Orthogonality of Phenomenality and Content

Gottfried Vosgerau, Tobias Schlicht & Albert Newen
Institut für Philosophie, Ruhr-Universität Bochum

Corresponding author:
G. Vosgerau
Institut für Philosophie
GA 3/153
44789 Bochum
gottfried.vosgerau@rub.de

Acknowledgements:

The authors are thankful to the VolkswagenStiftung for financial support and to John Barresi, David Chalmers, David Papineau, and an anonymous reviewer for very helpful comments.

Abstract

This paper presents arguments from empirical research and from philosophical considerations to the effect that phenomenality and content are two distinct and independent features of mental representations, which are both relational. Thus, it is argued, classical arguments that infer phenomenality from content have to be rejected. Likewise, theories that try to explain the phenomenal character of experiences by appeal to specific types of content cannot succeed. Instead, a dynamic view of consciousness has to be adopted that seeks to explain consciousness by certain ways of processing representations. Therefore, only empirical methods that are able to investigate the dynamics of the mind can be used for the "quest for consciousness" proper. Moreover, the central intuitions about consciousness are best explained when phenomenality and content are clearly distinguished.

1. Introduction

We do regard it as a fact—as most people do—that there is something phenomenal to our conscious experiences, something it is like to have an experience or to be in a conscious mental state (cf. Nagel 1974). However, since there are very different views on this phenomenon associated with very different terms and since we do not want to predefine the ontological status of the phenomenal aspect of experience, we have chosen the somewhat cumbersome word "phenomenality". Various questions concerning phenomenality arise in the philosophical debate (see, for example, Chalmers 2002). However, independent of the specific question being asked, there seems to be one background assumption shared by many philosophers, namely that phenomenality is a matter of the content of a representation. According to this view, a mental representation is phenomenally conscious if it has a certain kind of content (or if there is another representation with a specific kind of content). In our view, this background claim deserves closer examination.

We will thus not discuss the several possible positions one can take with respect to the several possible questions, nor will we start with the thought-experiments that sharpen our intuitions about the concepts in question. Instead, we will begin to review some theories and empirical facts about conscious experience that will lead us to the conclusion that phenomenality has nothing to do with content. Phenomenality is better explained without the presupposition of what is sometimes called "phenomenal content" (e.g. Block 2005). Instead, phenomenality should be understood as a uniform relational property of conscious representations which is orthogonal to content. We will then apply this consequence for a fundamental critique of some classical arguments and theories of phenomenal consciousness. After having explored some of the implications for empirical research, we will give an account of some central intuitions we have about phenomenal experiences.

2. Theories of Phenomenal Consciousness

2.1 Neuroscience: The Ventral Stream Hypothesis

One of the most successful neurological hypotheses of phenomenality is the ventral stream hypothesis which is confined to visual experience. As formulated by Milner & Goodale (1995), it states roughly that only activation in the ventral stream leads to phenomenally conscious states, whereas activation in the dorsal stream remains unconscious and is used for the visual guidance of action. In other words, activation in the ventral stream is regarded as a necessary and sufficient condition for phenomenal visual experiences. Jacob & Jeannerod (2003) showed, however, that this picture is too simple for empirical reasons. They report data from neglect patients suffering from a lesion in the right inferior parietal lobe, which should be regarded as belonging to the dorsal stream (and is definitely not part of the ventral stream). These patients loose their phenomenal consciousness of visual attributes (color, shape, orientation, etc.) although their ventral stream is working properly. That is, they have the "right" activation in the ventral stream yet they lack phenomenal consciousness of the stimuli. According to Milner and Goodale, however, this ventral activation should suffice for rendering the visual representations conscious. Moreover, Weiskrantz (2006) cites evidence from blindsight subjects to the effect that neither dorsal nor ventral activation is sufficient for conscious vision, concluding that Milner and Goodale's picture is too simplified. Hence, the ventral stream hypothesis is wrong in claiming that activation in the ventral stream is sufficient for phenomenality. In general, it seems that phenomenal consciousness cannot be simply reduced to the activation of one region or circuit (physiologically individuated) in the brain.

2.2 Cognitive Science: The Enactive View

The so-called enactive view of visual phenomenal consciousness has gained much attention in the last decade. According to O'Regan and Noë (2001a, 2001b), we have to distinguish the perspectival properties (P-properties) of a perceived object from its actual properties (A-properties) which are both revealed to us in visual perception: An object, for example, that is actually round (A-property) appears elliptical (P-property) from most perspectives. When we change our perspective, i.e. when we move through space, the P-properties of the object change as well (we see different elliptical shapes), but the A-properties (being round) do not. Moreover, the P-properties change systematically with our movements: given the same movement, the same change in P-properties occurs. Thus, there is a systematic relation between the perceiver's movement and the perceived P-properties, which Noë and O'Regan call sensorimotor contingencies. They hold that the mastery of these sensorimotor contingencies is the ability to perceive A-properties. With respect to the example this means: Once I "know"[i] the sensorimotor contingencies between my movements and the round object, I can predict every change in P-properties for each movement, and this means that I am able to see the actual roundness of the object. So, conscious perception of A-properties is nothing else than mastering the sensorimotor contingencies. The peculiar phenomenality of such perceptions is hence explained on the grounds of systematic dependencies between movements and perceptual input, and the "knowledge" about them.

In this paper, we cannot discuss the enactive account introduced by Noë and O'Regan in any more detail. Therefore, we will neither give arguments against nor in favor of their account. For the sake of argument, we will assume that their theory can adequately describe and explain phenomenal consciousness for visual A-properties. If so, the following problem arises nevertheless: the phenomenality of P-properties cannot be explained by the same account, although we are (can be) clearly phenomenally conscious of P-properties (e.g. the

elliptical shape). The reason is that P-properties themselves do not systematically covary with our movements.[ii] Quite on the contrary, each time we move, the P-property switches (and this means: there is a new property). Hence, each time we move, the P-property just perceived "vanishes" and is replaced by a new P-property. Therefore, there simply are no sensorimotor contingencies for P-properties. Nevertheless, we are able to consciously perceive them.

Another argument shows that the attempt to explain the conscious experience of P-properties in terms of the mastery of sensorimotor contingencies leads to an infinite regress (cf. Schlicht & Pompe 2007). According to Noë (2004), we learn (are informed) about A-properties via the conscious perception of P-properties (which are both construed as elements of the same perceptual content), and thus by mastering the relevant sensorimotor contingencies. But since P-properties are supposed to be consciously perceived, they call for even more basic P-properties informing us about the P-properties (call them PP-properties), which in turn require the assumption of PPP-properties and so on ad infinitum. To evade this infinite regress, we either have to assume that P-properties are not consciously perceivable (which is *de facto* wrong), or that PP-properties are not consciously perceivable (but there are no PP-properties, as shown in the argument above), or to abandon the idea that the phenomenality of P-properties can be explained by the enactive theory.[iii]

Even if it is true that the phenomenality of A-properties can be explained by the enactive view, the phenomenality of the more basic P-properties cannot. Therefore, phenomenality of visual properties cannot be reduced to or be explained by one type of content, namely the one constituted by the mastery of sensorimotor contingencies. There has to be at least another type of content (the one P-property percepts have) which also contributes to phenomenality.

2.3 Philosophy: The Fine-Grainedness Argument

In arguing for the existence of non-conceptual representations, quite a few authors have relied on the representations involved in phenomenally conscious visual perception (Raffman 1995, Peacocke 1998, 2001, etc.). The main argument is that visual experience is much more fine-grained than our conceptual apparatus.[iv] It is concluded that the representations underlying visual experience are non-conceptual. Therefore, there are two types of content: conceptual and non-conceptual. Some philosophers conjecture on these grounds that only the non-conceptual content is responsible for phenomenal consciousness (Peacocke 1983, Tye 1995).

However, there are examples in which the phenomenal content of visual experience is altered by concepts. An engineer who was conceptually informed by the manual about a certain machine will have a different content of conscious visual experience when first looking into the machine compared to a layman who never had training. The engineer will see a lot of parts that work together whereas the layman will only see many screws and the like: There is empirical evidence that our perceptual experience is influenced by top-down processes in the way that concepts we have acquired change the production processes that lead to perceptual experience (Frith & Dolan 1997, Bar 2003). In general, it seems straightforwardly possible that we learn about a certain thing theoretically, i.e. by verbal instructions, such that we form a clearly conceptual representation of this thing. However, we are normally able to consciously perceive the theoretically introduced thing as a thing of this very kind when we first encounter it. Hence, conscious experience does have some conceptual parts. It cannot be explained by non-conceptual representations alone.

Therefore, phenomenal consciousness cannot be reduced to the level of non-conceptual representations—we can also be phenomenally conscious of conceptual (visual)

representations. Again, phenomenality cannot be reduced to or be explained by one type of content.

3. The Orthogonality Hypothesis

All this points to the conclusion that no theory is able to explain phenomenality by appeal to a certain type of content. Phenomenality is not bound to a certain type of content—not to the contents processed in the ventral stream, nor to the content based on sensorimotor contingencies, nor to non-conceptual content. Thus, the relationship between phenomenality and content is orthogonal: Various types of content can be phenomenally conscious or unconscious, but there is no type of content which has the exclusive right to become phenomenally conscious.[v] (Of course, there are some contents which cannot become phenomenally conscious, and maybe there are even contents which we can have only consciously, but this does not play a role for the present argument.) In principle, every type of content can become conscious or remain unconscious.[vi]

3.1 There is No Phenomenal Content

Representations are the kind of things that make information available to us. They have a content (often called the representational content). We can distinguish red from yellow *because* we have different representations: red-representations and yellow-representations. More precisely, it is the content RED and the content YELLOW that distinguish the two representations. Independent of consciousness, this discrimination ability remains the same. For example, if red and yellow spots are presented as the masked stimulus in a masked priming experiment, subjects are clearly able to discriminate the two colors although they are not consciously seeing them (Schmidt 2000, 2002). Under normal conditions, however, the same color spots can be seen consciously and be discriminated by subjects, of course. Hence,

the discrimination ability is a matter of the so-called representational content of the involved representations.

We cannot only distinguish red from yellow spots, but we can also distinguish red-experiences from yellow-experiences. How come? A lot of philosophers answer this question analogously to the question why we can distinguish red spots from yellow spots: Because there is some content, namely a "phenomenal content", of our representations which brings about this ability. Doubtlessly, experiencing red differs phenomenally from experiencing yellow. Hence—so these philosophers argue—, there is a "phenomenal content" which is responsible for this difference. As we will show in this paper, this explanation of the discrimination ability for phenomenal experiences is mistaken.

If the argument were right, then every conscious representation would have (at least) two contents: a representational content which makes available to us some information about the world, thereby providing us with the ability to distinguish different things, and a "phenomenal content" which marks the phenomenal difference between different experiences, thereby enabling us to distinguish different experiences. In order to "prove" that the two contents fall apart, many philosophers introduced a whole family of thought-experiments including inverted spectrum scenarios (e.g. Block 1990), different kinds of zombies (cf. Kirk 2005), swampmen (originally introduced in Davidson 1987; see also Dretske 1995), and Twin Earth cases with and without traveling scenarios (e.g. Tye 1998). The main idea is always the same: either, the representational content is held constant while the "phenomenal content" is changed (inverted spectrum) or removed (zombies, Swampman), or the "phenomenal content" is held constant while the representational content is changed (Twin Earth). The first kind of thought-experiments is based on intuitions, namely the intuition that inverted spectra and zombies are possible. We will postpone the discussion of this intuition to the last section. We

will now show why Twin Earth stories are no good argument for the distinction between the two supposed contents.

The Twin Earth stories are based on the assumption of externalism (Putnam 1975, Burge 1979, etc.), i.e. the thesis that the representational content of a representation is partly determined by factors external to the representing system. If externalism is true, then it suffices to change these external factors in order to change the representational content. Exactly this is done in the Twin Earth experiments. It is then argued that there are cases where certain external factors are changed without the representing person noticing it (usually by letting the person travel between Earth and Twin Earth). It seems absurd to conjecture that the phenomenal aspects of the person's representation changes as well, since the change is unnoticed, and "phenomenal contents" are exactly the kind of things with which we are immediately acquainted (Tye 1998). Hence, to argue that such travels change the "phenomenal content", we would have to argue that there can be unnoticed differences in the "phenomenal content". This contradicts the assumption that the "phenomenal contents" give us the ability to distinguish different experiences. So it is concluded that representational content and "phenomenal content" fall apart, since the one can change independently of the other.

In our view, there is something wrong with the thesis of externalism of representational content. And if externalism is false, then it is not the case that the representational content changes by traveling to Twin Earth, and the thought experiment does not show that representational content and "phenomenal content" fall apart.

What is wrong with externalism? In one sentence: Externalism is based on the false assumption that representational content is identical with linguistic meaning (the content that sentences have). This thought-language principle is widely accepted in the philosophy of mind. However, we have argued elsewhere (Newen & Vosgerau 2007) that the meaning of

linguistic utterances and the content of mental representations should be sharply distinguished. Externalism holds for linguistic utterances but not for mental representations. Apart from empirical evidence (mentioned in Newen & Vosgerau 2007), the main argument for the sharp distinction is that the meaning of utterances is determined by a linguistic community (we will not argue for this claim since we take it to be common ground), while the content of mental representations is determined by the behavioral dispositions of the representing system. Mental representations are introduced in order to explain flexible behavior. An ant, for example, is able to go back to its nest on a straight line after an unsystematic search for food (Gallistel 1993). If the ant is displaced on its way back, it will look for the nest in the place where it would be had it been displaced to the same amount. In order to explain this behavior, we assume that the ant has a representation of the nest enabling it to find it even in absence of nest-stimuli. Why is it a representation of the nest and not, say, of the tree? Because the ant is disposed to go to the nest and not to the tree. (Note that the representation can be wrong, as it is the case when the ant is displaced. Hence, the nest does not need to be there for the representation to be a nest-representation.) The content of the representation is, unlike the meaning of utterances, determined by the behavioral dispositions of the representing system. Hence, representational content is not identical with linguistic meaning.

In the Twin Earth case, the dispositions of the traveling persons do, by presupposition, not change. Therefore, their representational content cannot change, although the meaning of the linguistic utterances, with which they express their mental representations, does change. Thus, the Twin Earth thought experiments do not show that representational content and "phenomenal content" fall apart. They can show, at most, that linguistic meaning and content of mental representation fall apart. However, we are not able to distinguish things because we are able to know what we say (linguistically express), but because we have mental

representations. And since we cannot tell H_2O from XYZ when traveling to Twin Earth, we neither have an H_2O- nor an XYZ-representation but rather a water-representation (for details see Newen & Vosgerau 2007; Newen & Bartels 2007).

Thus, the Twin Earth stories give us no reason to believe that there are two kinds of content. Therefore, the best explanation for our ability to discriminate different phenomenal experiences is simply the assumption that they have different (representational) contents—an assumption which we already made on different grounds independent from phenomenality. Hence, our ability to discriminate different states (and thereby discriminate different properties) relies only on one type of content, namely representational content. There is no further kind of content, in particular no "phenomenal content". Henceforth, we will use the term "content" to mean the only content we have to assume: representational content[vii].

So far, we have only sketched criteria of content individuation. However, it is crucial to our argument that, on the one hand, mental representations do not have the same content as their linguistic expressions and that, on the other hand, representations with one and the same content can be conscious or unconscious. Therefore, we will have a closer look at the individuation criteria of (representational) content.

3.2 Representational Content and Phenomenality

The content of representations is determined by behavioral dispositions, or—more precisely—by the target of the behavior to which we are disposed. Because the nest is the usual target of the ant's behavior, we say that it has a nest-representation rather than a tree-representation. Of course, the behavioral dispositions of humans are far more complex than those of the ant. For example, humans are, under normal conditions, disposed to linguistically express the contents of their conscious mental representations (and this might explain the fact that they were often mistakenly equated with the contents of these linguistic expressions).

Simply saying that the behavioral dispositions determine the content of a mental representation will hence only postpone the problem to the problem of the individuation of behavioral dispositions. What do we have to add then?

There are different families of theories formulating criteria for content individuation (similarity-theories, causal theories, functionalism, teleosemantics, etc.). Each family has its specific problems which we cannot discuss here in detail.[viii] What we want to highlight here is that any theory of content has to start with or rely on the behavioral dispositions in the following sense: The reason to ascribe a system a contentful mental representation is the wish to explain (and predict) the *behavior* of this system. More specifically, the content is what is taken to explain the behavior. In this way, the content necessarily depends on the behavioral dispositions of the system to which a mental representation is ascribed. At least, the object of the mental representation has to be the object of the behavior to be explained (at least in standard cases, i.e. leaving aside cases of misrepresentation). Assuming that a cow-representation never leads to some sort of cow-directed behavior means to give up the explanatory role of content. Or formulated the other way round: because we want to explain cow-directed behavior, we assume that the system in question has a cow-representation.

Nevertheless, there are two caveats to be added to this story. First, the content of representations cannot simply be equated with the behavioral dispositions or be too tightly bound to them, since the behavior of humans is far too complex. If the content of a representation would change with every slight change in behavior, we would end up with a unique content for every piece of behavior. This clearly contradicts the aim to explain and predict behavior. Therefore, it is just the object of the behavior one is disposed to that enters into the individuation of content, and not the whole fine-grained complexity of behavior.[ix] Second, the content cannot simply be equated with the objects, i.e. be individuated externally. The reason for this is that there are cases of misrepresentation, and in these cases there is no

object to be directed at.[x] However, if we want to explain such cases by the same mental representations, we have to assume that the content is not dependent on an actual target object of the behavior. The content of a representation is hence determined by the target object of the behavior, which has to be explained and to which the bearer of the representation is disposed, under normal circumstances. The clause "under normal circumstances" is, of course, introduced to exclude cases of misrepresentation and it is clear that it cries out for further explication. However, for the present purpose, this characterization will suffice, since we will not be dealing with cases of misrepresentation in this article. [xi]

Since the content of a representation is determined by behavioral dispositions, it cannot be an intrinsic feature of the representation (it cannot depend on properties of the representation that the representation would also have in isolation). Rather, it is an extrinsic property, i.e. a relational property which depends on other entities distinct from the representation in question. However, this does not imply content-externalism since the relational features playing a determining role for content can be and are indeed internal, i.e. the entities to which the representation stands in certain relations are themselves in the mind (or—if you like it more physicalistic—inside the skull). (They are features of the processing which leads to behavior.)

Moreover, since one and the same representation can be unconscious or conscious, the phenomenal character of the representation cannot be an intrinsic feature of it either.[xii] Rather, phenomenality arises in virtue of some relational features of a representation. The fact that we can tell apart two experiences from each other is still explained in terms of the (representational) content of a representation. In the case of an unconscious representation, our ability to distinguish one representation from another cannot be deployed because some relational features are missing: the representation is not available to introspection. However, availability to introspection is not dependent or even grounded in a special content.

To say that a content C is conscious is to say that we have a representation with content C that we consciously experience. But there is no difference between a conscious and an unconscious content C—they are the same! The difference is only that in the first case the representation is conscious and in the second case it is not—due to the lack of some relational feature(s) and not in virtue of the conscious content having some further content. The content can explain our ability to distinguish things and experiences, but it cannot explain the phenomenal aspects of conscious experience.

Many empirical studies have shown that representations with the same content can be processed either consciously or unconsciously. For example, people suffering from *blindsight* do not consciously perceive an object but are nevertheless able to grasp it and reliably make a "forced choice" guess as to what object it is, whether its color is A or B, or whether it is located in C or D (Weiskrantz 1986, 2006). Blindsight-patients have residual visual capacities with respect to a great range of stimuli, from simple forms and colors to emotional facial expressions (deGelder et al. 1999). Because of such behavioral dispositions, these patients are assumed to have representations, namely object-representations. Hence, it is clear that a representation with content C can be conscious or unconscious.

Split-brain patients provide a second example since they reveal (under experimental conditions) the existence of causally efficacious yet unconscious mental representations independently of forced-choice situations. In these patients, the *corpus callosum* is cut such that the main connection between the two hemispheres is severed, resulting in largely independent information processing in the two hemispheres. Due to lateralization, information presented in the right visual field is processed in the left hemisphere and vice versa. After the information-flow is interrupted, most mental representations are computed "intra-hemispherically" (Colvin & Gazzaniga 2006, p. 182) such that neither hemisphere receives the information processed in the other.[xiii] In the famous "chicken-claw experiment"

(Gazzaniga 1985), it was shown that a picture presented to the left eye has been computed in the right hemisphere, but due to the disconnection of the hemispheres this representation did not become conscious although it has obvious causal effects that can only be explained if we assume it to have the same content as it has in healthy subjects (who perceive the picture consciously). Thus, the difference between the conscious representation of a picture in healthy subjects and the unconscious representations of the same picture in split-brain patients is not a difference in content. It is rather a matter of a representation's being integrated in the right kind of way.

Similar cases of unconscious processing of otherwise conscious contents due to pathologies can be found in various forms of agnosia. For example, patients suffering from simultanagnosia fail to consciously register a letter that is composed out of small tokens of another letter although this representation affects their behavior. It can be shown that patients register unconsciously the form of the large letter (Karnath et al. 2000). Patients suffering from prosopagnosia learn the right names for unrecognized faces faster than wrong names (Sergent & Pocent 1990). Moreover, some can even correctly select the right face to a presented name, although a conscious feeling of familiarity is missing (De Renzi & Di Pellegrino 1998). Hence, it seems that the contents of representations can be left unaffected by pathological changes in phenomenal consciousness. Thus, again, phenomenality cannot be a matter of content.

A further argument can be made by demonstrating the core functional role of attention, namely to influence phenomenality without changing content. First of all, it has to be noticed that attention and consciousness should be sharply distinguished (cf., e.g., Koch & Tsuchiya 2006, Lamme 2003, 2004). The view we would like to put forward here is that attention serves mainly the function of integrating some contents into a wider field of processing, thereby making them phenomenally conscious.[xiv] Attention itself can be driven bottom-up by

the saliency of an object or top-down volitionally controlled (Koch 2004, pp. 153-171). Regardless of the direction of influence, attention usually brings about a change in phenomenal consciousness since it makes available the contents to other modules. Because these other modules comprise conceptual modules and other interpreting modules, in most cases attention also goes hand in hand with a refinement of content. For example, when registering something moving in the periphery of my visual field, I will automatically (bottom-up driven) direct my attention to the according point and, in this way, gather more information which ultimately results in a refinement of content such that I perceive the black cat of the neighbor in the garden, say. However, such a refinement of content cannot be explained by attention,[xv] for there are cases where attention is detached from such a change: When you are attending to a certain object in the sky, for example, because you just registered something moving there, this shift of attention does not necessarily bring about a refinement in content. It happens often enough that we are still left with the (now conscious) content "There is something moving" without further refinement. Only after minutes of active investigation (gathering of information), we suddenly come to recognize that it is, e.g., a helicopter. Therefore, the usual refinement of content is a "byproduct" of focusing attention due to the availability to further modules, which is the core function of attention.

To summarize, contents can be made available and thus phenomenally conscious by attending to them. The often occurring change in content is rather a consequence than a necessary feature of attention. If this picture is true, then it provides further evidence for our claim that phenomenality is orthogonal to content. Evidence for the picture itself comes from pathologies of attention, namely neglects. In the famous case of patient P.S. suffering from hemineglect, the contents that could not be attended to remained unconscious although they reliably directed decisions and were therefore represented (Marshall & Halligan 1988). Moreover, it has been shown that attention can even alter the "visibility", i.e. the

phenomenality of contents in a precisely measurable manner. Carrasco, Ling & Read (2004) showed that attention enhances particularly contrast sensitivity and spatial resolution, which in turn helps to specify representational content. When observers focused attention to a certain stimulus location via a peripheral cue, then they "reported that stimulus as being higher in contrast than it really was, thus indicating a change in appearance with attention" (Carrasco et al. 2004, p. 310). In this way, focal attention "boosts" the stimulus contrast such that a subthreshold (invisible) contrast stimulus appears as if it were above threshold, i.e. focal attention can make an otherwise unconscious representation of one and the same stimulus conscious. In accordance with our view, Carrasco et al. take this to be evidence for a "contrast gain model, in which attention allows for greater neuronal sensitivity (decreased threshold), suggesting that attention changes the strength of a stimulus by enhancing its 'effective contrast' or salience" (Carrasco et al 2004, p. 312).

Koch and Tsuchiya (2006) present empirical evidence to the effect that consciousness and attention are two different brain processes and sketch a framework similar to the threefold distinction between subliminal, pre-conscious, and conscious processing proposed by Dehaene et al. (2006). The crucial difference is that Koch and Tsuchiya allow both for conscious perception without attention and for selective attention in the absence of consciousness. The first case can plausibly be interpreted in terms of familiar examples such as the distracted truck driver (Armstrong 1997). In general, we can be conscious of states in the world without attending to the details of them.[xvi] But this perceptual content is typically neither *specific* nor conceptualized. Rather, it is merely *generic* or *sparse*, i.e. a sort of "summary" of the scene, such that we consciously perceive only the "gist" of what could be attended to.[xvii] The second case is more interesting. It is claimed that top-down attention on an object or event does not always result in conscious perception of it. Among the cited examples are aftereffects induced by an invisible object although the spatial region is attended

to (He, Cavanagh & Intriligator 1996). In this case, it seems that attention is not being directed at a single object (which causes the aftereffect) but to the whole spatial region not discerning several objects. Therefore, the conclusion that attention is possible without consciousness is not drawn in the original paper. Instead, the authors suggest, in perfect concordance to our view, "that the attentional filter acts in one or more higher visual cortical areas to restrict the availability of visual information to conscious awareness" (He et al. 1996, p. 335).

Another case in which attention is said to be possible without consciousness is that of masked priming. This, however, only shows that attention is not by itself sufficient to render a representation conscious (moreover, we need a stable representation which cannot be formed when the stimulus is masked); it does not contradict the interpretation that attention is a means of rendering representations conscious (if the background conditions are fulfilled). In visual search, when, for example, someone is looking for her car keys in a messy room, attention is focused on an invisible object. But it is not clear whether such examples show that attention does not usually serve the function of making events or objects conscious, since the latter case is, of course, also possible if the relevant object does not even exist. So, it is not at all clear what it could mean that we attend to things that we do not perceive.[xviii] Rather, attention seems to be most naturally described as a selection mechanism that filters out certain parts of what we *perceive* (and not what we want to perceive) for further processing. In the searching case, we are attending to every single corner of the room trying hard to find the key, but in no sense we attend to the key (unless we have found it).

Taking together the conceptual and empirical considerations, content and phenomenality seem to be dissociated: One and the same content can be conscious in one subject or at one time and unconscious in another subject or another time. Moreover, attention

can be described as the mechanism that makes available different contents to other cognitive modules, thereby rendering otherwise unconscious contents phenomenally conscious.

One direct consequence of this view is that phenomenality does not come in pieces: it is rather a uniform relational property of representations. The difference between various phenomenally conscious experiences is accounted for by the difference in representational content. However, there is no "phenomenal content" over and above the representational content. In particular, there are no intrinsic properties of the representations that help to explain phenomenality. Thus, phenomenality is not constituted by pieces of phenomenal experience but rather by a uniform relational property which is the same for all conscious representations. This is to say that it is not possible to type conscious experiences by phenomenality (or phenomenal content), since the phenomenality is the same for all conscious experiences. All that can be done is typing it by (representational) content: The difference between what it is like to see red and what it is like to see green lies exclusively in the content of the two different representations, namely in the difference between RED and GREEN. There is no phenomenal aspect over and above these contents that differs for the two experiences: The phenomenality is the same for both (it is a uniform property). Many philosophers have used the term "qualia" for describing phenomenality, and they have used it in various ways. Independent of the exact meaning of "qualia"[xix], one presupposition that must be shared by all those who use the term is the following: that there is a difference between a red-quale and a yellow-quale (i.e. that experiences can be typed by phenomenality). Thus, whoever uses the term "qualia" is hence committed to the view that there are pieces of phenomenality (e.g., one piece for each color). But phenomenality is a uniform relational property which does not come in pieces. Therefore, qualia (in whichever sense) do not exist. The difference between, e.g., a red-experience and a green-experience is due to the difference

in representational content, whereas their phenomenality is due to the uniform relational property of being made available to other cognitive modules.[xx]

Before turning to a positive account of phenomenality, we will apply the results of the discussion to some classical arguments.

4. Reply to Classical Arguments

If the orthogonality thesis of phenomenality and content is right, we cannot infer from the content of a representation alone that it is phenomenally conscious. This allows us to attack some classical arguments.

Ned Block (1995, forthcoming) argues that there is what he calls P-consciousness (phenomenal consciousness as opposed to A-consciousness, access-consciousness). The arguments all run as follows: When being asked, you realize that there has been some "experience" before (e.g. the clock going tick-tack all the time although this fact was not A-conscious to you, i.e. you didn't consciously register it).[xxi] At the time the question is posed, you suddenly become aware of a fact you had not been aware of before although the same fact was present. So, when asked, the experience of the fact becomes available to you and you can access it (A-consciousness). Now it is presumed that this mental state you are in at the moment you were asked has the same *content* as the non-A-conscious representation of the fact earlier. On this basis, it is concluded that the non-A-conscious representation must have been conscious as well. This new kind of consciousness, which does not bring about "global availability" in your mind, is called P-consciousness, and it is conjectured that it has some phenomenality to it as well.[xxii]

Given the orthogonality of phenomenality and content, this is a non-sequitur. Just because one content is conscious at one time, it cannot be concluded that the same content is

also conscious at any other time. Therefore, from the fact that the tick-tack sound becomes conscious when being asked about it, we cannot conclude that it has been conscious before the question was posed. The better explanation is that there is some content, which becomes phenomenally (as well as access-) conscious at the moment of being asked.[xxiii] In this picture, although there might have been a tick-tack-sound-representation before the moment the question is posed, this tick-tack-sound-representation was indeed unconscious. At a later time (when being asked), it became conscious and thereby exhibited phenomenality (presumably *because* it was integrated into a wider field of reasoning or "broadcast" into a "global workspace"). There is no reason to believe that there is P-consciousness in addition to and separate from A-consciousness.

Fred Dretske (1993, 2006) argued that, when looking at a wall from some distance, say, we are phenomenally aware of all different bricks in the wall. The argument runs roughly as follows: The content of my experience is the wall. However, the wall is composed out of bricks which are (for present purposes) clearly visible. Hence, whenever I experience the wall, I thereby experience the bricks. Since I'm consciously aware of the wall, I have to be consciously aware of *all* the bricks as well. Dretske's argument is that if we are asked whether there is one blue brick among all the orange ones, we can answer *No* with great confidence. Here, again, from the content (the wall being composed out of bricks) it is inferred that the whole content (all elements of the content, namely every single brick) is phenomenally conscious. Although we might have a representation of each and every single brick, this representation need not be conscious even if there is a conscious wall-representation at the same time. Again, we cannot conclude from the sameness of content (being a wall-representation) that both representations are conscious (or unconscious). One may reply to Dretske's argument that *at the time* the confident judgment is made, attention was focused on the task of answering the question and information was thereby made available which wasn't

available and thus not phenomenally conscious before. As long as the information about the bricks is not integrated into some wider field of reasoning in some appropriate way, it will not become conscious. Of course, it is possible to integrate only the information about the wall but not the one about the bricks. Therefore, Dretske's argument is invalid.

More generally, "experience" seems to be a notoriously unclear notion. In one way of using it (Dretske's way) it refers to the information our sensory organs receive. Just because every brick is visible, our eyes receive information about them. In another sense, it refers to what we consciously perceive. The two are mixed up in the argument: Of course, if we want to talk about the information I get, it is clear that I experience the bricks by experiencing the wall. Nevertheless, it is not true that I'm consciously aware of everything that I have information about. Therefore, one may concede a conscious experience of a generic content like "wall made of bricks" while denying a conscious experience of each single brick.

David Rosenthal (1997) and others have defended the so-called higher order thought (HOT) theory of consciousness. The central claim is that a representation becomes phenomenally conscious if and only if it is itself represented by a distinct and separate higher order thought. In other words, the fact that a first-order state is the content of a higher order state makes the first order state conscious. A virtue of this view is that it explicitly allows for any mental state (and thereby for any representational content) to occur consciously or unconsciously, since it denies that consciousness is an intrinsic feature of mental states. But the relational feature it postulates as necessary for a representation to become conscious is again a certain form of representational content: The representation itself has to be represented by another mental state, a higher order thought. But in general, the fact that something, a table, say, is the representational content of a mental state does not make this very thing conscious. If being the representational content of a mental state were sufficient for something to become conscious, this should make the perceived table itself conscious, just

like higher order thoughts are supposed to make mental states conscious (Dretske 1995, Byrne 1997). This is clearly nonsensical. Another problem is the explanation of the functional role of conscious thoughts (as opposed to unconscious thoughts): if I have a conscious thought that p, this leads to different dispositions to act compared to only having the unconscious thought that p (e.g. it leads to "spontaneous" actions as opposed to mere re-actions[xxiv]). The HOT-proponent would have to explain this observation completely in terms of the content of the HOT. But the content of p and the related higher-order thought cannot explain these differences, since the HOT-component "I am now having the experience…" (or something similar) does not give us a better understanding of spontaneity, let alone phenomenality. The reason is, as argued for above, that the content of a representation is (in the first place) determined by the target of the behavior it should explain. However, spontaneity of action is not a target of behavior, and phenomenality is even less so. No matter what content we introduce over and above the content that explains our behavior, it can never explain the changes in behavioral dispositions (since the targets do not change) nor can it explain phenomenality. (The second point can be regarded as a reformulation of the Zombie-intuition: I can have the same behavioral dispositions without phenomenality. We will come back to this point at the end of the paper.) Again, contents cannot give us any explanation of phenomenality since phenomenality and content are orthogonal.

Some have proposed that conscious representations are self-representing (Kriegel 2006). In other words, they propose that the conscious representation and the HOT collapse into one single representation. However, the same problem already discussed above arises for this variant of theories: Contents are orthogonal to phenomenality, and therefore no kind of content can explain phenomenality.

5. The Dynamic View

So far, we have criticized some popular arguments and theories in the philosophy of mind. However, if the orthogonality thesis is right, what could serve as a theory of phenomenality then? It seems that the crucial point in all the examples discussed above of phenomenally conscious experiences is some kind of integration of information.[xxv] Such integration may be achieved to varying degrees such that one extreme is totally generic content while the other extreme is maximally specific content. Thus, the degree of information integration is really a matter of content but not of integration: When a representation with fully determined content is integrated, we may speak of the integration of detailed information, in contrast to integrated representations that have only generic content and so constitute cases of lesser degrees of content integration. If this is plausible then it offers a new interpretation of the famous distinction between phenomenal and access consciousness: The contents which we have access to and are phenomenally conscious of are those of integrated representations; however, in one case, they have a fully determined representational content, while in the other case, they have only a sparse or generic content. Nevertheless, both can play a causal role in reasoning, are poised for further cognitive processing by the belief system, and so on. In the case of Sperling's experiment (discussed in detail in Block forthcoming), for example, the sparse content that there were 12 numbers on the screen is conscious, accessible, and reportable, although the specific content (which numbers were on the screen) is only accessible after attention has been directed to one line (and it seems to be a constraint of attention that we cannot determine more than the specific numbers of one line). However, we are not at all conscious of (i.e. cannot integrate the specific representation of) the unattended lines.

Contents of the ventral stream, for example, become phenomenally conscious only if they are integrated with contents of the right inferior parietal lobe, which codes the objects'

locations in allocentric terms (see Jacob & Jeannerod 2003). The contents that are based on sensorimotor contingencies are themselves integrated information from visual input and motor information. The P-properties, however, can become conscious as well, presumably when they are attended to, i.e. when they are integrated into some process of reasoning. The contents of visual experience can be influenced by concepts exactly because they are integrated into some space of reasoning. My conceptual understanding of the machine forms part of my visual experience because I integrate what I see into what I know.

What we have sketched here is often called a dynamic view of consciousness. There are many philosophical, neurological, and physiological models of where and how such a dynamic integration can take place. There is, for example, the hypothesis that contents become integrated by phase locking of the signals (Crick & Koch 1990, Engel et al 1999). A more computational theory is the "global workspace" model (Baars 1988, Dehaene & Naccache 2001, Dehaene et al 2006), according to which some representations can be projected into a global workspace making their contents available to all different modules.[xxvi] This is close to what Block (1995) calls A-consciousness. Again, we do not wish to discuss the varying models here. We would just like to highlight that dynamic models of consciousness do not conflate phenomenality and content, nor do they leave the change in behavioral dispositions untouched: The integration of the contents just does explain the availability of the contents to all other modules which brings about a massive change in the possibilities to behave.

Of course, there are also famous hybrid theories that combine a dynamic view with constraints on content, such as Tye's PANIC theory (Tye 1995). According to this theory, only poised abstract non-conceptual intentional contents are conscious. Obviously, certain kinds of contents are made responsible for phenomenality. However, according to our argumentation, no kind of content can carry the burden of explaining phenomenality. Except

for being poised, all attributes listed by Tye specify constraints on content and therefore cannot contribute to an explanation of phenomenality. Only the property of being poised is a relational (functional role) property that refers to some (potential) integration of the content. Yet, this part of the theory points in the same direction as the other dynamic theories. Of course, a lot more has to be said about what it is to be poised, or be accessible, or be available in global workspace. We do not want to imply that this notion is sufficiently clear—it is not. However, according to our argument it has the potential of being developed into a clear notion explaining phenomenality. Unfortunately, this cannot be done in this paper.

According to our argumentation, dynamic theories of consciousness are not only much more promising but are in principle the only theories that could possibly provide a better understanding of phenomenal consciousness. The overall picture we want to draw is thus the following: Representations with contents can be unconscious, and some of them can become conscious. One crucial condition for becoming phenomenally conscious is the integration of the representations into some bigger framework. In other words, the way of processing (namely integrative processing) seems to be the crucial property, which leads to phenomenality. Therefore, we have to find different explanations for the contents of states independent of whether they are conscious or not, on the one hand, and for the phenomenality of some of these states, on the other hand.

6. Empirical Questions

The so-called “quest for consciousness” becomes two-fold under the premise of the orthogonality of phenomenality and content. First, it should be noted that the notion of “phenomenal content” doesn't make any sense anymore. There is content on the one hand, and phenomenality on the other. This leads to the assumption that one of the classical explananda for a theory of consciousness, namely the content of “qualia”, falls out of the range of

consciousness studies proper. The content of mental states should be studied independently of their phenomenal status, since it is not the content that makes them phenomenal. The very same contents could be non-phenomenal (this is the basic intuition underlying zombie arguments). Trying to explain the content of phenomenally conscious states (as it is done by e.g. Milner/Goodale, Noë/O'Regan, and a lot of people investigating the visual/auditory cortex, and so on) is still an important part of consciousness studies but cannot possibly shed any light on phenomenality.

Second, the "real" quest for consciousness, i.e. the attempt to find the nature of phenomenality, should focus on the processes and mechanisms in virtue of which certain states or contents become conscious. The dynamic approaches we mentioned above point in this direction. If they are correct, consciousness and hence the neural correlates of consciousness are dynamic in nature. Thus, empirical research has to find (and use) methods that are able to measure the dynamics in the brain (which, of course, already happens). However, still a lot of consciousness research is done with static brain imaging tools like fMRI. These tools are not suitable to measure dynamical properties of the processing brain and hence can only be used to find the correlates of the contents of consciousness; they cannot provide us with an idea about the very nature of consciousness and phenomenality itself.

7. Analyzing Some Central Intuitions about Consciousness

One pressing question remains: Why do we have the qualia intuition expressed by the claim that my (conscious) experience as of something red involves a special content that the unconscious registration of the color does not involve? We accept the phenomenon that a conscious experience of a red object is different from an unconscious registration of a red object. But the difference is not a difference in the content—there is no special content involved in the experience, e.g. the quale red. What is different is the way in which the

content is accessible and—according to our view—this is realized by the integration of the representation. The content is (ultimately) determined by a relation between the cognitive apparatus of a person and the physical conditions in the environment (the targets of behavior). This explains the inter-individual differences in color experiences in the same situation (e.g. in the case of color blindness). The intuition that there is an aspect of the content of my color experience that I cannot share with anyone else can best be accounted for either by the mentioned uniqueness of my cognitive system or by the trivial fact that something which I consciously register is part of *my* consciousness. One special way in which my cognitive apparatus may be unique is the individual semantic network of concepts I have compared to other people. If consciousness is an information process that constitutes the integration of information in a special way, then the content of my color experience is a shade of blue but it may constitute a different content from your experience because I conceptually compare it with the picture of Yves Klein while you compare it with the impression of the blue sky in Italy.

Furthermore, consciousness may play a central role because some contents are only available on the basis of consciousness, e.g. an engineer having conceptual knowledge of a machine has different perceptual contents than a layman who is not able to integrate his perceptual information into such a detailed conceptual frame. But this is compatible with the view that for many contents—especially for perceptual contents—it is true that one and the same content may be available consciously or unconsciously.

At least in the philosophical debate, the Zombie-intuition has occupied a central place. It is the intuition that one and the same content which is accompanied by the same behavioral dispositions can be conscious in one system but unconscious in another (the Zombie), i.e. the content is not (metaphysically) necessarily connected to phenomenality. If this is the core intuition, it is right (according to our view). However, the intuition is often interpreted in a

broader way: All behavioral dispositions could be the same while the Zombie still does not have phenomenality. This, however, seems to be wrong. Blindsight-patients, who are, in a way, visual Zombies, do not have the same behavioral dispositions as normally sighted subjects; they only share some behavioral dispositions with them, namely those that can be explained by the content of the representations involved. All other dispositions (like spontaneity) are not shared. So, the Zombie-intuition stems from the right fact that all content-related behavioral dispositions can be equal and still phenomenality vary, but it is tacitly generalized to the wrong thesis that all behavioral dispositions could be the same in the absence of phenomenality.

Therefore, the orthogonality thesis is at least compatible with the core intuitions about consciousness.

8. Conclusion

Content and phenomenality are two distinct and independent features of mental representations. Both features are relational and not intrinsic, and each one can change while the other remains constant. Therefore, content and phenomenality are orthogonal. Thus, every theory that tries to explain consciousness by appeal to some contents is doomed to fail, and arguments that infer phenomenality from content are invalid. The only theories that can possibly shed light on the phenomenon of phenomenality are so-called dynamic theories which explain consciousness by a certain way of processing representations. Likewise, empirical research on phenomenality can only be successful if the dynamics of the mind are studied.

References

Amstrong, D.M. (1997), What is consciousness?, *in* N. Block; O. Flanagan & G. Güzeldere, eds., 'The Nature of Consciousness. Philosophical Debates', Cornell University Press, Cambridge, MA.

Baars, B. (1988), *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge.

Bar, M. (2003), 'A Cortical Mechanism for Triggering Top-Down Facilitation in Visual Object Recognition', *Journal of Cognitive Neuroscience* **15**, 600-609.

Block, N. (1990), 'Inverted Earth', *Nous Supplement: Action Theory and Philosophy of Mind* **4**, 53-79.

Block, N. (1995), 'On a confusion about the function of consciousness', *Behavioral and Brain Sciences* **18**, 227-247.

Block, N. (2005), 'Two neural correlates of consciousness', *Trends in Cognitive Sciences* **9**, 46-52.

Block, N. (forthcoming), 'Consciousness, accessibility, and the mesh between psychology and neuroscience', *Behavioral and Brain Sciences*.

Burge, T. (1979), 'Individualism and the Mental', *Midwest Studies in Philosophy* **4**, 73-122.

Byrne, A. (1997), 'Some like it HOT: Consciousness and higher-order thoughts', *Philosophical Studies* **2**, 103-129.

Carrasco, M.; Ling, S. & Read, S. (2004), 'Attention alters appearance', *Nature Neuroscience* **7**, 308-313.

Chalmers, D.J. (2002), Consciousness and its place in nature, *in* D. J. Chalmers, ed., 'Philosophy of Mind: Classical and Contemporary Readings', Oxford University Press, Oxford, pp. 247-272.

Colvin, M.K. & Gazzaniga, M.S. (2006), Split-brain cases, *in* Max Velmans & Susan Schneider, eds., 'The

Blackwell Companion to Consciousness', Blackwell, Oxford.

Crick, F. & Koch, C. (1990), 'Towards a neurobiological theory of consciousness', *Seminars in Neuroscience* **2**, 263-275.

Davidson, D. (1987), 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association* **60**, 441-458.

deGelder, B.; Vroomen, J.; Pourtis, G. & Weiskrantz, L. (1999), 'Non-conscious recognition of affect in the absence of striate cortex', *NeuroReport* **10**, 3759-63.

Dehaene, S. & Naccache, L. (2001), 'Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework', *Cognition* **79**, 1-37.

Dehaene, S.; Changeux, J.; Naccache, L.; Sackur, J. & Sergent, C. (2006), 'Conscious, preconscious, and subliminal processing: A testable taxonomy', *Trends in Cognitive Sciences* **10**, 204-211.

Dennett, D.C. (1988), Quining Qualia, *in* A. Marcel & E. Bisiach, eds., 'Consciousness in Contemporary Science', Oxford University Press, Oxford.

De Renzi, E. & Di Pellegrino, G. (1998), 'Prosopagnosia and alexia without object agnosia', *Cortex* **34**, 403-416.

Dretske, F. (1993), 'Conscious experience', *Mind* **102**, 263-283.

Dretske, F. (1995), *Naturalizing the Mind*, MIT Press, Cambridge, MA.

Dretske, F. (2006), Perception without awareness, *in* J. Hawthorne & T. S. Gendler, eds., 'Perceptual Experience', Oxford University Press, Oxford.

Engel, A.K.; Fries, P.; König, P.; Brecht, M. & Singer, W. (1999), 'Temporal binding, binocular rivalry, and

consciousness', *Consciousness and Cognition* **8**, 128-151.

Frith, C. & Dolan, R.J. (1997), 'Brain mechanisms associated with top-down processes in perception', *Philosophical Transactions of the Royal Society* **325**, 1221-1230.

Gallistel, C.R. (1993), *The Organization of Learning*, The MIT Press, Cambridge MA.

Gazzaniga, M.S. (1985), *The social brain. Discovering the networks of the mind*, Basic Books, New York.

He, S.; Cavanagh, P. & Intriligator, J. (1996), 'Attentional resolution and the locus of visual awareness', *Nature* **383**, 334-337.

Jacob, P. & Jeannerod, M. (2003), *Ways of Seeing: The Scope and Limits of Visual Cognition*, Oxford University Press, Oxford.

Karnath, H.O.; Ferber, S.; Rorden, C. & Driver, J. (2000), 'The fate of global information in dorsal simultanagnosia', *Neurocase* **6**, 295-306.

Kirk, R. (2005), *Zombies and Consciousness*, Oxford University Press, Oxford.

Koch, C. (2004), *The Quest for Consciousness: A Neurobiological Approach*, Roberts and Co. Publishers, Englewood, CO.

Koch, C. & Tsuchiya, N. (2006), 'Attention and consciousness: Two distinct brain processes', *Trends in Cognitive Sciences* **11**, 16-22.

Kriegel, U. (2005), 'Naturalizing subjective character', *Philosophy and Phenomenological Research* **71**, 23-57.

Kriegel, U. (2006), The same-order monitoring theory of consciousness, *in* U. Kriegel & K. Williford, eds., 'Self-Representational Approaches to Consciousness', MIT Press, Cambridge, MA, pp. 143-170.

Lamme, V.A.F. (2003), 'Why visual attention and awareness are different', *Trends in Cognitive Sciences* **7**, 12-18.

Lamme, V.A.F. (2004), 'Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness', *Neural Networks* **17**, 861-872.

Marshall, J.C. & Halligan, P.W. (1988), 'Blindsight and insight in visuo-spatial neglect', *Nature* **336**, 766-767.

Milner, A.D. & Goodale, M.A. (1995), *The Visual Brain in Action*, Oxford University Press, New York.

Nagel, T. (1974), 'What is it like to be a bat?', *Philosophical Review* **83**, 435-456.

Newen, A. & Bartels, A. (2007), 'Animal Minds and the Possession of Concepts', *Philosophical Psychology* **20**, 283-308.

Newen, A. & Vosgerau, G. (2007), 'A Representational Theory of Self-Knowledge', *Erkenntnis* **67**, 337-353.

Noë, A. (2004), *Action in Perception*, MIT Press, Cambridge MA, London.

O'Regan, J.K. & Noë, A. (2001a), 'What It is Like to See: A Sensorimotor Account of Vision and Visual Consciousness', *Synthese* **192**, 79-103.

O'Regan, J.K. & Noë, A. (2001b), 'A sensorimotor account of vision and visual consciousness', *Behavioral and Brain Sciences* **24**(5), 939-73.

Peacocke, C. (1983), *Sense and Content*, Clarendon Press, Oxford.

Peacocke, C. (1998), 'Nonconceptual Content Defended', *Philosophy and Phenomenological Research* **58**(2), 381-388.

Peacocke, C. (2001), 'Does Perception Have a Nonconceptual Content?', *The Journal of Philosophy* **98**, 239-264.

Putnam, H. (1975), The Meaning of "Meaning", *in* H. Putnam,'Mind, Language, and Reality', Cambridge University Press, Cambrigde, pp. 215-271.

Raffman, D. (1995), On the Persistence of Phenomenology, *in* Thomas Metzinger, ed., 'Conscious Experience', Imprint Academic, Thorverton.

Rosenthal, D.M. (1997), A theory of consciousness, *in* N. Block; O. Flanagan & G. Güzeldere, eds., 'The Nature of Consciousness. Philosophical Debates', MIT Press, Cambridge, MA.

Schlicht, T. & Pompe, U. (2007), 'Rezension von Alva Noë: Action in Perception', *Zeitschrift für philosophische Forschung* **61**, 250-254.

Schmidt, T. (2000), Visual perception without awareness: Priming responses by color, *in* T. Metzinger, ed., 'Neural correlates of consciousness', MIT Press, Cambridge, MA, pp. 157-170.

Schmidt, T. (2002), 'The Finger in Flight: Real-Time Motor Control by Visually Masked Color Stimuli', *Psychological Science* **13**, 112-118.

Sergent, J. & Poncet, M. (1990), 'From covert to overt recognition of faces in a prosopagnosic patient', *Brain* **113**, 989-1004.

Simons, D.J. & Rensink, R. (2005a), 'Change blindness, representations and consciousness. Reply to Noë', *Trends in Cognitive Sciences* **9**(5), 219.

Simons, D.J. & Rensink, R. (2005b), 'Change blindness: Past, present and future', *Trends in Cognitive Sciences* **9**(1), 16-20.

Tononi, G. (2004), 'An information integration theory of consciousness', *BMC Neuroscience* **5**, 42.

Tye, M. (1995), *Ten Problems of Consciousness*, MIT Press, Cambridge, MA.

Tye, M. (1998), 'Inverted Earth, Swampman, and Representationism', *Nous Supplement: Language, Mind, and Ontology* **12**, 459-477.

Vosgerau, G. (2007), 'We do not always express what we think. Comments on Lynne Baker', *Erkenntnis* **67**, 301-304.

Vosgerau, G. (forthcoming), *Mental Representation and Self-Consciousness*, mentis, Paderborn.

Weiskrantz, L. (1986), *Blindsight. A Case Study and Its Implications*, Oxford University Press, New York.

Weiskrantz, L. (2006), The case of blindsight, *in* M. Velmans & S. Schneider, eds., 'The Blackwell Companion to Consciousness', Oxford University Press, Oxford, pp. 175-180.

[i] "Knowing" is used here in a very broad sense including knowing-how, not in the narrow sense of true justified belief.

[ii] Covariation requires the constancy of the factors in the sense that the term cannot be applied if one of the factors vanishes instead of changing its "value". In the case of A-properties, however, the A-property is constantly present but only changes its appearance.

[iii] Noë (2004, 87) recognizes this problem but does not provide a satisfying answer.

[iv] Typically, the content of a phenomenal visual experience at one moment is much richer than we could possibly express linguistically (Peacocke 1998). Also, the difference between two colors which differ only slightly can be perceived when they are presented together, but it is impossible to tell which one is which when they are presented in isolation (Raffman 1995).

[v] Following Kriegel (2005) and others, the same point can be expressed by distinguishing the "qualitative character" from the "subjective character" of conscious experiences. Only the qualitative character can be explained in terms of special kinds of content, whereas the subjective character is independent of content. However, the term "subjective" points to a dependence on self-consciousness, which should not be assumed without further arguments. We will hence continue to use the term "phenomenality".

[vi] Note that this is even true for procedural content: Although we cannot become conscious of many procedures (e.g. speech production), we can become conscious of some (e.g. our problem-solving strategies).

[vii] This does not mean that representational content has necessarily only one level of individuation. Some might want to subsume, e.g., the linguistic meaning (or wide or public content) under representational content. If so, then the Twin Earth stories only show us that different levels of representational content fall apart. Nevertheless, the crucial point here is that none of these levels can contribute to an understanding of phenomenality, such that none of these levels can be equated with "phenomenal content".

[viii] For a discussion see Vosgerau forthcoming.

[ix] The problem of fine-grainedness is faced by classical functionalist theories that are based on abstract automata.

[x] Causal theories of content have basically this problem, since causation presupposes the causing object to exist.

[xi] Indeed, the consideration of misrepresentation leads to the assumption that the target object of a behavior is determined by the internal (contentless) processing of the behaving system. In this way, the above characterization that content can be determined by the target object can be true, although the fuller story tells us that the content is internally determined: The *acquisition conditions* for a certain (perceptual) content require actual target-directed behavior and are, in this sense, externalistic. However, the *possession conditions* for the same content do not include such a requirement—they rely on the internalistic conditions formulated above (see also Vosgerau 2007). The same is true for behavior: Think of the displaced ant mentioned above which shows the *same* going-straight-back-to-nest-behavior as the not displaced ant, even though there is no nest (if we would deny the sameness of behavior-type here, we would have to ask for two completely different explanations of the two kinds of behavior and thus deny the very possibility of behavioral sciences). Therefore, the *possession conditions* of behavior are also internalistic, whereas the *acquisition conditions* are equally externalistic (the ant could never acquire a nest-directed behavior if it never encountered nests).

[xii] As noted above, it might be the case that some contents can never become conscious and some are only available consciously. Nevertheless, the fact that some content can be both conscious and unconscious is sufficient for the conclusion that phenomenality is not intrinsic but orthogonal to content. (Possible content

classes of principally conscious or principally unconscious representations simply have no explanatory value regarding phenomenality.)

[xiii] Importantly, these patients do not experience two visual fields, neither do they completely ignore one side like neglect-patients do; indeed, usually they can compensate for their impairment by head movements etc.

[xiv] As will become apparent later, attention is neither sufficient nor necessary for phenomenality.

[xv] As we will argue below, attention can alter the quality of a stimulus and reduce external noise which serves but is not itself refinement of content.

[xvi] It is often assumed that the truck driver has no conscious experience of the road whatsoever. However, we think this is a bad description of the phenomenon that his phenomenally conscious representation has only a sparse content which is something like "There is a road". So, in our view, he is not totally unconscious but only conscious of very abstract and "generic" aspects of the situation.

[xvii] Some have argued against this possibility on the grounds of phenomena such as change blindness (e.g. O'Regan & Noë 2001a). However, in recent debates, there is more and more agreement that these phenomena neither support the „blindness"-interpretation nor necessitate sparse representation (Simons and Rensink 2005a, b).

[xviii] At least in a broad sense, so that introspection is not excluded.

[xix] Other philosophers deny that "qualia" is a meaningful term at all (e.g. Dennett 1988).

[xx] It is widely assumed that showing that phenomenality and content fall apart entails that there are qualia (because phenomenality cannot be explained by content). However, this argument is clearly incorrect, since there is a third possibility: that phenomenality is a uniform relational property.

[xxi] When being asked, you will form a representation with a content like "I have had the experience all along". This representation can, of course, be false: Indeed, you are not a-conscious of the sound before being asked such that you have no chance to know about this experience (if you would know about it, then you would be able to report it – but you are not at that time), such that it seems mysterious where you get the information from. Nevertheless, if this way of reconstructing the argument is not right because it is already part of the story that you had the experience all along, then we are faced with a *petitio*.

[xxii] Indeed, A-consciousness is assumed to be (at least conceptually) possible without P-consciousness, such that phenomenality is the mark of P-consciousness but not A-consciousness (see Block 1995).

[xxiii] You might thereby form a false representation with the content that you have been conscious of the tick-tack all along, but we doubt this. Rather, it is likely that you form a (correct) representation with the content that you have been registering the tick-tack all along such that it influenced you.

[xxiv] This can be demonstrated, for example, by the various pathological cases such as blindsight, hemineglect, etc. (see above).

[xxv] Tononi (2004) puts forward the view that consciousness corresponds to the brain's ability to integrate information where information is defined as the reduction of uncertainty. But his complex view entails the claim that experience—information integration—is a fundamental quantity like mass, charge, and energy. We do not want to commit ourselves to such an ontological claim.

[xxvi] Crucially, "global workspace" is primarily a functional notion, not a physiological one. It does not amount to the claim that representations are "broadcast" into a certain brain region. Instead, many neuronal networks may be supporting such a workspace.