



UvA-DARE (Digital Academic Repository)

Sculpting the space of actions: explaining human action by integrating intentions and mechanisms

Keestra, M.

[Link to publication](#)

Citation for published version (APA):

Keestra, M. (2014). Sculpting the space of actions: explaining human action by integrating intentions and mechanisms. Amsterdam: Institute for Logic, Language and Computation.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

4 MODEST, ALL TOO MODEST: THE SEARCH FOR NEURAL CORRELATES

Singing, we have seen earlier in chapter I.2, is impossible to define in such a way that it delineates a conceptual space in a comprehensive, consistent and coherent sense without remaining ambiguous cases or blurred boundaries between, for example, vocal signalling, infant crying and novice vocalizations. In the previous sections a somewhat more liberal approach aiming to combine several perspectives on a given function was discussed. In agreement with the earlier approach that depended upon conceptual analysis, Marr's approach emphasized the relevance of an account of the task being studied, while leaving room, however, for there being several such accounts instead of only one. Constraints on the task definition or its computational theory could subsequently be derived from the algorithmic theory, devoted to answers concerning the kind of tasks involved in singing, its particular goals, the distinction of subtasks that merit separate study, and the like. Similar relations between levels of analysis may involve the third level of analysis, focusing on the neural implementation of the function. The plausibility of a particular answer can be partly based upon its coherence with the results of investigations of the task at the other levels or other perspectives involved.

As singing is a complex and difficult task, it is convenient to use a strategy of divide and conquer. Through dividing it into several subtasks that can be investigated separately, researchers have made the problem more manageable. Indeed, as we have seen in section I.3.5, a working hypothesis behind this strategy is that tasks and subtasks are to a large extent modular in nature. Basic to this approach is our distinction of the functions involved and their subtasks, and our subsequent finding of empirical support for that distinction. Importantly, we should not overlook the possibility that our classificatory distinction leads us astray and eventually turns out to be implausible.⁶⁶ On the other hand, not making any classificatory distinctions at all may result in a situation where we have no difficulty in expanding our research topic 'signalling-crying-singing' by including ever more social communicative functions in it, like gesturing, sobbing, shivering, and so on. Expanding the computational theory of our research topic would make it ever harder to distinguish a specific neural implementation for this function other than the brain as a whole. Designing a single

⁶⁶ That classifications and taxonomies can misguide us in dividing a particular set in two or conversely lumping together two sets that had better remain apart has been a topic of discussion since Plato and Aristotle put emphasis on definition as a scientific principle. A lesson to be drawn from this is not the entire rejection of definition – which would leave us with even greater problems – but to allow for some pluralism corresponding with different questions, as is convincingly argued in (Dupré 2001).

algorithm that would allow the brain to perform this comprehensive function in all its distinct expressions would be even more difficult. In sum, some classification is necessary as a starter for research, as is some kind of modularity – even if they need to be corrected at a later point.

The tension between the investigation of a complex and comprehensive function that is hard to study as such, and the assumption of it being subserved by subtasks that do allow separate study is prominent in consciousness research – which is arguably a more prominent, larger and even less defined domain than the study of singing. For the next methodological approach, we will turn to the study of consciousness and in particular to the study of so-called Neural Correlates of Consciousness (NCC). This approach does not present a method to study the comprehensive function directly and instead suggests studying only particular components or particular phenomena related to consciousness. It allows researchers to focus on a particular phenomenon that is related to consciousness or that represents a specific form of consciousness, while remaining relatively silent about consciousness as a whole. When some form of decomposability is assumed, the distinguished components are not separated as modules from the comprehensive function, as was the case when Marr studied stereopsis as part of early vision. What is important is whether the levels of analysis that Marr distinguished play a role in this method of the study of NCC. For example, is there a requirement of explicitly articulating computations or algorithms, such that it enables researchers to relate different NCC's to each other? Or is the focus different, perhaps on the implementation level? Is there an interest in constraints that mutually help to determine the theories at different levels of analysis, as Marr suggested?

4.1 Identifying a minimal yet sufficient neural correlate – of what?

Considering the great variety of phenomena and concepts associated with the 'last surviving mystery' of consciousness, Dennett raises the possibility that: "[p]erhaps the various phenomena that conspire to create the sense of a single mysterious phenomenon have no more ultimate or essential unity than the various phenomena that contribute to the sense that love is a simple thing" (Dennett 1993 23).⁶⁷ If that is the case, then it seems that the preliminary requirement of defining a cognitive task or providing a task analysis is to be avoided. Instead of seeking to identify and explain such a simple thing, NCC researchers have been focusing on a series of particular

⁶⁷ Classifications of consciousness are rich and to some extent divergent among authors – compare those offered in the invited review of (Zeman 2001), according the conceptual analysis of (Bennett and Hacker 2003), or in the argument for investigating computational correlates alongside neural correlates of consciousness in (Atkinson, Thomas et al. 2000).

phenomena generally considered to belong to the class of events to which the concept of consciousness applies. To this end, a method has been developed which allows them to study a series of cognitive events in an animal or human subject such that they assume to be capable of distinguishing these events according to their property of being conscious, or not. Let us explain this curious method.

As we have seen, explanation can occur in various ways. Not only can we facilitate the explanation of a phenomenon by taking up different perspectives or levels of analysis, or by dividing the phenomenon into components that facilitate the project of investigation and explanation, as we have witnessed above. Another method that can simplify explanatory work is to describe a phenomenon not by providing a comprehensive account of it, but by contrasting it with a comparable phenomenon which differs from it in a specific respect: why does X occur instead of Y (Ruben 1992)?⁶⁸ Instead of explaining the riding of a car, we can explain the contrast with the car's riding back and forth or we can aim to explain the difference between singing mere melodic lines and singing a song.⁶⁹ When offering such a 'contrastive explanation', it is much easier to decide which information is relevant and which is not, since we aim to explain a specific difference between comparable phenomena and not a single phenomenon in its entirety.⁷⁰ For example, contrastive explaining of the car's riding direction will direct us to the gearbox while taking the activities of the motor, wheels and breaks for granted. Explaining the difference between vocalizing and singing may inform us specifically about the neural activities that correlate with semantic processes in vocal expression.

Returning to consciousness, there is broad consensus that consciousness refers to a phenomenon or set of phenomena that is absent when subjects are asleep, unconscious or in coma and which reappears when they wake up.⁷¹ However, that transition is too intricate and ramified to allow detailed study, as it encompasses several distinct sensory and cognitive functional differences. Accordingly,

⁶⁸ Ruben warns against overstretching the importance of contrastive explanation: he argues that not all explananda are contrastive and that contrastive explanations are not all equally apt for traditional non-contrastive explanation (Ruben 1992 39-44). What is relevant in the present context, though, is that contrastive explanation corresponds particularly well with the subtraction method in many neuroscientific experiments, in which the neural activities correlated to two - test and control - conditions are compared.

⁶⁹ Analogously, the causal relation to be investigated is not just a relation between a cause and a particular phenomenon, but between two comparable causes, where one of them is causing that particular phenomenon while the other is responsible for another phenomenon (Schaffer 2005).

⁷⁰ Indeed is such contrastive explanation quite common in consciousness research, probably because of its relative modesty as (Hohwy and Frith 2004) argue.

⁷¹ Revonsuo has a different approach, holding that the dreaming brain is similar to consciousness with respect to phenomenal awareness and thus a good place to start consciousness research (Revonsuo 2000).

researchers have been keen to identify cases in which a similar, specific and more minute transition is observable. That transition should consist of at least a single component associated with consciousness, that is to say with conscious experience. This component requires subjects to report their experience behaviorally or verbally, excluding sleeping or comatose subjects who are principally unable to report on the transition. Indeed: “the fundamental methodological problem faced by any rigorous research program on consciousness is the subjectivity of the target phenomenon” (Metzinger 2000 1). Given this and in view of Dennett’s reminder that we should not focus our study of consciousness on a single phenomenon, the question is how we can isolate specific and minute transitions from non-conscious to conscious experience that are reportable by subjects. If researchers gather evidence about a number of such transitions, hope is that together they may teach us something about the structure and neural implementation of consciousness more generally.

Probably the first study to focus on such a minute yet reportable transition was with monkeys, investigating the ‘Neuronal Correlates of Subjective Visual Perception’ (Logothetis and Schall 1989). Interestingly, it did focus on visual motion yet not merely on the task of visual motion processing according to Marr’s approach. Instead, it targeted the transition of visual motion perception into and out of reportable experience. It consisted of a binocular rivalry task, in which the eyes of the monkey were continuously presented with two different stimuli, which could not be perceived simultaneously. If the left eye is presented a downward movement and the right eye an upward movement, the reportable perception will alternate between the two stimuli, even though they themselves remain stable. The question of the authors was whether the neurons involved in processing visual motion are the same that correlate with the reportable perception of movement. They found that indeed there were single cells whose firing rates correlated with the subjective visual perception as reported by the monkeys.

This correlation between recorded firing of specific neuronal cells and the subjective visual perception is fascinating but still leaves room for alternative explanations, as we are not sure whether transient perception correlates with other neural activities as well.⁷² Echoing Marr’s worry that his approach of investigating components of vision could fail if vision turned out to be a highly interactive process, the authors note that the correlation itself needs further interpretation because: “the

⁷² (Noë and Thompson 2004) argue that this analysis of the binocular rivalry task obscures the fact that the experience of this rivalry – or similarly in perception of a bi-stable figure like the Necker-cube – is a perceptual experience in itself, which is temporally extended and encompasses the two distinct percepts between which that rivalry exists. Although we concur with their observation, it does not exclude the option to focus on the rivalry as such.

perception-related modulation observed in these neurons may be a result of feedback from higher centers” (Logothetis and Schall 1989 763).⁷³ It would require additional research to exclude the possibility that this reportability may be due to such a higher center and not rely just on the neural areas investigated in this study. In its aim to investigate a minute and specific transition of a percept in and out of a reportable status, this approach seeks correlating neural activities that are as specific for that transition as possible.

The method of searching for a neural – or neuronal, as it is sometimes called – correlate of consciousness has become ever more articulated since this first monkey study. The general hypothesis behind such research has remained that: “it is useful to think of consciousness as being correlated with a special type of activity or perhaps a subset of neurons in the cortical system,” without assuming that consciousness is a single phenomenon, nor that its neural correlates always have to be identical (Crick and Koch 1990 266). In fact, researchers are looking for a specific neural correlate that is not a mere side-effect of a case of conscious experience but that is itself responsible for a specific case. An often-used definition captures the NCC research goal more technically: “An NCC is a minimal neural system N such that there is a mapping from states of N to states of consciousness, where a given state of N is sufficient, under conditions C, for the corresponding state of consciousness” (Chalmers 2000 31). In the definition, the clauses of minimality, sufficiency and mapping relation are obviously the most remarkable. However, given the fact that minimality and sufficiency conditions are not exceptional for this research method, we will not discuss them further.⁷⁴ The other condition, that one is looking for a mapping relation, is relevant here and deserves further scrutiny.

The mapping relation may remind us of the ‘loose relations between levels’ that Marr defended (Marr 1982). In the present context, the result of the search for an NCC is not the development of a causal model but a result that allows researchers: “a mapping from states of N [= a minimal neural state, MK] to states of consciousness” (Chalmers 2000 31). Such a mapping or correlation between neural and conscious states gains in relevance as the specificity of those states becomes more and more

⁷³ As discussed more below, recurrent processing has more recently indeed been found to be correlated with reportability of visual stimuli – see e.g. (Lamme and Roelfsema 2000).

⁷⁴ Obviously, considering the brain as a whole as an NCC is uninformative, which is why: “it makes sense to ‘start small’ in the search for an NCC” (Chalmers 2000 32). However, if a particular NCC were not sufficient for producing a specific conscious experience, the explanation would probably still be incomplete, as other neural activities would be equally necessary. Single cell studies like the one by Logothetis e.a. mentioned above exemplify this strategy (Logothetis and Schall 1989). fMRI activations patterns are less specific but may in turn inform us about task modulations by less relevant processes related to motivation, attention, learning and memory, as one of the authors argues (Logothetis 2008). Excluding such ‘neuromodulations’ from the NCC is an important methodological task.

explicit. One way to achieve this is to focus not on consciousness generally, but on a specific content of consciousness, as in the study of binocular rivalry, when the aim was to identify the neural correlates of the reported visual percept. However, this requires a preliminary basis for identifying a particular state of consciousness as being indeed just that: a state of consciousness.

Such an identification is not as simple as it seems. In the case of binocular rivalry, for example, one could ask what exactly the unit of consciousness or the precise state of consciousness is for which researchers have been trying to identify a correlating neural state. Is the conscious experience of binocular rivalry in fact a couple of experiences that alternate between two different conscious states, or is it the singular and comprehensive conscious state of experiencing the switch between percepts? Should we identify the neural correlate of two different, snap-shot like conscious states – involving two non-identical and non-overlapping percepts – or should we look for the NCC of a temporally extended conscious state which includes the alternation between two different percepts (Noë and Thompson 2004)? Arguing in favor of the latter, Noë & Thompson emphasize that researchers have to pick the right level of analysis when looking for an NCC: “the content of perceptual experience is personal-level content, not subpersonal-level content” (Noë and Thompson 2004 18).

The distinction between a personal-level and a subpersonal-level of analysis here distinguishes between a phenomenological account and a neurophysiological account of an experience. The methodological distinction between levels proposed by Marr may be more informative in this case: what is the computational theory or task analysis of the conscious state for which researchers are trying to identify an algorithm and implementation theory?⁷⁵ Is it plausible to expand the computational theory of bistable percepts such that it includes both alternating percepts and their alternation, or disregard the bistability as such? Obviously, both cases will yield different results in terms of the correlated implementation theories, especially in terms of the temporal structure of the alternation. When bistability is disregarded, the temporal dynamics of alternation may be left out of the equation completely – simply as a consequence of a particular task analysis of bistable perception.⁷⁶

⁷⁵ The authors reproach the NCC methodology for not providing us with a causal explanation of an experience by a neural state (Logothetis 2008). However, a reliable correlation between two states at different levels of analysis can in itself be considered a part of a scientific explanation. Moreover, it is not so much a causal but a constitutive relation between the neural state and the conscious state that we should expect in this case. This is acknowledged in the mechanistic account of explanation, discussed in the next section of this Part I.

⁷⁶ A framework that aims to explain the fact that it is continuity and discreteness together that make up our phenomenal experience is offered in (Fingelkurts and Fingelkurts 2006). These authors also contend that the phenomenal level at which continuity and discreteness are both experienced should be the starting point for the development of an explanatory framework.

The foregoing may warn us against avoiding the issue of first analyzing the function or task to be investigated. This warning was already expressed in Marr's methodology and therefore not new to us. Nonetheless, the NCC approach to some extent avoids the development of a computational theory, since consciousness is accepted as being a research topic so hard to determine that the hope is that its delineation may be reached by evidence converging on one or more neural correlates of states accepted as being conscious – even without a more technical analysis or definition of those states. This hope has led to another strategic choice, aimed at limiting the phenomenon for which an explanation is required. It has been generally accepted in NCC research to distinguish between research of the 'background state of consciousness' like wakefulness or dreaming and research of specific contents of consciousness (Chalmers 2000). Even though the two are likely related, their investigation is facilitated by separating them. Clearly, the experience of binocular rivalry or bistable percepts is possible only on the condition of being conscious and awake, but these aspects of the background state are left out of the explanation of the contents of the experiences. Although the reason for the omission may be convincing, the consequences should not be neglected.

4.2 Further limiting the phenomenon and its correlates

We started the previous section by quoting Dennett's consideration that the phenomenon of consciousness is so complex that we may be well-advised to accept that it in fact consists of various phenomena that do not conceal a single one (Dennett 1993). The search for a NCC concurs with that consideration, as it usually aims to determine a minimal neural correlate for a particular conscious content. Given the wide-ranging variety of phenomena associated with consciousness, this strategy avoids both the conceptual and the empirical challenge of presenting a unifying definition of some sort. Rejecting the 'simple thing' assumption about consciousness also provides a practical advantage, since it allows researchers with as many possible ways of studying and accessing consciousness as there are associated phenomena (Churchland 2005).⁷⁷ Even though these problems appear particularly intricate for the phenomenon of consciousness, several of the lessons learnt from these regarding the methodology of research are relevant in other domains of cognitive neuroscience as well.

Common to these lessons is that limiting the phenomenon under scrutiny increases our potential for explaining it. This obviously comes at a cost, however. In the previous

⁷⁷ Obviously, recognition of the variability and the multifaceted nature of consciousness does not preclude researchers to attempt a unified theory of consciousness. However, as Seth and others argue in presenting their framework, such a theory should include measures for both various quantitative and the qualitative aspects of consciousness (Seth, Izikevich et al. 2006).

section, contrastive explanation was invoked, as it allows a focus on explaining the difference in reportably perceiving either picture of a binocular rivalry task, without having to explain all neural conditions necessary for perceiving any picture at all. In the domain of consciousness studies, a similar distinction has been made: “between the neural correlate of background state of consciousness (wakefulness, dreaming, etc.) and the neural correlate of specific contents” (Chalmers 2000 33). Of course, it should be realized that the NCC resulting from the study of a specific content of consciousness may produce that state only against such a background state, while the conditions of this background state are not themselves included in that particular NCC. Accordingly, the NCC was defined as a minimal and sufficient neural system for producing the contents of a very specific conscious state – minimal, that is, against the presence of such a background state. Since this background state remains unspecified, the interpretation of the associated NCC is troublesome, because such an NCC by itself is not really sufficient for any conscious experience to emerge. Indeed, “momentary activations in local neural populations in visual system, which the definition of the NCC is carefully constructed to pick out, are not sufficient for any actual experience had by any actual subject”(Neisser 2012 683-684).⁷⁸ In spite of these reservations, it is not uncommon in cognitive neuroscience research to limit the phenomenon under study and its explanation in this manner.

Not surprisingly, the strategy has raised serious criticism. Following their line of critique aimed at the delineation of binocular rivalry, the same authors draw attention to the background state and the activities of an animal or human subject that has conscious perception: “if perceptual content depends on the skillful activity of the whole animal or person, making use of its capacities for eye, head, and whole body movements, and for directed attention, then it becomes questionable whether there is any such thing as a minimal neural substrate sufficient to produce experience” (Noë and Thompson 2004 17).⁷⁹ If this should be taken to suggest that an NCC for a particular experience should include all correlates responsible for this rich texture, then it is unhelpful since that would likely involve the most part of the brain. Again,

⁷⁸ Consider for example Searle’s conception, involving a rather different task analysis – or computational theory, in Marr’s words – of consciousness: “[w]e should not think of perception as creating consciousness, but as modifying the preexisting conscious field” (Searle 2004 81). In addition, he writes that the NCC thesis regarding correspondence between states is in fact a trivial one.

⁷⁹ As Noë and Thompson remark –while referring to Varela- this line of research is impossible to carry out “independent of the sensorimotor context of the animal as a whole” (Noë and Thompson 2004 13). That brings them to the conclusion concerning the NCC approach that “[w]hat makes this account internalist is that it views the experience as supervenient on neural processes alone” (Noë and Thompson 2004 20) and consequently independent of interactions with the body and the external world. However, as imaging techniques are used to investigate whether apparently unconscious or comatose patients are in fact capable of some kind of voluntary responses, such bodily interactions are circumvented indeed (Schnakers, Perrin et al. 2008). Unsurprisingly, this approach of sidestepping normal behavioral evidence of consciousness by using imaging evidence has been criticized by Hacker et al. who fears that at some point a neural correlate may even be accepted to overrule behavioral evidence (Nachev and Hacker 2010).

however, we can just accept that skillful activity and directed attention are generally involved and still aim for the limited NCC that would allow us to merely distinguish between the specific contents of two different conscious states. Obviously, this additional limitation does not make the interpretation of the function of the NCC any easier. A third limitation deserves mention.

As much as NCC research does generally depend upon disregarding the background state of consciousness and avoids a definition or task analysis of conscious experience, it must also ‘screen off’ the resulting NCC such that it captures the relevant condition. For if we find some activation – via single cell recording or fMRI, for example – in correlation with a particular consciously experienced content, we still must decide which feature of that activation is responsible for the correlated conscious state. We cannot just be satisfied by pointing out the involvement of a neural area or network, for it may still not inform us what property of that area yields the conscious state. As a case in point, Hardcastle observes that there is an ongoing dispute between ‘smallists’ and ‘largists’ about how to define the neural correlate. On the one hand, there are those who expect that small scale quantum effects are responsible for consciousness, on the other hand there is increasing interest in large-scale dynamics within the brain (Hardcastle 2000).⁸⁰ So is it the particular small scale quantum effects in specific neural cells that are relevant, or are these common to all cells in all conditions and thus irrelevant, whereas it is their firing in synchronicity with other cells that determines the conscious state? Researchers hope that this uncertainty can be settled by comparing several different NCCs, assuming that this will eventually tell us which of their properties overlap in most cases, and which do not.⁸¹

In sum, the search for NCCs depends upon strategic limitations in several senses or directions: limiting both the target phenomenon of consciousness, distinguishing it from a background state of consciousness and from other relevant activities and cognitive functions of the subject. A way to do this can be to engage in contrastive explanation of why a particular conscious state obtains instead of another. In another direction, the limitation involves the determination of the relevant neural state that correlates with the conscious state under scrutiny. These limitations can be found in

⁸⁰ Hardcastle mentions the necessity of a ‘screening-off analysis’ which amounts to determining the probability of potential causes A and B co-occurring with the event, and comparing these with the probability of the event itself, in order to decide which cause is doing the actual work (Hardcastle 2000). As far as we can see, this works best with a rather simple branching structure of causal relations, but is less adequate in a complex and interacting system including recurrent processes.

⁸¹ For example, reviewing existing evidence for NCCs of conscious visual percepts, Lamme argues that it is a specific - recurrent - type of processing that is the relevant and overlapping NCC (Lamme 2006). More recently, it has been suggested to look for recurrent processing in vegetative, anaesthetized and dreaming subjects as a way of looking for residual conscious experience (van Gaal and Lamme 2012).

other strategies of research as well, to be sure. However, they are relatively prominent in NCC research, which is the reason why they were discussed here. Moreover, they enable us to formulate more specifically what in fact we expect from an explanation and in what sense a NCC may not be adequate.

4.3 Lessons on explanation drawn from the NCC research

Despite considerable research efforts, consciousness remains an elusive phenomenon. This is due to several factors. As mentioned earlier, researchers have not been able to agree upon a particular definition of consciousness. In fact, it has been argued that consciousness should not be considered as sharply distinguishable from non-conscious processing, but as a gradual phenomenon instead (Cleeremans and Jiménez 2002). Similarly, a different taxonomy has been proposed in order to explain and distinguish different forms of conscious and related phenomena (Dehaene, Changeux et al. 2006). Divergence has not only affected the definitions of consciousness or the distinction between conscious and other states, but there is also no consensus about the criteria for empirically determining whether a specific instance represents consciousness (cf. Lamme 2006 497, fig. 1 on measuring conscious visual experience). Not surprisingly, consensus regarding the functional role of consciousness is equally lacking, as much as it is absent with regard to neuroscientific explanatory accounts of it. So how do researchers respond to this situation and what progress in explaining consciousness do they aim at?

First, researchers generally accept the diversity and heterogeneity of conscious phenomena and experimental approaches to consciousness, as we mentioned in the previous section. Doing so, they fail to comply with the requirements of preliminary conceptual analysis or task analysis, as we saw above. Indeed, NCC researchers don't assume their object of research to be a 'simple thing' and appear to be relatively liberal in accepting phenomena into the domain of consciousness research, expecting that with such a wide net they will eventually be able to distinguish relevant from irrelevant phenomena as suggested in (Churchland 2005). And of course, they subsequently develop strategies to exclude phenomena from that net while simultaneously trying to find some overlapping property of those phenomena that are accepted as belonging to the space of consciousness.

Since evidence of consciousness in everyday life or in research generally depends upon behavioral or verbal activities, in many cases consciousness has become associated with properties of behavior, reflection, language, and so on. The second strategy contributing to progress in research therefore consists of showing that consciousness does not always depend upon the co-occurrence of these phenomena

and can even be dissociated from them. For example, a recent review of NCC research concludes from evidence that conscious experience “does not require sensorimotor loops involving the body and the world, does not require self-reflection (or language), and does not reduce to attention” (Tononi and Koch 2008 240). Such a dissociation of these types of response from consciousness itself is a useful result. It should not be forgotten, however, that any such dissociation is only possible on the basis of some sort of a previously accepted delineation of phenomena that attest to consciousness. Then again, the first strategy was to be liberal in accepting phenomena in the domain of consciousness research. So how is a limitation eventually brought about?

Apart from being liberal in admitting phenomena to the domain of consciousness research and then trying to dissociate these phenomena from other, associated cognitive functions, a third strategy is to look for some overlapping property on which the remaining phenomena converge. It is not surprising that such a property will be relatively abstract or general. For example, one proposal that finds widespread recognition is that consciousness is related to information processing.⁸² On top of this - not unexpectedly in computational neuroscience - several authors who employ the NCC approach stress the role of consciousness in learning processes or information integration processes (cf. Cleeremans and Jiménez 2002 ; Crick and Koch 1995 ; Dehaene, Changeux et al. 2006 ; Lamme 2006 ; Tononi and Koch 2008). For example, the distinction between consciousness and pre-conscious or subliminal information processing is claimed to be the accessibility of the relevant information by certain processes (Dehaene, Changeux et al. 2006). And another theory claims that: “the level of consciousness of a physical system is related to the repertoire of causal states (information) available to the system as a whole (integration)” (Tononi and Koch 2008 253, italics in original).

Although Marr referred to an only loose interdependency between the three levels of explanation, as we discussed in section I.3.4, he was aiming for more than just correlations between the theories formulated at those levels. He argued that even though a task could be performed by different algorithms and researchers should therefore allow for potentially different neural implementations, researchers could still use the theories available at a particular level to constrain the number of – theoretically or empirically – plausible theories at another level, and vice versa.

⁸² Given that NCC research can be distinguished as focusing on either background states of consciousness or on consciousness of specific contents, we would expect that information is a crucial element at least in the latter strand of research. Indeed, information theory is considered to be a major development in the scientific study of consciousness, since “information theory is also the first step in solving the difficult problem of bridging the mental and the physical domains” (Rees and Frith 2007a 14).

Moreover, in this search for mutual constraints he assigned a primary role to the top level or the computational theory (Marr 1982). This is a more specific aim of combining theories at different levels of analysis than the mapping relation involved in NCC research seems to prescribe (Chalmers 2000). Nonetheless, within the NCC we can also discern a fourth strategy: the aim to constrain potential theories on the basis of an overlapping neural correlate that research has yielded.⁸³ Although there may not have been a very strict delineation of phenomena admitted to NCC research in advance, a post factum comparison of relevant results can suggest a potentially defining neural correlate of those phenomena. Consequently, NCC researchers may be tempted to define the psychological function in terms of this neural correlate and in so doing reverse the lessons that were presented in the first chapters of this part, by the approach that assigned priority to conceptual analysis or Marr's approach. For example, Lamme goes so far as to claim that ultimately, one should redefine consciousness in terms of recurrent processing, since it is the 'key neural ingredient of consciousness' for most – though not all – phenomena associated with the problem of consciousness (Lamme 2006 499).⁸⁴ However, this still leaves the demand to explain how this key ingredient can be responsible for the particular phenomena that are associated with the conscious states. For that, a more elaborate specification of the relation between the levels of analysis and integration of such an ingredient with other relevant components is needed than the NCC requires. And eventually, as Chalmers states elsewhere, research of consciousness phenomena cannot stop with the detection of neural correlates but should instead result in: "specifying a *mechanism* that performs the function" (Chalmers 2007 227, italics in original). What such an explanatory mechanism might look like will be discussed in the next section.

⁸³ Here, the mapping relation is not just unidirectional, but bidirectional. Even though the definition mentioned earlier seemed to suggest otherwise, there was also hope that eventually an NCC might help in reaching conclusions about undecided cases like patients in coma or 'locked-in' patients, and that an NCC might help explain the functional phenomena associated with consciousness (Chalmers 2000).

⁸⁴ As neurobiological as this key neural ingredient appears to be and as remote from a phenomenological or psychological theory of consciousness, it does bear a similarity with the theory of consciousness at the psychological or computational level which Dennett presents as his concluding hypothesis about consciousness: "our capacity to relive or rekindle contentful events is the most important feature of consciousness – indeed, as close to a defining feature of consciousness as we will ever find" (Dennett 2005 171). Common to both definitions is the fact that contents or information are not just forwarded to further stages of processing.

To sum up this chapter on NCC research, we have seen that even a higher cognitive function like consciousness is open for empirical research aiming to combine different sources of information. In doing so, the NCC approach appeared to be more liberal in certain methodological respects than the other approaches thus far, while some requirements still stand out. Taking the requirement of a strict definition or task analysis of consciousness at first more lightly, the approach still has to use some provisional concept of consciousness in order to decide whether a particular state can be admitted. Furthermore, finding a neural correlate itself is not yet sufficient for explaining its role in the conscious phenomenon in which it is involved. Obviously, words like ‘mapping’ or ‘correlation’ that capture the relation between conscious and neural states are mere indices or filler terms that are in need of further specification.⁸⁵ What such specification might look like and how these states can be related is left open in the NCC approach and needs other resources for its articulation. Additional resources are also needed in order to acknowledge the differences in relevance between the many neural correlates that will accompany any cognitive function. In the next section, we will discuss such further resources when we focus on the mechanistic explanatory approach. One of the advantages of that approach is that it provides insight into the differences in relevance of the various components of a mechanism responsible for a phenomenon – including the dynamics that underlie the shifts in components’ relevance, as their causal relevance may also change under changing environmental conditions.

In conclusion, the NCC approach can at least play a role as a heuristic in commencing research of a complex cognitive phenomenon, inviting research of its neural correlates – in the form of neural localization or otherwise - in the absence of its delineation. However, sooner or later the requirements that were passed over will resurface and ask for fulfillment.

⁸⁵ Again, “correlation studies cannot determine whether such neural activity plays a causal role in determining the contents of consciousness” {(Rees and Frith 2007b 560)}.