg miss an important component in their studies.

2. Therefore, if we assume a conscious decision of the subjects in all cases considered (Soon, Wang, Goldberg, child), we still stumble with the question of the delay in Soon's study. Why should brains sometimes need six to nine seconds to "produce" decision awareness in the subjects (Soon) and in the other cases (Wang, Goldberg, child) not?

The problem of the delayed brain's action in Soon's study persists and seems to be related to the question of awareness. Obviously we have to dig into the "hard problem"[7] and ask what the functional role of the

[7] Chalmers (1996) differentiated the "easy problems" of consciousness, i.e. to explain the functional role of specific aspects of "consciousness" (attention, working memory, learning, processing of visual stimuli, etc), and the hard problem: to explain why consciousness, why phenomenal states exist at all. In any (scientific) explanation

(phenomenal) awareness of a decision could be or in other words: What do we actually mean by "entering awareness" or "consciousness"?

David Gamez considered amongst the benefits of the work on machine consciousness the possibility to test theories about consciousness and to deepen our understanding of consciousness in the brain, besides of course the advantage of constructing more intelligent machines.[8] The investigation of these benefits might help us to solve our still persisting decisive question on a scientific level without digging too deep into the philosophical debates on consciousness.

Gamez differentiated machine consciousness starting with systems that simply replicate human behavior (labeled MC1) or show "cognitive characteristics" commonly associated with consciousness (MC2) up to attempts to create "real artificial consciousness" i.e. "phenomenally conscious machines" (labeled MC4).[9] With the label "MC3" Gamez summarizes

"machines with an architecture that is claimed to be a cause or correlate of human consciousness"

and here we meet again neuronal model architectures similar to those reviewed by Wang (2008). Gamez refers for example to the work of Stanislas Dehaene and colleagues who create computer models of the neuronal correlate of consciousness (NCC)and hence these models might provide an essential hint regarding our question of the delayed awareness and the hidden question of consciousness.

6 Consciousness and it's neural correlate

Dehaene and colleagues suggested a neuronal model which implements a kind of "global workspace"[10] to explain a range of psychological effects as "inattentional blindness"[11] or the "attentional blink"[12], for example.

of conscious aspects "phenomenal states" do not and cannot play any explanatory or causal role.

[8] Gamez (2008, p.907)

[9] Gamez (2008, p.888)

[10] The workspace theory was initially suggested by Baars (1988)

[11] Refers to the circumstance that experimental participants are unable to report on stimuli presented unexpectedly outside the focus of their attention but nevertheless these stimuli can be shown to affect their answers ("priming"). Dehaene and Changeux (2005)

[12] Experimental participants seem to be unable to detect a second target within about half second after the presentation of a first target. But again priming effects are measurable. Dehaene et al. (2003)

Within their model "consciousness" is closely associated with the ability to memorize hence with working memory. This working memory is thought to be implemented by a set of neuronal pools which represent the "global workspace" and which have access to distributed cortical areas, processing sensory input or generating motor actions.

Whenever there are neuronal activities or "processes" in the brain, these activities are "broadcasted" to gain access to the global workspace with a winner-take-all competition. Access to this global workspace equals access to consciousness and limitations in the operation and access to this global workspace might explain the referenced psychological phenomena. For the brains in the study of Soon et al. (2008) the global workspace hypothesis would thus imply that the neuronal processes, having found the which-finger-to-move decision (fig. 2 (a), x_1), needed about six to nine seconds to access this workspace and "produce" the awareness in the subject (fig. 2 (a), x_2). This seems to be a comparatively long period for a brain which is asked to do nothing more than find this decision.

Moreover, the concept of a "global workspace" to understand consciousness is not commonly accepted. Ned Block (2005), for example, criticized the global workspace concept of Dehaene and colleagues and suggested to look for two NCCs: The phenomenal NCC (a minimal neuronal base for the content of an experience, in so far as the experience of "red" differs from the experience of "green", for example) and the access NCC (the neuronal base of contents of experiences with respect to working memory, reasoning, decision-making, etc). But Dehaene and colleagues argue that it is only possible to study the access NCC (thought as the global workspace) as in the other case we have no reports or data: As soon as subjects are able to report and act on an event, this event, or this neuronal process, won the winner-take-all competition and gained access to the global workspace. Hence further "experiential" *content* or phenomenal states are beyond the possible focus of scientific research. However, so far we detected the *necessity* to differentiate between a decision and its (phenomenal) awareness although Block's differentiation does not really help us in this respect: We would have to assume a kind of dependence of the access NCC (the brain's decision making, the brain's movement initiation) on the phenomenal NCC (awareness of the brain's decision).

Lamme (2006) considered another posibility. He suggests to move towards a true neural stance on consciousness and define consciousness only from the neuronal point of view with the advantage to avoid inconsistencies. In his work he lists a range of psychological and neuropsychological phenomena and summarizes inconsistencies with respect to a neuronal correlate of consciousness. The experimental phenomena observed with split brain patients, for example, would imply that the NCC "sits" solely in the left hemisphere whereas the right hemisphere would have to be termed "unconscious".

Lamme suggests as a solution to define consciousness as "recurrent processing". I.e. whenever neurons are activated not only in a feedforward manner but also via feedback loops we would state that the person belonging to these neurons can be deemed "conscious" of something. Lamme admits that in this way we might lose the "explanatory power" of a definition of consciousness but we gain the possibility to focus research on the *purpose* of consciousness: Consciousness i.e. recurrent processing is necessary for learning and synaptic plasticity, for example. In this way it is not possible to solve the "hard problem of consciousness" but, following his opinion, we seem to be a lot better off than traditional NCC approaches. Furthermore it is possible to separate the aspect of consciousness from other aspects as attention, reportability or working memory. Thus we could state that someone is conscious of something but it is not necessary that he is able to report on this state.

Unfortunately Lamme's suggestion does not help us with respect to our decisive question: If the brains in Soon's study were "conscious" by definition about the brains' decision already six to nine seconds before they were able to report and act on this decision what happened in between? Why did the brains somehow "wait" for this reportability before initiating the corresponding movement?

In summary the considerations regarding the neuronal correlate of consciousness do not take us any further; there are a range of models and potential explanations but so far none is able to cover all relevant phenomena and none is at least able to explain the delay observed in Soon's study: Neurons operate in the scale of milliseconds but we have at least six seconds from the decision to awareness and action.

Furthermore it seems to be at least unclear how the attempts to model decision making processes could fit into the so far named frameworks: Are the processes of the integration of sensory data *conscious* or only their results or both? Or is the awareness of the brain's decision a phenomenal state beyond scientific treatability? Or are phenomenal state (subject level of a decision) and neuronal state (brain level of a decision) identical?

At this point it might seem inevitable to enter the more philosophical debates on consciousness and consider - to begin with - a fundamental criticism Daniel Dennett expressed almost two decades ago: Are we failing to understand and explain consciousness, are we failing to understand the brain because we are still bound to the "Cartesian devil" and try to understand the brain with the support of Descartes' concept of Mind?

Following Dennett's consideration we have to discard all ideas of a Cartesian self or a kind of "controller" of the brain if we ever want to understand the brain and its operations.[13]

Actually his criticism hits all models or neuronal explanations listed so far, as already the simple question what is "in" our "out" of consciousness has to be deemed misdirected by, in Dennett's terms, a "Cartesian materialism". Following Dennett, a "Cartesian materialism" is the resulting view following the discovery that the Cartesian dualism has to be discarded but the concept of a Cartesian Mind still blurs the understanding of the *brain*.

However, the concept of a global (neuronal) workspace found recently some support by Dennett himself though criticizing single items and asking for caution with respect to terms as "top-down" and, of course: global availability or the global workspace does not *cause* consciousness "but is in itself a conscious state".[14] Nevertheless, the concept of a global workspace did not help us with respect to our decisive question. Which brings us back to our question regarding Soon's delay between decision, movement and (phenomenal) awareness. Hence we are to ask:

What is a phenomenal state? What is its purpose? Does it really exist?

---

[13] Dennett (1991, p.106-113)
[14] Dennett (2001)

7 Tracing phenomenal states

In this respect we might take again a look to our conscious machines and the review of Gamez (2008) as even Dennett himself aimed to construct a (phenomenal) conscious robot.[15] With the area labeled MC4 we meet the question of phenomenal states as well. This research area is named "synthetic phenomenology" and claims to have its roots in the phenomenological considerations of Edmund Husserl (1977). Whereas Husserl's phenomenology dealt with the description of human consciousness, "synthetic phenomenological" deals with the description of machine consciousness. Following Gamez, synthetic phenomenology refers to the identification and description of machines that are capable of "conscious states".[16] Hence we can expect to find some explanations regarding existence and purpose of phenomenal states within "synthetic phenomenology".

However, Gamez introduces the major difficulties to first of all identify machines that have this capability and further on describe in which respect. For the identification of machines capable of "conscious states" or "synthetic phenomenology" there are several suggestions. Whilst some argue that the machine must have a representation of the world with *itself* in it and claim to have developed successfully such a system others argue that it will never be possible to decide whether a system is phenomenologically conscious as we simply do not know the factors needed for phenomenal consciousness.

Thus a commonly accepted definition of a phenomenal conscious machine is not available. Following the "representational" account the global workspace model is deemed entirely unconscious, for example. A view the defenders of this approach would certainly not accept. Nevertheless, Gamez summarizes the second problem following the identifications of systems with such a capability to discriminate between internal states that are phenomenal states whilst other are not.[17] In this respect he outlines an overlap to another research direction labeled "neurophenomenology" which might be interpreted as "human phenomenology from a third-person-perspective". In this research area patterns of brain activity measured by fMRI or similar techniques are

[15]Although this project meanwhile effectively stopped. Gamez (2008, p. 897).
[16]Gamez (2008, p. 904)
[17]Gamez (2008, p.904/905)

related to the phenomenal state of the subject. Gamez expects a "considerable potential" of collaborations between synthetic phenomenology and neurophenomenlogy.[18]

However, there are many neurons within the brain, how to identify the ones responsible for a phenomenal state? And what would be their special property responsible for the generation of this state? After all the problems we face with respect to human phenomenology might be almost identical to those faced by synthetic phenomenology: Why are some internal states conscious whereas others are not and how to differentiate them? What is the special sign of a conscious internal state, being a machine or a brain? And what is its purpose? As Gamez also refers to the ethical consequences of constructing machines "with real feelings"[19] we might at least assume that a phenomenal state could be achieved if someone or something has "real feelings".

Again, these considerations do not seem to take us any further: The root of all our problems regarding consciousness seem to be the difficulty to detect this consciousness, phenomenal state or "real feeling" from outside, from a scientific "third-person-perspective". And in this respect machine consciousness stumbles with exactly the same difficulties and controversies as human consciousness and has at least so far no hints *why* phenomenal states should exist at all.

Hence Gamez addresses as well the principal criticism raised with respect to the project of "machine consciousness": Dreyfus' considerations that human intelligence is not reducible to a number of rules;[20] the objections of Roger Penrose essentially with respect to phenomenal awareness[21] and Searle's famous chinese room.[22] However, Gamez considers these criticisms principally as non-essential and argues that at least the human brain is itself a biological machine.[23] But what *are* "phenomenal states", what is their purpose?

Why the awareness of a decision in the subjects? And if is not necessary to become conscious of a brain's decision why the large delay

[18] Gamez (2008), p. 905
[19] Gamez (2008, p. 890)
[20] Dreyfus (1992)
[21] Which is, following Penrose closely linked to the question of "understanding": "I do not see how to say much that is scientifically useful about the qualities of a 'free will' or 'awareness', but 'understanding' seems to me to be the one of the more tangible aspects of consciousness." Penrose. (1996)
[22] See for example Searle (1992)
[23] Gamez (2008, p.892/893)

between decision and action for the biological machines participating in the study Soon et al. (2008)?

We seem to move in circles and might be a lot better off if we start to question the results obtained by Soon et al. (2008): Possibly the brain areas of activity "shaping the up-coming decision" were just "shaping" and there was a lot more work to do to find the (brain level) decision. Following the completion (of the brain level) of the "which-finger-to-move" decision the brain initiated the movement and the awareness of this decision in the subject in a comparable time scale as in the other examples considered (child; Goldberg et al., 2008; Wang, 2008). However, even if we state that there was probably more work to do for the neurons to find the decision we might still ask: What exactly? And why were other brains able to act much more quickly (child; Goldberg et al., 2008; Wang, 2008)?

And what about the question whether the ability to discriminate between stimuli is identical to a *decision*? Is the "integration of sensory data" a consistent definition of a *decision*? What are phenomenal states? What is their relationship to neuronal states? Are we hunting for a phantom, still blurred by the "Cartesian devil"?

Ironically we might find a conclusive hint in Husserl (1977). In Husserl's "Cartesian meditations" we do not only find a reference to Descartes' "ego cogito", the item heavily criticized by Dennett, but moreover a "transcendental subjectivity". For Husserl then the method of "epoché", or "bracketing", a method obviously now applied to the field of synthetic phenomenology, implicated an "eidetic reduction" which leads directly to the cognition of the transcendental self, to the cognition how the world is *constituted* by consciousness.

But what could be meant by such a "transcendental self"? Does it represent a kind of presupposition for the constitution of consciousness itself? Can synthetic phenomenology thus ever lead to some success? And how to find such a "transcendental self" amongst the neurons?

However, without the cognition of this root appearing as a result of the "eidetic reduction" Husserl deemed science unable to produce *true* knowledge, but finds itself lost in contradictory experimental results, fundamental problems and "incomprehensibilities",[24] which

[24] "Nicht anders in den positiven Wissenschaften. Sie sind Naivitäten höherer Stufe, Werkgebilde einer klugen theoretischen Technik, ohne daß die intentionalen Leistungen, aus denen alles letztlich entspringt ausgelegt worden wären. (...) Daher

might describe exactly the situation we currently find ourselves in.

Thus let us relax and see how to sort the stuff concerning decisions, brain and consciousness research and let us start with simple but fundamental observations.

8 Observations

Phenomenal states seem to be something that neurons *themselves* are not capable of; neurons are thought to operate according (at least) physical laws and we lack physical laws that incorporate feelings. Hence, neurons are not sad or happy, they do not wait, attend or get bored, they spike, as far as we are able to *observe*. However, there seems at least to be a kind of relationship between neuronal activities and these rather mystic phenomenal states. To differentiate them let's, for the beginning, thus simply *distinguish* between "neuronal states" and "phenomenal states".

Daniel Dennett would probably disagree and point to the "Cartesian devil": Neurons do not "produce" consciousness and by no means they "produce" phenomenal states as something different from or additional to a neuronal state but they *are* this state. However, for the beginning we are just talking about potentially empty terms we got used to use blurred by the Cartesian devil. Hence phenomenal states and neuronal states might be the same but we detected severe difficulties with their relationship while seeking for a decision within the brain. Therefore let's begin with this potentially empty differentiation and try to understand their relationship or the non-existence of phenomenal states.

Following Gamez' considerations machines were occasionally thought to be capable of phenomenal states if they have a certain representation of the world with themselves in it. Thus when asking for phenomenal states we would actually ask for the relationship between the world, the neurons and our representation of the world.

However, here we face the risk to enter a circle: If we consider phenomenal states to be or constitute a representation of the world we are

haben wir gerade auf der hohen Stufe der modernen Wissenschaften, Grundlagenprobleme, Paradoxien, Unverständlichkeiten. Die Urbegriffe, die, durch die ganze Wissenschaft hindurchgehend, den Sinn ihrer Gegenstandssphäre und ihrer Theorien bestimmen, sind naiv entsprungen, sie haben unbestimmte intentionale Horizonte, sie sind Gebilde unbekannter, nur in roher Naivität geübter intentionaler Leistung." Husserl (1977, p.157) to claim the identity of the world and our representation of it because these representations *include* our observations about neurons as a kind of premise. On the other hand we are to claim the identity of neuronal and phenomenal states because a neuronal state actually *is* a phenomenal state if we follow Dennett. Thus we would have to assert that the outside world actually *is* a neuronal state.

To clarify this circle we are to outline what we mean by "the world" and how neurons might be incorporated in processing this world or some of its aspects. Thus we have to clarify what we mean by an "outside world", by "neuronal states" and how both might relate to our representations or perceptions.

In this respect a computer seems to be a more subsumable component than a brain. Facing similar difficulties both, computers and brains, are considered as almost equal hot candidates to possess or generate "phenomenal states" or "consciousness". But why is this the case? Why is a computational device considered as an almost equal hot candidate for phenomenal states as a brain?

Potentially the primary reason for this comparability is constituted by the aspect that both of them seem to process signals from an outside world, both seem to "compute"; a kind of processing or computing which might enable the capability to have internal states, may they be conscious or not.

However, every computational device, every machine, has a user and further devices to access its "inner states" from outside. Computers were constructed by humans but brains developed obviously by evolution without an explicit constructor. If we consider a computational device to be an equivalent applicant for consciousness as a brain we certainly think about a computer without an user; hence we think about an autonomous actor who does not require any (user) access devices as a screen, for example. Keeping these restrictions and thus especially the inventions of Searle (1992) in mind let us therefore first of all just use a computer and simply translate our *observations* about neuronal states and the world to a very simply computer metaphor, let us simply ask for comparable relations. Hence we do not ask for a method to measure, detect or generate consciousness within a computational device but we ask how to use a computer metaphor to understand, first of all, the relation between the world, the neurons and ourselves on our way to a potential answer to our decisive question. To form this basic question more precisely:

In which way can we find ourselves (brains i.e. neuronal and potential phenomenal states) represented within a computer analogy?

To address this question we need to start the computer analogy from the scratch.[25]

9 Computational steps from neuronal to phenomenal states

First of all (Searle, 1992, see above) we are to consider computers as autonomous actors in *their* environment as brains are considered as autonomous agents as well. Furthermore we should, for the beginning, avoid to apply knowledge about computers and their operation to the brain and its operation. Hence we should not try to find a processor, hard disk, etc within the brain.

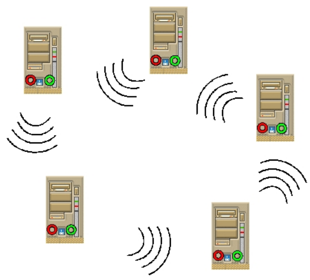

Abbildung 3: Autonomous actors. When considering a computer to represent the brain, we are to consider computers as autonomous actors in their environment as we consider brains to be autonomous actors.

Considering these initial restrictions the primary comparability between computers and brains seems to be constituted by the *codes* with which they operate: Computer operate with binary codes where brains operate with neuronal codes. We might not know much about the concrete decoding of these codes, but our observations tell us, that neuronal respectively binary codes are used within the brain respectively within the computer and *differ* from their environment.

One aspect of the primary environment of computers is of course constituted by other computers with which they communicate (see fig.

[25]See f.e. Stemme (2009) for an overview of the usage of computer metaphors in the mind/brain debate.

3). Within this communication they might debate as well about rather philosophical issues provided that the corresponding questions arise *within* their communication and environmental access.

For the purpose of the environmental access we might consider them connected to something like an outside world, "world" in the sense of different from neurons, different form transistors, via various "interfaces". One of these interfaces might be a kind of simple scanner transforming pictures or light waves into the Zeros and Ones with which the computer operates, another interface might be a microphone, transforming sound waves into the binary code used within the computer.

In a comparable manner we find eyes and ears connected to a brain. Hence, the comparability of the interface components for brains and computers results from the aspect that both transform environmental signals into a neuronal respectively binary code which *differs* from this environment (fig. 4).

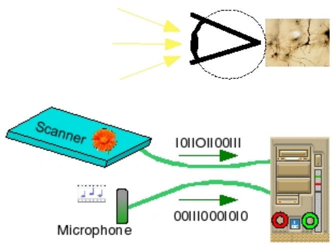

Abbildung 4: Computer Interfaces. In a similar manner as the brain is connected to an environment via eyes (top) and ears, for example, a computer might have environmental access: Peripheral devices transform something that is different from transistors (e.g. light, sound) into the binary code with which the computer operates.

What about phenomenal states?

At least we have the impression that we do not experience neuronal activities but rather a world outside of us and somehow inner feelings. These aspects are only accessible via a first-person-perspective and thus not observable from a third-person point of view. These considerations and our physical laws tell us, that feelings are not experienced by the neurons themselves. Similarly, we do not assume that neurons themselves have perceptions. So where to find these phenomenal states within the computer?

Our computer operates on binary codes only. Whatever he might process or represent with "higher"[26] binary states these will still be binary states, so far incapable of any *qualitative* differentiation between sound and vision, for example. Whereas it turns out to be comparatively easy to state that one code III000 represents "red" and another code 000III might represent "green", so that the computers *within* their communication might be able to differentiate red and green and communicate about these issues, with respect to sound and vision we fail: It is possible that the computer states the identity of a binary code received via a scanner and a binary code received via the microphone because coincidentally both codes were identical.

Although I consider it very difficult to describe the phenomenological qualities of sound opposed to vision, I am quite sure that I would never state that what I see is actually *identical* to what hear except that I receive one impression via the eyes and the other via the ears. Hence some kind of qualitative, or dimensional "jump" is obviously missing so far although this "jump" appears to be related to binary or neuronal states, respectively.

With this qualitative aspect we actually face phenomenal states and are now able to raise a key question: Would our computers so far and without an external metaphysical programmer be able to *begin* debates about the *existence* of phenomenal states? Would the computers within and out of themselves be able to debate about what actually is "outside"? What would be *their* concept of an outside world? A binary world constituted out of Zeros and Ones?

If, just if, our computers so far, just based on a kind of "natural" development, are really able to prefer or dislike something, for example, they are only able to debate about something like three Zeros being much nicer than three Ones. *All* our computer might get to know is constituted out of Zeroes and Ones. Provided that we avoid a metaphysical programmer but just assume some kind of evolutionary development based on environmental interactions.

Hence in order to enable our computer to enter into philosophical debates in a similar manner as *we* debate about phenomenal aspects we need to assume that somehow additional phenomenal states must have developed evolutionary; to just illustrate these states let us investigate

[26]See f.e. Tononi (2004) for a "higher-order concept" in terms of information integration.

the usability of the screen.

A screen represents as well a kind of "qualitative jump" compared to the binary states used within the computer. Presentations on the screen are different from but dependent on the operational activity of the computer although these presentations do not influence the operational activity of the computer as an autonomous actor. Hence in decisive aspects a screen of an autonomous actor computer is comparable to the phenomenal states of a brain: as different, as dependent and as superfluous for its actual operation, at least following our *observations* so far.

Our phenomenal states, our experience of colors, sound, pain, anger, happiness, etc. are - as far as we can observe - not *experienced* by the neurons themselves and are, again following our observations, of no causal relevance for further neuronal activities. Also, between us and the outside world there are a couple of neurons. We do not have access to some outside world directly, no access to the picture on the retina. Without our visual cortex we would not see anything. Hence there are always some further neuronal operations required to actually *see* something.

To illustrate the aspect that *we see* a flower, for example, that *we* experience some feelings, which are neither experienced by the neurons themselves nor necessary for their actual operation (by premise of a physical monism) we might thus use the screen of our computer (fig. 5).

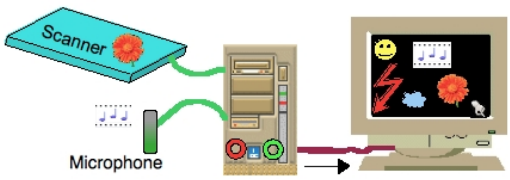

Abbildung 5: "Phenomenal states" in a computer analogy. We use the computer screen to illustrate phenomenal states: Phenomenal states differ from neuronal states in a comparable manner as presentations on the screen differ from binary states ("qualitative jump"). Furthermore, a computer, thought as an autonomous actor, does not need a screen for its autonomous action and similarly we consider phenomenal states to be irrelevant for neuronal operations.

Thus we use the computer screen to illustrate our experiences which are in either theory of mind - based on a physical monism - of course of no causal relevance. As is the screen for the activities within the computer. The computer, considered as an autonomous actor, is determined by the binary activities within its cabinet as we consider us to be determined by the neuronal activities within our brains. For either activities, binary or neuronal, phenomenal states are unnecessary though (still only potentially) existing.

The screen of a computer just constitutes a device representing essential properties of phenomenal states and supports us in verifying their relationship to neuronal states and the world. For the purpose of this empirical validation let us now investigate a famous example: The "color phi" phenomenon - before we finally come back to our decisive question.

We consider this example on the one hand because it was one of the examples Dennett (1991) and more recently Blackmore (2003) used to illustrate that our common concept of consciousness as a continuous stream of perceptions is misleading, that we have to deconstruct our ideas about self and consciousness. On the other hand optical illusions constitute principally a very striking example to understand the relationship between world, neurons and our representations of this world or perceptions, respectively.

10 The phenomenal "Color Phi" phenomenon

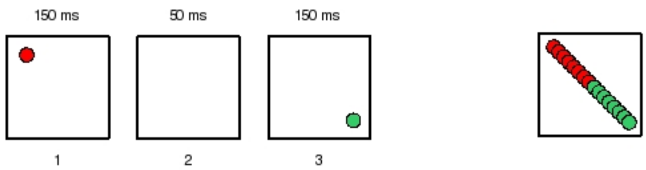

Abbildung 6: The color phi phenomenon. Left: Two spots are presented to the observers. Right: The timing and spatial conditions of the presentation lead to the perception of a single moving dot.

In the "color phi" experiment[27] (fig. 6) the participants are seated in front of a screen on which two spots in different colors at different

[27] Initially conducted by Kolers and von Grünau (1976)

positions are presented. Due to the spatial and timing conditions of the presented spots, the observers reported that they saw a moving spot changing its color in the middle of the way.

The key question raised in conjunction with this experiment turned out to be: How can the observer *know* the color of the second spot already in the middle of the way i.e. *before* the presentation of the second dot? Dennett argued that this is neither possible nor necessary because there is no central observer within the brain, there is no place where everything has to come together. We rather have to consider the perception of a single moving spot as kind of "story" told by the brain, as an afterward explanation provided by the brain, *following* the presentation of the two spots rather than a true perception.[28]

Hence "Color Phi" appears to be a good example to verify our assumptions regarding phenomenal states, in this case the potential (phenomenal) perception of one moving spot opposed to the presentation of two static spots "outside".

When moving the focus from the *spatial* dynamics of the perceived spot to the *timing* dynamics of the presented spots neuronal explanation for this phenomenon turn out to be rather easy: for almost half of the presentation *time* a red spot was presented, for the other half a green one. Thus the observer does not need to know the color of the second spot in the middle of the *way* but he certainly might know the color of the second spot after roughly half of the presentation *time*. The remaining aspect of the experiment, namely the aspect that the spot is moving in perception but stationary outside is a common phenomenon observed daily when watching TV and thus obviously lacked any mysterious aspect. Considering neuronal explanations for the color phi phenomenon these might even meet neuronal dynamics following the presentation of a *really* moving spot (fig. 7).

Hence when moving the focus from the *way* the perceived spot put aside to the presentation *times* of the outside spots, the *perception* of one moving spot in this experiment becomes plausible and does not undermine a concept of consciousness or phenomenal state per se: We see this moving spot in this experiment as we see a *really* - outside - moving spot. A phenomenon which is actually very obvious for everybody when watching TV: Only a sequence of static images

[28] Dennett (1991, p.136)

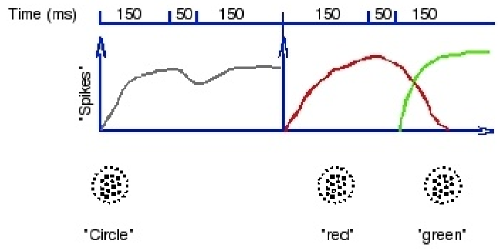

Abbildung 7: Color phi in a neuronal perspective. A pool of neurons (labeled "Circle") responsible for the perception of a spot might stay activated (Y-axis, "Spikes") due to the spatial and timing dynamics of the presentation ("movement detectors", cortical area MT). Furthermore, for 150 ms, almost half of the presentation *time* a red spot was presented followed by the presentation of the green spot for another 150 ms. See as well Stemme and Deco (2008).

is presented on the TV screen but we *see* people moving around as we might see people moving outside, in the "real" world, for example. Thus these phenomena, perceptual illusions, underline the necessity to differentiate between the outside world, between our *environment* and our *perceptions*. And modern brain research taught us that the brain builds a kind of connection, a kind of bridge between our environment and our perceptions.

Thus we are to separate separate optical illusions and with them perceptions and our phenomenal states from the outside world and our brains and are able to use the computer analogy to understand these relations (fig. 8).

Furthermore, with the support of the computer analogy we are able to balance two extremes, to hold a kind of paradox between our observations, which at least constitute our physical laws, and the circumstance *that* we observe. We are able to conceive the potential correlation between the outside world (the world on the scanner) and our subjective experience (the screen). We are able to acknowledge that whatever we perceive appears to be dependent on neurons, to be *mediated* by neurons.

Obviously, we have no access to the "things in itself"[29], as Kant might have addressed this circumstance, no access to the world on the

[29]Immanuel Kant, Kritik reiner Vernunft, 1781/1787

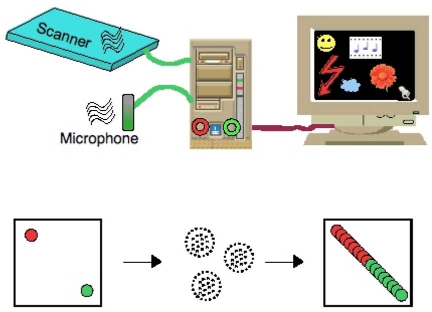

Abbildung 8: "Color Phi" in a computer analogy. Two stationary spots outside lead to some neuronal processing and thus the *perception* of one moving spot. Hence our perceptions are obviously *generated* by the neurons.

scanner or beyond the microphone if we express it with the computer analogy. But with the support of this analogy we might be able to understand relations and are, furthermore, able to acknowledge the most astonishing aspect about this world: *that* we perceive it, that we see this world, hear, touch and feel it. An aspect which is - following our observations - neither present nor required for the neuronal operation of brains.

At this point it might be possible to consider the principal criticism raisable with respect to a physical monism. The physicist Jan-Marcus Schwindt (2008), for example, pointed to the substantial inconsistency of a physical monism: Everything we are talking about, everything we try to describe with scientific theories relies on symbols which are *relative* to their interpreter. Every symbol, every structure or every scientific theory is a theory with respect to an interpreter which is *not* part of this structure.

Following the philosophical considerations of Erwin Schroedinger Schwindt analyzes the curious double role of consciousness: On the one hand constituting the world (see as well Husserls view, mentioned above) on the other hand thought to be constituted by the brain. Schwindt comes to the conclusion that neither resulting extreme, neither a physical monism nor a "mental monism", helps to understand consciousness but that many "puzzle pieces" are spread over many different disciplines and are to be critically reviewed with "the right mixture of skepticism and openness" in order to gain an effort in understanding consciousness or the world.[30]

Especially with respect to brain research such an effort seems extremely necessary as based on the wide spread *believe* in a physical monism researchers seem to be *forced* to look for almost everything within the brain. And this aspect, the blind believe in a physical monism without any observers, without ourselves, might have had a substantial contribution to the irritations we discovered amongst the studies with respect to the question of "Decisions".

11 A more complete computer analogy

Before completing our observations regarding computers and brains let us start with a summary of previous considerations:

When we consider our brain as kind of computer, as a kind of computational or (neuronal) coding device, then we are to consider our retina as an interface component, as a kind of scanner in the analogy, and we are to consider the screen of the computer as a device whose contents represent our phenomenal states.

Given the necessity to distinguish between outside world and perceptions, and between neuronal and phenomenal states, we are able to ask what exactly we might find on the screen: What is the content of a phenomenal state and how to differentiate between conscious and unconscious events, for example. Phenomenal states include not only a visual or auditory scene, but as well the body representation, feelings, thoughts, etc., hence all this "stuff" which we cannot find within the brain. How are these items arranged on the screen? Which contents are in the center of the screen? And are items more at the border of the screen to be considered as kind of "preconscious"?[31] What exactly appears on the screen?

Switching back to some knowledge we have concerning real computers, we may, first of all, consider the searched neuronal correlate of consciousness as a kind of "device driver" for the screen. A device driver is necessary for the control of screen presentations and is actually a kind of reverse implementation of the device driver for a scanner. Hence, with respect to brains we would have to look for kind of "reverse" components compared to retinal cells, for example. Are there neurons within

[30] Schwindt (2008, p.26)
[31] Dehaene et al. (2006)

the brain which receive a lot of input but do not appear to spike nor drive muscles? *Which* neuronal operations lead to conscious percepts? And much more importantly *why* does this happen?

With real computers used in our everyday work we usually have presentations on the screen because some kind of user interaction is necessary. Furthermore the contents presented on the screen appear in a format understandable to the user: We do not have to operate on Zeros and Ones directly.[32]

Using this picture we might understand the relationship between neuronal and phenomenal states as a kind of *translation* provided by the neurons. Obviously neurons do more than absolutely necessary to operate within the world at least under the premise of a physical monism: They translate the stuff "out there" into perceptions, sound, smell, taste, etc. And they generate some kind of feelings. The screen seems to go somehow beyond the stuff necessary for the binary operation of the computer if all actions are determined by the binary states within the computer. For an autonomous actor computer a screen is as unnecessary as for a brain a phenomenal state.

Actually we are so far very well in line with modern theories of mind: If we consider - with identitytheorists - phenomenal states to be actually identical to neuronal states the screen would have to be considered as the result of the (neuronal) first-person-perspective with the open issue whether all activities within the computer lead to presentations on the screen: (Why) are there neuronal states which do not lead to phenomenal states?

With Searle, who considered considered consciousness to be the result of a biological process,[33] we might consider the presentations on the screen as the result of a special kind of process within the computer. With Chalmers, who considered consciousness to be situated at the level of particles, the presentations on the screen would be the result of potentially more complex particle arrangements. Hence so far only eliminative materialists or Daniel Dennett would probably (still?) disagree and state that phenomenal states do not exist, *are* actually neuronal states and our computer analogy would thus just illustrate our misleading picture of the world.

[32] A picture use in a similar manner as well by Schwindt (2008)
[33] Searle (1992)

But so far we discovered the brain to be an amazing component: Following our observations and intending to keep at least a roughly consistent picture of the world we are to recognize the brain as a kind of interface component to the world, to whatever there might be outside. We are able to understand our dependence on neurons and are able to ask old questions with a new background: What actually *is* an object, what happens that two objects outside merge to one perceived objects? How are object properties bound by the brain and what might be the role of the "device driver" in all these cases? What actually *constitutes* an object and what does this tell us with respect to the world on the scanner? And *why* is there something like a "screen" at all?

Hence, so far a central item remains to be addressed: The answer to our decisive question and - with respect to the computer metaphor - the question of a self (fig. 9): Has the self to be considered as an element on the screen, "something" generated by the binary states within the computer, by a separate process? Or has the self to be considered as the sum of all experiences, the screen as a whole? And *where* to find a decision?

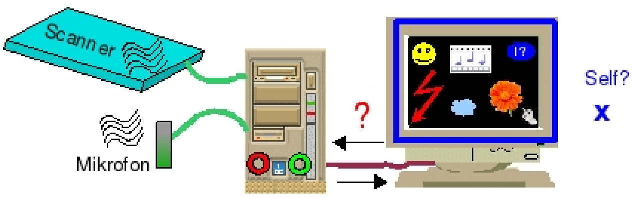

Abbildung 9: The Complete Computer Analogy. Given the screen for the illustration of the relationship between phenomenal states, brains and the outside world, we are able to think about further relationships.

A major advantage of the, let's call it "complete computer analogy", as it considers not only an equivalence of neuronal and binary states but illustrates as well phenomenal states and their relationship to the brain and the world, is hidden in the possibility to ask as well for the role of a kind of "computer user".

Hence we face the ability to ask for the existence and role of such an user, compare this possibility with other options and look for correspon- ding empirical findings substantiating either view.

We are able to think about this computer user as being someone or something who experiences the outside world, someone who *has* perceptions and who experiences as well some feelings potentially generated solely by the brain. Hence, is it possible that the self corresponds to "something" in front of the screen? An unknown "**x**" that potentially even hits Husserl's "transcendental self"?

Considering a range of neurological diseases (Anosognosia, Neglect, etc.) we are at least not able to exclude this last option but find ourselves enabled to understand why these phenomena might disappear without any neurological changes observable: We do not have access to the outside world directly; it is not possible to have a quick look to the "things in itself", to the picture on the scanner, to understand that a leg is no longer represented within my bodypicture but nevertheless belongs to my body. We depend on the neuronal translations which might have led in these cases to an inconsistent picture of the world with my body in it.[34]

Another consequence is the ability to reinvestigate the old question regarding the source of attention:[35] Might *attention* indeed be a *cause* of neuronal activities rather than just an *effect*, a kind of epiphenomenon of neuronal processes?[36] Hence, is it possible that *we* are able to actively direct our attention to certain issues amongst our experience, amongst our phenomenal states, to trees, apples, music or some pain, for example? Is it possible that as soon as we direct our attention or our thoughts to our left hand or a potential movement of our left hand, for example, is it possible that this attention *leads* to an increased neuronal activity?

Is it possible that an increased neuronal activity *following* such changes in the focus of attention is measurable in terms of the famous readiness potential, for example?[37] In a similar manner as widely ob-

[34] See as well Stemme (2009)

[35] A question already addressed by William James (1890) and Wilder Penfield (1975), for example.

[36] See as well Blackmore (2003) who neglected the possibility that attention might be a cause.

[37] As measured by Libet (1985). This interpretation is actually supported by the work of Trevena and Miller (2010): They measured a readiness potential in conjunction with decisions to move or *not* to move, hence as a consequence of the bare *concentration* on a movement.

served?[38] Or in terms of activated cortical regions identified by Soon et al. (2008)?

Hence, is it possible that these measured activities simply reflect the circumstance of directing the *attention* to either finger and that the final decision is something far beyond these activities? Is it possible that a *decision* is something we *cannot* find within the brain? Is it possible that we are to assume a kind of "brain user" to be responsible to take decisions? Hence a "computer user" in the computer analogy, directing the *attention* to various items on the screen. Thus is it possible that this assumption would be able to solve all our perplexities regarding the studies of Soon et al. (2008), Goldberg et al. (2008), Wang (2008) and of course with respect to our child?

In this case we would have to consider the screen in the analogy as a kind of touchscreen and the connection between the screen and the computer to be bidirectional (fig. 9). Thus in this last case we would enter a domain which previously had to be considered to contradict a scientific view of the world: A *kind* of "dualism"; but a kind of dualism as reflected in our observations if we do *not* neglect the circumstance *that* we observe.

And with the support of the computer analogy we should even be able to think beyond the idea that "dualism" requires the concept of a funny ghost moving particles around the brain.[39]

13 Decisions

Actually we finally (re)opened a common but nevertheless new perspective in the debate regarding the relationship between brain and mind. Based on the consideration that any "non-material" or "mind stuff"[40] can never be able to move particles around the world we initially considered ourselves *forced* to believe in a physical monism as the premise for all plausible theories of mind.

However, as outlined in this essay, this premise leads to different but nevertheless tough kind of problems even when considering simple empirical findings with respect to decisions: *What* is a neuronal decision?

[38] E.g. Egner and Hirsch (2005) detected increased neuronal activity depending on the attention of the subjects in a fMRI study.

[39] Dennett (1991, p.35): Caspar, the funny ghost as an illustration of our misconceived concept of "mind".

[40] Dennett (1991, p.35)

What is the awareness of a decisions? *Why* do we have phenomenal states? Do we really have them at all? How could phenomenal states be related to neuronal states?

With the support of the computer analogy we were able to generate some respective ideas which obviously led us directly back to a kind of "dualism".

Did we make a mistake?

Where is the world, where are our neurons, what is the *relationship* between the world and our perception of the world - *where* is the moving spot in the color phi example? We *see* this spot as we *see* a really moving spot. Whatever we see seems to be dependent on our brain. Our brains appear to be responsible for our perceptions, seem to actually constitute the way we perceive this world. But neurons do not see trees.

We do. And we are unable to ignore this item because it simply constitutes the world we are talking about.

If we aim to understand the brain we are not able to ignore the aspect that it appears to be just the brain which *constitutes* the way we perceive this world, that it appears to be the brain which *constitutes* our observations. Thus especially with respect to brain research we need to acknowledge the contents of our observations *as well* as the aspect *that* we observe. The suggested computer metaphor just enables us to see these relations. Obviously we are to face a *kind* of dualism back again. But maybe our computer analogy might take us a step further in understanding.

Previous interactional theories focussed in the first instance on the concrete mechanism of a potential interaction; they searched for a certain *location* (e.g. pineal gland, microtubuli) of or physical mechanism (e.g. quantum mechanics) for this interaction. However, considering the circumstance that *if* we think dualism as a dualism of two very different *substances*, mind and matter for example, should we consequently not look for an interface between them somehow in the *middle* of the way?

Considering phenomenal states in their relationship to the brain and the world with the help of the computer metaphor enables us to consider them as a kind of *translation* provided by the neurons, potentially because the *computer user* only understands this language: colors, sounds, smell, taste etc., rather than Zeros and Ones, or lightwaves and particles, or whatever there might be outside. These phenomenal states themselves go somehow beyond the physical stuff and its natural laws following the *contents* of our observations: Neurons do not need these concepts, they seem to operate and determine our behavior on a spiking level from sensory to motor neurons without any perception.

Thus, the brain does something more than absolutely necessary under the premise of a physical monism, something that goes beyond itself. Hence these phenomenal states themselves acknowledge the aspect *that* we observe and thus form a candidate for one part of a kind of "interface" between "substances".

And *attention* is at least a promising candidate for an interaction in the "opposite" direction.

If we are forced to look for *neuronal* processes somehow distributing "attention" within the brain we face in any thinkable implementation the problem that neurons themselves do not know anything about a *content* of this processing. [41] They do not know that it might be much nicer to listen to the sound of some music than concentrating on some pain in the foot, for example. But to ignore a current pain we use exactly this method - we try to focus on something *else*. Why should any neuronal process do this? From an evolutionary point of view it makes sense to focus on a pain in the foot until the nail is removed. Which kind of knowledge would neuronal processes need to initiate the change in the focus of attention from some pain to some music?

Further more, a range of phenomena allow far more plausible *neuronal* explanations if we do not have to look for the source of attention within the brain: E.g. the examples Daniel Wegner (2002) outlined concerning the "illusory" will. He assumed rather complex "ironical" processes to be initiated within the brain; processes that find far more easy *neuronal* explanations if we consider that we can never direct attention to *nothing* but only to something *else*.[42]

With the question regarding the source of attention and the inability to assume any kind of *brain* mechanism to distribute attention amongst the neurons we are able to acknowledge Dennett's objection with respect to the missing central controller in the brain.[43]

But moreover, the computer analogy as presented so far allows to meet Penrose's criticism with respect to the question of understanding

[41] Consider Searle's objections, Searle (1992)

[42] See as well Stemme (2009)

[43] Dennett (1991)

and Searle's famous Chinese room with the underlying question of the missing meanings on a neuronal level: Neurons do not attend, they do not know. They spike.

Any other assumption forces us to question our scientific view of the world. And with the support of the computer analogy we are even able to respond to Dennett's fundamental criticism: We do not have to try to understand the brain with the concept of a Cartesian Mind. We are able to leave brain and mind as we observe and experience and are now able to ask for relations.

But if we want to understand the brain and are to look for unreasonable components *within* the brain we might fail to understand anything as outlined in the first sections of this essay; neuronal models seem to be a powerful tool but we have to know their limits. We are able to simulate the ability to *discriminate* between stimuli, we might even be able to simulate various other psychological phenomena. After all this might help us with understanding the brain, provided that we know what consciousness actually means, what phenomenal states are and what the role of the "computer user" might be in this picture.

Hence *we* are to **decide**: Do we aim to achieve a progress in brain research and the debates on consciousness or do we prefer to continue to move in circles? Do we prefer to continue with the exchange of well-known arguments on all sides:

- "Nobody at home in the brain"[44]
- "*Meanings* are assigned from *outside* the system"[45]
- "(no) first-person-perspective, (no) quale, (no) zombie, (no) explanatory gap"[46]

which at least, and again and again, simply express the impossibility to understand the aspect that we experience this world, that we talk about this world, that we form theories like "physicalism" about this world *within* this physical theory. [47]

[44] A central argument in Dennett's argumentation: The brain has no central commander, we are like complex computer programs, constituted by simple Zeros and Ones, Dennett (1991)

[45] Searle's central objection against Dennett's view, Searle (1992)

[46] Initial arguments: Place (1956); Feigl (1958); Smart (1959); Nagel (1974); Levine (1983); Churchland (1986); Nagel (1986); Dennett (1991); Searle (1992); Chalmers (1996); Kim (1998). See Dulany (2008), for a recent overview with respect to a "science of consciousness"

[47] Compare as well the objections of Schwindt (2008) mentioned above.

Form *your* perspective and following *your* observations, **dear reader**, all beings but you *could* be philosophical zombies, lacking any phenomenal state, lacking any feeling. That all these other beings are indeed no zombies you can only know from your, your very own, personal experience: *You* see this world, hear it, smell and feel it.

But if you still prefer to consider yourself as being nothing but a brain driving a body, as a kind of computational device, with your seeing this world being nothing than "image processing", then ask yourself what a computational device like a computer, for example, is able to "see". What is the picture or representation of the world such a computational device is able to have *naturally*, i.e. without a metaphysical programmer, assigning his own meanings. How can you know anything about any *picture*? How can you know anything about "*image* processing"? So if you are not sure whether there are other "Non-Zombies" besides you, ask them to paint an apple from their imagination, if they speak your language; or ask them to sing a song. A computational device never saw an apple and never heard a song. Computers simply process, as neurons simply spike. At least as far as we are able to *observe*.

Considering the aspect *that* we experience this world provides us with an imagination of phenomenal states, of the screen in the computer analogy. An analogy which is nothing more than a metaphor providing the possibility to *see* relations and thus might provide the base with respect to essential *decisions* in brain research. But it might be a useful metaphor, permitting to build ideas about world, brains and minds beyond inconsistent one-sided theories.

Of course, it is still possible to cave occasionally into the one or the other extreme:
Physicalism: "Everything including me is at least physical." or
Solipsism: "Brains cannot do this stuff. Everything is mental. There is nothing but me."

The introduced computer metaphor provides the possibility to balance extreme positions, to acknowledge and illustrate the relationship of our observations and the fact *that* we observe. This very simple metaphor has at least the potential to support us in understanding brain, mind and world; and this understanding might even help us to construct indeed intelligent machines that *support* our daily work by considering *our* role in the game of physics.

However, it is just a metaphor, a possibility to understand relationships but no metaphysical explanation. Your might use it to find your own explanations.

But this is at least **your** decision.

Bibliography

Armstrong, D., 1968. A materialist theory on consciousness. Routledge, London.

Baars, B. J., 1988. A cognitve theory of consciousness. Cambridge University Press, Cambridge.

Blackmore, S., 2003. Consciousness - An Introduction. Hodder Arnold, London.

Block, N., 2005. Two neural correlates of consciousness. Trends in Cognitve Science 9, 46–52.

Chalmers, D., 1996. The conscious mind. In search of a fundamental theory. Oxford University Press, New York, Oxford.

Churchland, P., 1986. Towards a Unified Science of the Mind/Brain. MIT Press, Cambridge.

Dehaene, S., Changeux, J.-P., 2005. Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. PLoS Biol 3 (5), e141.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., Sergent, C., 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in Cognitve Science 10, 204–211.

Dehaene, S., Sergent, C., Changeux, J.-P., 2003. A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proc Natl Acad Sci U S A 100 (14), 8520–8525.

Dennett, D. C., 1991. Consciousness explained. Little, Brown and Company, Boston, Toronto, London.

Dennett, D. C., 2001. Are we explaining consciousness yet? Cognition 79, 221–237.

Dreyfus, H. L., 1992. What computers still can't do. MIT Press.

Dulany, D. E., 2008. How Well Are We Moving Toward a Most Productive Science of Consciousness? Journal of Consciousness Studies 15 (12), 75–98.

Egner, T., Hirsch, J., 2005. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nature Neuroscience 12, 1784–1790.

Feigl, H., 1958. The "Mental" and the "Physical". University of Minnesota Press, Minneapolis, reprint 1967.

Gamez, D., 2008. Progress in machine consciousness. Conscious Cogn 17 (3), 887–910.

Goldberg, I., Ullman, S., Malach, R., 2008. Neuronal correlates of "free will" are associated with regional specialization in the human intrinsic/default network. Conscious Cogn 17 (3), 587–601.

Haggard, P., 2008. Human volition: towards a neuroscience of will. Nature Reviews Neuroscience 9, 934–946.

Haggard, P., Eimer, M., 1999. On the relation between brain potentials and the awareness of voluntary movement. Brain Research 126, 128–133.

Husserl, E., 1977. Cartesianische Meditationen. Felix Meiner, Hamburg, (Méditationes Cartésiennes, 1931, Paris).

James, W., 1890. The principles of Psychology. Macmillan, London.

Kim, J., 1998. Mind in a physical world. An Essay on the Mind-Body-Problem and Mental Causation. MIT Press, Cambridge, 2. Auflage 1999.

Kolers, P., von Grünau, M., 1976. Shape and color in apparent motion. Vision Research 16, 329–335.

Lamme, V. A., 2006. Towards a true neural stance on consciousness. Trends in Cognitive Science 10, 494–501.

Levine, J., 1983. Materialism and Qualia: The Explanatory Gap. Pacific Philosophical Quarterly 64, 354–361.

Libet, B., 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. The Behavioral and Brain Sciences 8, 529–566.

Nagel, T., 1974. What is it like to be a bat? The Philosophical Review 83, 435–450.

Nagel, T., 1986. The View from Nowhere. Oxford University Press, Oxford, New York, Toronto.

Penfield, W., 1975. The Mystery of the Mind. Princeton University Press, Princeton, London.

Penrose., R., 1996. Beyond the doubting of a shadow. PSYCHE 2.

Place, U., 1956. Is Consciousness a Brain Process? British Journal of Psychology 47, 44–50.

Schwindt, J.-M., 2008. Mind as hardware - matter as software. Journal of Consciousness Studies 15, 5–27.

Searle, J. R., 1992. The Rediscovery of the Mind. MIT Press, Cambridge.

Smart, J., 1959. Sensations and Brain Processes. Philosophical Review 68, 141–146.

Soon, C. S., Brass, M., Heinze, H.-J., Haynes, J.-D., 2008. Unconscious determinants of free decisions in the human brain. Nat Neurosci 11 (5), 543–545.

Stemme, A., 2009. Deus et machina: Der Geist und die Naturwissenschaften. PhD thesis, Ludwig Maximilian University Munich.
URL http://edoc.ub.uni-muenchen.de/9413

Stemme, A., Deco, G., 2008. Neuronal and cortical dynamical mechanisms underlying brain functions. Elsevier Handbook of Cognitive Science, 221–242.

Tononi, G., 2004. An information integration theory of consciousness. BMC Neuroscience 5.

Trevena, J., Miller, J., 2010. Brain preparation before a voluntary action: Evidence against unconscious movement initiation. Consciousness and Cognition 19 (1), 447 – 456.

Wang, X.-J., 2008. Decision making in recurrent neuronal ciorcuits. Neuron 60, 215–234.

Wegner, D. M., 2002. The illusion of conscious will. MIT Press, Cambridge, London.

Wegner, D. M., 2003. The mind's best rick: How we experience conscious will. Trends in Cognitive Science 7, 65–69.

Wegner, D. M., 2004. Precis of the illusion of conscious will . Behavioural and Brain Sciences 27, 649–692.