Do Pseudonormal Persons Have Inverted Qualia? Scientific Hypotheses and Philosophical Interpretations

by

Uwe Meyer

University of Osnabrueck
Dept. of Philosophy (FB 02)
D-49069 Osnabrueck, GERMANY
uwmeyer@rz.uni-osnabrueck.de

in: Facta Philosophica 2 (2000), 309-325

Draft

Keywords: Inverted Qualia, Pseudonormal Vision
Vertauschte Qualia, pseudonormales Sehen

1. Introduction

In a recent paper[1], Martine Nida-Rümelin argues that there seems to be some scientific evidence for the actual existence of a phenomenon which has been considered as an object of pure philosophical speculation up to now: inverted qualia. Widely accepted theories about color vision and the genetics of color vision deficiencies, she argues, make it plausible to claim that some people are 'pseudonormal' in the sense that their visual system is slightly different from that of normally sighted persons, whereas their behavior concerning colors does not differ very much or perhaps not at all from ours: in general, those people (approximately 14 of 10000 males, according to an estimation of Piantanida, one of the neurophysiologists who brought up the hypothesis[2]) would make the same color discriminations and color judgments as we would. Insofar they *seem* to be normal, although their visual system is slightly different. In particular, pseudonormal persons would *call* the same things "green" and "red" as normally sighted people. However, it seems reasonable to assume that their different visual system makes pseudonormal people *see* green where we see red and see red where we see green, i. e. red things would *seem* to them as green things seem to us and vice versa.[3]

This would indeed be an exciting discovery, since it is generally assumed that cases of the described type of qualia inversion would pose a serious problem for some prominent philosophical theories of mind, in particular (but not only) for functionalism. It also seems to have implications for some more specific philosophical topics; playing a certain role in Nida-Rümelin's rejection of an indexical approach to the knowledge argument, for instance.[4] It is furthermore of quite some importance in the controversy between Nida-Rümelin and Wolfgang Spohn about the character of color predicates.[5]

In sum, it seems as if a purely scientific, philosophically 'unsuspicious' result could help settle some highly controversial philosophical debates. In this paper, I shall argue that this is not so. In fact, the scientific hypothesis that there may be pseudonormal people allows for different philosophical interpretations, these interpretations being dependent on more basic philosophical presuppositions.

2. Scientific Results and Hypotheses

First of all, let's have a brief look at the relevant results. Color vision as a whole is a very complex matter, but for the moment it is not necessary to go too far into detail: a schematic and simplifying overview about some basic facts will do. To begin with, there are three different kinds of photoreceptor cells on a human retina that are relevant for color vision: B-, G-, and R-cones. The letters "B", "G", and "R" stand as abbreviations for "blue", "green", and "red", but this naming is a bit misleading. It's true that the degree of activation of B-, G-, and R-cones is *typically* conjoined with sensations of blue, green, and red respectively, but the different cones are *not defined* by this connection. Instead, modern neurophysiology defines them by their morphological properties.

The cones normally contain three chemically different photopigments with different absorption properties. For the sake of convenience, let's call them B-, G-, and R-pigments (for the pigments which are *typically* but, as we shall see, *not necessarily* contained in B-, G-, and R-receptors respectively). The different absorption properties of the photopigments have the effect that the cones containing them are activated in different degrees by different kinds of light. For example, the maximum activation of cones containing G-pigments (i.e. typically G-cones) is reached at a wavelength of 535 nm of monochromatic light, whereas the cones containing R-pigments (typically R cones) are maximally activated at a wavelength of 570 nm. At a wavelength of 600 nm, cones containing R-pigments will be activated at approximately 75 per cent, whereas cones containing G-pigment will be activated only at 20 per cent.[6] How the respective cones would react to non-monochromatic light, which is the standard case, can be calculated from their response to monochromatic light.

Now, if light falls on a certain area of the retina of a human being, the question if she experiences the light as reddish, greenish or neither reddish nor greenish is, according to a simple model of the so-called opponent process theory[7], dependent on the *relation* between the average activation of R-cones and G-cones in that area. Speaking very schematically, this relation is 'worked out' in two steps by the nervous system 'behind' the cones. First of all, nerve cells connected to the R-cones in the area in question 'calculate' the average activation of these cones, and the same goes for the G-cones. At the second step, these average values are 'compared' with each other: If the average activation of the R-cones is higher than that of the G-cones, the person will experience the light as reddish; if, on the contrary, the average activation of the G-cones is higher than

that of the R-cones, the light will be experienced as greenish; and if there is no difference in activation, the light will seem neither reddish nor greenish. Furthermore, the degree of reddishness depends on the amount by which the activation of R-cones exceeds the activation of G-cones, and the opposite is true of the degree of greenishness.
Human color vision is not restricted to the red-green dimension, of course, but the other mechanisms (in particular the 'blue-yellow-channel' and the interaction between the red-green- and the blue-yellow-system) can be neglected here.
So far, these assumptions can be summarized in the following figure 1, which presents the different "levels" of color vision, beginning with the bottom level of pigments and ending at the top level of the "phenomenology" of color experiences.

5		if the average activation of G-cones exceeds that of R-cones, there is an experience of greenish light; if the average activation of R-cones exceeds that of G-cones, there is an experience of reddish light; if there is no difference between the average activations of G- and R-cones, neither a greenish nor a reddish shade will be experienced	
4		nerve cells compare the average activation of G-and R-cones in a certain area	
3	nerve cells calculate the average activation of R-cones in a certain area		nerve cells calculate the average activation of G-cones in a certain area
2	morphologically identified R-cones; activation is dependent on the kind of pigments the cones are filled with and the kind of light which falls on them		morphologically identified G-cones; activation is dependent on the kind of pigments the cones are filled with and the kind of light which falls on them
1	pigments		pigments

Fig. 1: Levels of color vision.

In the case of normally sighted people, the R-cones contain R-pigment and the G-cones contain G-pigment. This is different in people suffering from red-green-blindness. According to the prevailing theory about this defect, their R- and G-cones contain *the same* pigment, so that the average degrees of activation of R- and G-cones can never differ from each other: accordingly, nothing will ever appear reddish or greenish to those people. There are two different kinds of red-green-blindness: either R- and G-cones both contain R-pigment, or they both contain G-pigment.

Now, a widely accepted theory about the inheritance of this defect gives reason to suppose that there are some people in which the normal relation between R- and G-cones and pigments is inverted: their R-cones contain G-pigment, and their G-cones contain R-pigment. In those people, the genetic defects responsible for the different kinds of red-green-blindness both occur *simultaneously*, i. e. the genetic defect responsible for the first type has the effect that G cones contain R-pigments, and the genetic defect responsible for the second type causes the R-cones to contain G-pigments. As mentioned in the introduction, it is estimated that about 14 of 10000 males could suffer from this defect.

Of course, it would not at all be easy to identify those people, since, in contrast to individuals suffering from red-green-blindness, they *could* make a difference between reddish and greenish shades and, grown up in our linguistic community, they would *call* the same things reddish and greenish as we do. For that reason, they may be called *pseudonormal*. Indeed, it seems even possible that the differences in their visual system are *completely* compensated somehow, so that they would not differ from normally sighted people in any *behavioral* respect.

However, the theoretical assumptions which have been outlined in this section seem to imply that there is a psychological or mental difference between normally sighted people and pseudonormal individuals, even if their behavior is not at all affected by it. Remember that it is the *relative activation of morphologically defined R-cones and G-cones* which determines whether a given individual experiences something as reddish or greenish. This is assumed to be true *independently* of what kind of pigments the cones contain. Now since the distribution of R-pigment and G-pigment to R-cones and G-cones is inverted in the case of pseudonormal people and since the conditions under which the cones are activated is dependent on the pigments they contain, the theory seems to predict that pseudonormal people have experiences of reddishness when normal people experience something greenish, and vice versa. So, in comparison with normally sighted people, the experiences of pseudonormal individuals concerning reddish and greenish shades seem to be systematically inverted. At last, the old philosophical speculation about inverted qualia seems to have been given some scientific support.

3. Philosophical Conclusions from Scientific Results?

Nida-Rümelin argues that this inversion hypothesis, i. e. the hypothesis that the experiences of pseudonormal individuals concerning reddish and greenish shades are systematically inverted, poses serious problems for some well established philosophical theories of mind. Actually, it need not even be *true* for it to do so: it is already sufficient to consider the inversion hypothesis as *conceptually possible*.

In particular, this is true for a Wittgensteinian point of view, according to which "[r]ipe tomatoes look red to a given person iff it is appropriate according to the rules of the relevant language game to assert that they look red to the person at issue"[8]. In other words, if the (linguistic or non-verbal) behavior of a person in a normal life situation makes it seem fully appropriate to attribute to her experiences of red, it would be pointless to deny that she actually *has* experiences of red. Neurophysiological facts play no part here. For Wittgensteinians, this relation is not merely an empirical, but a *conceptual* matter: it simply seems meaningless to them to speak of two persons behaving equally in all relevant respects in a given normal life situation and yet being in two very different psychological states. If both of them look at a ripe tomato under normal conditions and utter truly, full-heartedly and consciously something like "This tomato looks red to me", it would be conceptually inconsistent to assume that the tomato *really* looks red to person A but green to person B, whatever the neurophysiological facts may be. It is obvious that this position is incompatible even with the mere conceptual possibility of the inversion hypothesis in its weaker form, according to which pseudonormal people behaving like normally sighted persons in ordinary life situations (though not necessarily *in all possible* situations) have systematically inverted experiences of reddish and greenish shades. So, after all, it would be enough to consider the inversion hypothesis as a comprehensible, conceptually possible assumption of color vision science to pose a serious problem for the Wittgensteinian point of view, perhaps even to refute it.

Functionalism is confronted with a very similar problem. There are, of course, many versions of this doctrine. Nida-Rümelin at first deals with a kind of causal role theory, according to which every psychological state is completely determined by the causal relations it bears to (1) a certain "sensory input", (2) other psychological states, and (3) a behavioral output. This thesis is often considered as *conceptual*, i. e. it is supposed to be a conceptual truth that psychological states are determined in such a way. But again, this "conceptual functionalism" seems incompatible with even the mere conceivability of the inversion hypothesis, for this hypothesis (in its stronger form[9]) implies the possibility that there can be two psychological states s_1 and s_2 in two persons p_1 and p_2 so that s_1 and s_2 completely resemble each other with respect to their causal roles but are nevertheless psychologically different. For example, s_1 and s_2 may both be caused by the perception of a ripe tomato and for their part cause the same beliefs, emotions, desires, behavior, etc., but whereas s_1 is a sensation of red, s_2 is (p_2 being pseudonormal) a sensation of green. So if the inversion hypothesis is conceivable, it seems possible that there are two *different* psychological states with exactly the same causal roles, which contradicts the conceptual functionalist's assumption that psychological states are *defined* by their causal roles.

Furthermore, Nida-Rümelin argues that the situation is similar even with a weaker form of functionalism called 'psychofunctionalism', but we don't need to go too far into details here: in the following, I shall concentrate on conceptual functionalism.[10] The crucial point should be clear enough, then: the mere conceivability of the (strong) inversion hypothesis seems to be a powerful argument against this well established philosophical theory.

However, hard-boiled advocates of conceptual functionalism could admit that their doctrine is incompatible with the (strong) inversion hypothesis but insist that this does not yet settle the question if this is bad for functionalism or for the inversion hypothesis. Why should we *presuppose* that the inversion hypothesis is conceptually possible? Couldn't we, on the contrary, conclude that this hypothesis finally turns out as conceptually impossible or incoherent, although it seemed coherent at first glance? After all, we found out that its conceptual possibility is incompatible with *functionalism*, our favorite and highly sophisticated theory of mind.

There seems to be an obvious reply to this objection. Functionalism is a *philosophical* doctrine, and the inversion hypothesis is an assumption of color vision *science*; and if there is a conflict between philosophy and science, it generally seems pretty clear that science will win. Science takes priority over philosophy, as it were. Accordingly, Nida-Rümelin supports the following principle:

> "*No hypotheses accepted or seriously considered in color vision science should be regarded according to a philosophical theory to be either incoherent or unstatable or false.*"
> (Nida Rümelin (1996: 145); italics original)

I consider the intuition underlying this principle to be quite all right, this wording carrying it too far, though.

First of all, it is certainly true that philosophy cannot ignore scientific results. For instance, modern quantum physics reveals that some processes are not causally dependent on other processes; this is a very good reason to doubt about the Kantian position that we know *a priori* that everything in nature is causally dependent on some other process. Or remember that Aristotle thought that primitive organisms could develop from mud - a position that has definitely been refuted by biology.

However, the latter case is not a very good example, as you already might have noticed. It's true that Aristotle was a philosopher (no doubt one of the greatest philosophers of the western hemisphere), and that he claimed that primitive organisms could develop from mud, but that does *not* make the assumption that primitive organisms can develop from mud a *philosophical* thesis. At least according to our present view, this is a biological assumption - and therefore it is not very surprising that refuting it is a matter of biological research.

Obviously, not everything a philosopher claims is a philosophical thesis - and an analogous fact is true for scientists: not everything a scientist claims is a (purely) scientific hypothesis, *even* if the claim has been developed in the context of scientific research. There may very well be hidden quite a bit of *implicit philosophy* in what a scientist asserts or considers. In this respect, Nida-Rümelin's principle seems too strong to me. Not *every* hypothesis accepted or seriously considered in color vision science is philosophically sacrosanct. In particular, this is not true for those hypotheses which contain or presuppose implicit philosophical assumptions. They can be - and actually they *should* be - objects of philosophical analysis and criticism. Accordingly, I propose a modified principle:

No *purely scientific* hypotheses accepted or seriously considered in color vision science should be regarded according to a philosophical theory to be either incoherent or unstatable or false; but hypotheses containing or presupposing philosophical assumptions should be objects of philosophical analysis and criticism.

However, sometimes it is not at all easy to tell whether a given hypothesis is "purely scientific" or loaded with philosophical assumptions. There are certainly lots of unambiguous paradigm cases, but there are also some highly problematic cases which need detailed consideration. I think it's an essential task of the dialogue between scientists and philosophers to find out which hypotheses contain or presuppose philosophical assumptions, i. e. to make the scientists' implicit philosophical intuitions explicit and open for criticism. Not only the philosophers would profit from this; such an interdisciplinary discussion can be just as fruitful for the scientists. In particular, they can become more conscious about the *conceptual frame* of their theories and might even consider its modification.

In our present case, there actually seems to be some evidence that scientists themselves are aware of the somewhat peculiar character of the hypothesis in question, i. e. the inversion hypothesis. According to Robert Boynton, pseudonormal people "would be expected to have normal color vision except that the sensations of red and green would be reversed - something that would be difficult, *if not impossible*, to prove".[11] Why could such a proof be *impossible*? Not because the empirical methods have not yet been developed, I guess; that would make it difficult, but not impossible. I rather think that Boynton doubts that the inversion hypothesis really belongs to the realm of pure empirical science at all. Perhaps what he wants to indicate is that, given the empirical results, the decision about the inversion hypothesis might be a matter of philosophy, at least in part. And that in my opinion is what holds true.

4. Pseudonormal Vision: Different Philosophical Interpretations

There seems to be a quite straightforward objection to this assumption. After all, the inversion hypothesis seems to *follow* from a well established *scientific* theory of color vision combined with some hypotheses about genetics, as we saw in section 2: how could a conclusion from scientific results be a matter of philosophical decision?

Well, there is one important aspect which was left out in the description of scientific results and hypotheses given in section 2 (as well as in Nida-Rümelin's paper): the described neural processes of course only lead to sensations of red or green if they are embedded in the far bigger neural system of the brain or the whole nervous system; perhaps even the rest of the body will play a part here, too[12]. Imagine the neural system described above were 'cut off' from the rest of the brain, somehow being kept alive and working: in this case, there would certainly never be any sensation of red or green, even if the photoreceptor cells in a given area were activated by the right kind of light. Obviously, the step from level 4 to level 5 in figure 1 will only be carried out if the signals from the nerve cells comparing the average activation of G- and R-cones in a given area are processed (at least) by the appropriate parts of the rest of the brain.

To be a bit more precise: What contemporary empirical color vision theory says is that *if* the processes explained at levels 1-4 are embedded in the neural structures of a *normal brain* (and perhaps in the system of a normal body as a whole), *then* the relationship described at level 5 holds true. This presupposition is so self-evident that it is hardly ever mentioned in color vision science, and so it is small wonder that Nida-Rümelin does not go into it in more detail.

But now notice that relevant parts of the brain of a pseudonormal person p_1 must *differ* from those of a normally sighted human being p_2 if they are assumed to produce different reactions to the same 'input' and the same reaction to different inputs. Suppose p_1 and p_2 both look at a ripe tomato. In p_1, the average activation of G-cones in the respective area exceeds that of the R-cones, whereas in p_2 the average activation of the R-cones exceeds that of the G-cones. So the inputs at level 4 are different. But they are both disposed to utter sentences like "This tomato is red", "This tomato looks red to me" etc., so that the outputs of the further processing are the same. This may also be true of the emotional reactions being caused by the sight of the color red. It follows that the processing in p_1 must differ from that in p_2. Suppose further that p_1 is looking at a ripe tomato, whereas p_2 is looking at an unripe (green) one. The relevant input at level 4 is the same, but the behavioral output is different: p_1 will be disposed to say things like "This tomato is red", whereas p_2 will say that is looks green to him. Again, it follows that the processing in p_1 must differ from that in p_2.

Because of these differences, we can say that the brains of pseudonormal people behaving in the same way as normally sighted persons must also be 'pseudonormal' with respect to the 'higher' processing of signals stemming from the visual system explained at levels 1-4.

But if this is true, it turns out that the inversion hypothesis does *not follow* immediately from color vision theory and genetics. This conclusion would only be correct if the relationship described at level 5 were true for pseudonormal people, too. But we cannot take that for granted, because standard color vision theory only maintains that this relationship is true *provided* that the processes described at levels 1-4 are embedded in a normal brain; and in pseudonormal people these processes are embedded in a pseudonormal brain. After all, if the higher processing in pseudonormal brains somehow manages to compensate the differences at the lower level visual system of pseudonormal persons so that their functional relations between the sensory input, other mental states, and the behavioral output do not differ from that of normally sighted people, why not suppose that this compensation also concerns the way things look like? Perhaps the functional compensation affects the sensory qualities as well, so that pseudonormal people see red tomatoes in exactly the same way as normal people do, although the lower level processings in their respective visual systems differ from each other decisively.

At first glance, it may seem that this latter assumption is a purely empirical hypothesis. Apparently, there are two possibilities. Either the suitable compensation is performed *before* the respective sensations become conscious; in this case, the inversion hypothesis would turn out as false: pseudonormal people would have the same sensations as normally sighted persons when looking at green or red things. Or the compensation is carried out *after* the sensation has become conscious; then the inversion hypothesis seems correct. So even if the inversion hypothesis does not follow immediately from the assumptions of standard color vision theory combined with genetics, it seems at least possible that further scientific investigation might confirm it. If this is true, the inversion hypothesis must still be regarded as a purely scientific and (therefore) conceptually possible assumption, so that the situation for *conceptual* functionalism has not become much better yet. The different cases can be represented in the following figures (parallel arrows symbolize "normal", crossed arrows deviant processing):

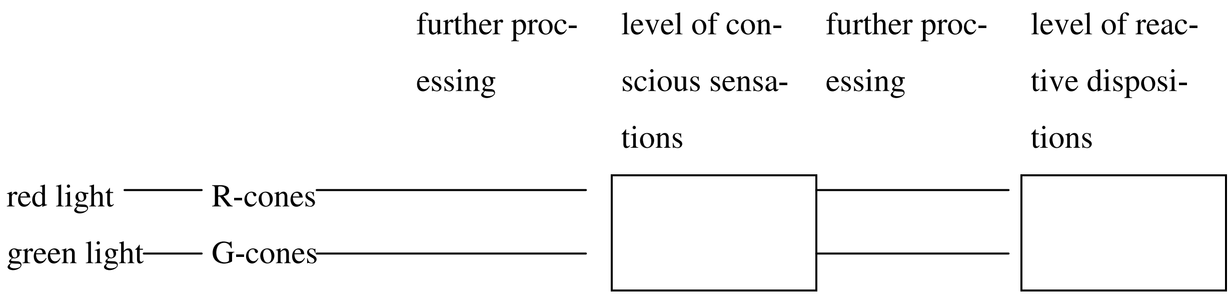

Fig. 2: normally sighted persons

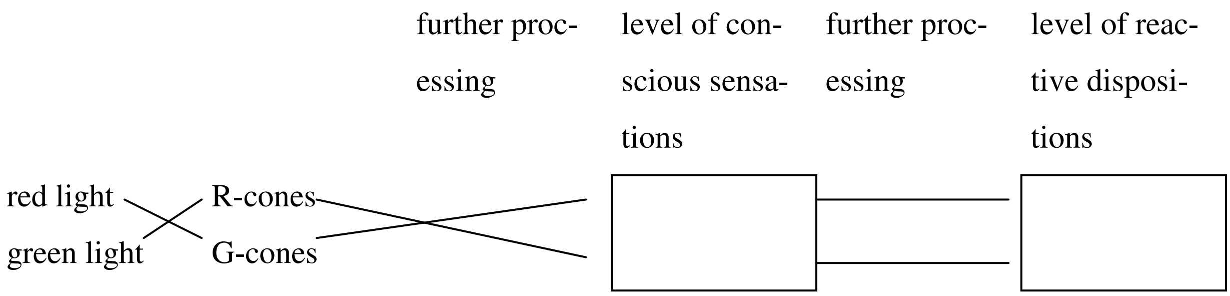

Fig. 3: pseudonormal persons, first possibility (normal sensations)

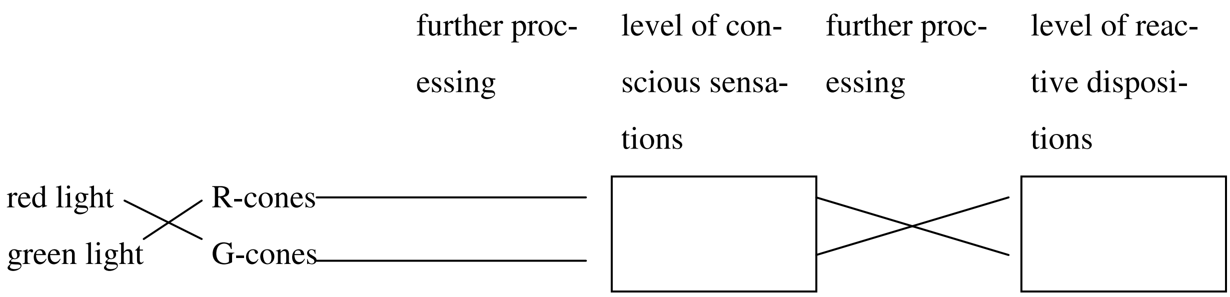

Fig. 4: pseudonormal persons, second possibility (inverted sensations)

However, this position presupposes that there is a clear-cut boundary between the levels of conscious sensations and reactive dispositions, an assumption having come under heavy attack during the last few years, in particular since Dennett (1991). In his book, Dennett criticizes what he calls the 'Cartesian Theater' theory of consciousness, according to which all sensory inputs of a given moment come together at a certain area of the brain (the Cartesian Theater) where a conscious 'picture' of the world is formed from them, this picture then giving rise to later reactive dispositions. The position I am presently considering does of course not necessarily presuppose that all sensory inputs come together at some defined area in the brain, i. e. it does not presuppose a complete Cartesian Theater, but it depends crucially on the assumption that there are at least some Cartesian Theater Workshops: even if there is no place "where it all comes together", there must at least be a certain point for *each* sensation at which it becomes conscious, and this point must be separated from all reactive dispositions.

Dennett raises a lot of empirical issues in his book, but I think that one main point is essentially philosophical. Prior to the question whether there *is* a clear-cut boundary between the levels of conscious sensations and reactive dispositions, one could ask if there *can possibly be* such a boundary. A positive answer would amount to the thesis that there is a sensible concept of *pure sensation*, i. e. to the assumption that one can sensibly imagine a sensation existing without any connection to reactive dispositions: after all, if the existence of a sensation were *necessarily* connected to the existence of a respective reactive disposition, one could never exclude the possibility that the disposition has an influence on the qualitative character of the sensation. However, this is a highly controversial point in philosophy. To quote an authority from the past, Kant seems to deny the existence of pure sensations by saying that sensations without concepts are "blind", i. e. that a sensory input must at least be subsumed under some (perhaps very general) concept in order to become a conscious sensation. The more recent discussion concerning the relation between sensations and reactive dispositions (including the disposition to subsume a sensory input under a concept) is very complex, and I don't want to go too far into detail here. My intention is just to emphasize that many authors discussing this matter bring forward *philosophical* arguments for their position.[13] And indeed it's hard to see how science alone could settle the question. The trouble is that conscious sensations or experiences cannot be *seen* or immediately and objectively identified by (e. g.) neurophysiological methods. In order to find correlations between neurophysiological processes and conscious experience, scientists must rely on neurophysiological *and behavioral* data about conscious beings, especially on their verbal reports, which are *always open for interpretation*.

As an example, think of the experiments with image-inverting glasses reported in 'Consciousness Explained'. Such glasses invert the way people see the world, i. e. they make people see things upside down; further, things on the left appear on the right, and vice versa:

> "When the adaptations of the subjects wearing these goggles have become so second nature that they can ride bicycles and ski, the natural (but misguided) question to ask is this: Have they adapted *by turning their experiential world back right side up*, or *by getting used to their experiential world being upside down*? And what do they say? They say different things, which correlate roughly with how complete their adaptation was. The more complete it was, the more the subjects dismiss the question as improper or unanswerable."
> (Dennett (1991: 397); italics original.)

This nicely fits in to the position that there is no sharp boundary between sensations and reactive dispositions: obviously, sensations and reactive dispositions change simultaneously during the process of adaptation. But now imagine that some of the test subjects answer like this after complete behavioral adaptation: "I have absolutely no problem in telling what is up and what is down. If an object A is located over an object B, I usually do not feel even the slightest inclination to think that B is over A. I immediately recognize that A is over B, without employing any conscious thought like 'It *seems* to me that B is over A, so in reality A must be over B'. That was only necessary at the beginning of the experiment. But in spite of all that, my visual sensations are still turned upside down."

Those who believe in (the possibility of) a sharp boundary between sensations and reactive dispositions could take this case as an example for an adaptive process *behind* or after the level of conscious sensation, but their opponents need not get embarrassed either: They could retort that in their opinion these reports are inconsistent and try to find out what it is that confuses the subjects. I don't think that this controversy could be settled by empirical research alone: both groups could agree about all empirical (i. e. behavioral, neurophysiological etc.) facts and yet disagree as to how to interpret them with regard to what the subject really *feels*. This is a matter of philosophical discussion: if one thinks that the concept of pure sensation is coherent, it is at least a sensible hypothesis to assume that the case in question is an example for adaptation after or behind the level of experience; if one belongs to the opposite philosophical camp, this hypothesis is not even conceivable.

The same point can be made with regard to the inversion hypothesis. If you find it conceptually possible to presuppose a clear-cut boundary between the levels of conscious experience and reactive dispositions (i. e. that the concept of pure sensation is coherent), then it makes sense to suppose that perhaps the necessary adaptations in a pseudonormal brain are carried out *after* or behind the level of experience, in which case the inversion hypothesis would be true. If you don't believe in the possibility of such a boundary, this hypothesis will not even seem coherent to you.

So, in the end, pseudonormal vision is not a matter of refuting philosophical positions by scientific results. In fact, the scientific hypothesis that there may be pseudonormal people allows for different philosophical interpretations, these interpretations being dependent on more basic philosophical presuppositions. If you are influenced by conceptual functionalism, you won't believe in pure sensations, since sensations are, like all other mental states or events, supposed to be *defined* by their causal relations to sensory inputs, other mental states, and behavioral outputs. There would be no point in postulating a mental state *isolated* from its causal connections. Accordingly, the inversion hypothesis would not even seem coherent to you. You would probably describe pseudonormal vision as just *another (physiological) way of having sensations of red or green*. Pseudonormal people have the very same sensations as normally sighted persons, but their *neural correlates* of these sensations differ from ours.[14]

On the other hand, if you prefer a philosophical position allowing for the possibility of pure sensations, you may conclude that there is good reason to suppose that pseudonormal persons have inverted experiences with regard to red and green, even if they turn out to be functionally indistinguishable from normally sighted people. In this case, you may ague that further empirical research is necessary to settle the matter.

This last point indicates that basic philosophical assumptions do not only have a crucial influence on the interpretation of scientific results, but may also give rise to particular research programs. Whereas those believing in the meaningfulness of the concept of pure sensation may find it promising to search for places in the brain which can possibly be considered as boundaries between the level of sensation and the level of reactive dispositions, others will consider this kind of enterprise as mistaken from the very start.

Note that this does *not* mean that philosophical preferences make preliminary decisions about *scientific* results. The *purely scientific* results of both research programs, i. e. those results concerning the pure neurophysiological structures and processes, the behavioral data etc., should be the same or at least not contradict each other, even if the respective scientists came to their conclusions by different ways, guided by different theoretical presuppositions. Again, the crucial point is a matter of *interpreting* the results philosophically. Even if scientists manage to discover a neurophysiological structure which *seems* to present itself as a suitable place for a boundary of the kind in question[15], one need not necessarily agree that this actually *is* such a boundary. As mentioned before, the problem is that all you can *immediately* investigate by scientific methods is neurophysiological structures and behavioral data, but not conscious experience: you simply can't *see* conscious experience, you just have it. And if you draw conclusions from scientific data with regard to conscious experience, there will always be some room for philosophically inspired interpretation.

So finally it turns out that the existence or the possibility of qualia inversion is still not a (purely) scientific, but a philosophical matter, albeit a matter of *philosophical interpretation of scientific results* as far as pseudonormal persons are concerned. I think the case of pseudonormal vision is an interesting lesson for scientists as well as for philosophers. Scientists may become more conscious of the implicit philosophical presuppositions their conclusions may contain, whilst philosophers may become more cautious when they try to refer to the authority of science in order to support their philosophical positions.[16]

References

Boynton, R. M.: 1979, Human Color Vision, New York et al., Holt Rinehart and Winston.

Damasio, A. R.: 1994, Descartes' Error. Emotion, Reason, and the Human Brain, New York, G. P. Putnam's Son.

Dennett, D. C.: 1991, Consciousness Explained, Boston, et al., Little, Brown and Company.

Dretske, F. 1995, Naturalizing the Mind, Cambridge/Mass., London, The MIT Press.

Hardin, L.: 1996, 'Reinverting the Spectrum', M. Carrier, P. Machamer (eds.), Mindscapes: Philosophy, Science, and the Mind. Pittsburgh, University of Pittsburgh Press/Konstanz, Universitätsverlag Konstanz, 99-112.

Knodel, H. et al.: 1983, Linder - Biologie, Stuttgart, J. B. Metzlersche Verlagsbuchhandlung.

Nida-Rümelin, M: 1995, 'What Mary Couldn't Know: Belief about Phenomenal States', Th. Metzinger (ed.), Conscious Experience, Paderborn, Schöningh/Imprint Academic, 219-241.

Nida-Rümelin, M.: 1996, 'Pseudonormal Vision - An Actual Case of Qualia Inversion?', Philosophical Studies 82, 145-157.

Nida-Rümelin, M.: 1997, 'The Character of Color Predicates: A Phenomenalist View', W. Künne, A. Newen, M. Anduschus (eds.), Direct Reference, Indexicality, and Propositional Attitudes, Stanford, CSLI, p. 381-402.

Nida-Rümelin, M.: 1998, 'Vertauschte Sinnesqualitäten', F. Esken, D. Heckmann (eds.), Bewußtsein und Repräsentation, Paderborn, Schöningh, 299-324.

Spohn, W.: 'The Character of Color Predicates: A Materialist View', W. Künne, A. Newen, M. Anduschus (eds.), Direct Reference, Indexicality, and Propositional Attitudes, Stanford, CSLI, 351-379.

Tye, M.: 1995, Ten Problems of Consciousness, Cambridge/Mass., London The MIT Press.

[1] M. Nida-Rümelin (1996).

[2] See Nida-Rümelin (1996: 146).

[3] For clarification, note that there are *two* essential hypotheses involved here. The first one claims that some persons are pseudonormal in the sense that their visual system differs in a certain way from ours, whereas their behavior concerning colors does not differ very much or perhaps not at all from that of normally sighted people. The second hypothesis claims that the

experiences of pseudonormal people with regard to red and green are systematically inverted in comparison with normally sighted persons. This latter assumption may be called the 'inversion hypothesis'.

[4] See Nida-Rümelin (1995: 231).

[5] See Nida-Rümelin (1997) and Spohn (1997).

[6] See for example H. Knodel et al. (1983: 215).

[7] See for example Boynton (1979: 351-358).

[8] Nida-Rümelin (1996: 148).

[9] According to the stronger version, the phenomenal experiences of pseudonormal people will be systematically inverted even if they behave like normally sighted persons *in every respect*. In this latter case, the differences between normally sighted people and pseudonormal persons could not even be revealed by ingenious psychological experiments, but only by direct physiological investigation of their retinas. Of course, it is an interesting empirical question if the behavior of pseudonormal people (given that there really are any) is completely identical with that of normally sighted persons or if there are slight but systematic differences between their respective behavioral dispositions concerning colors. For instance, they may differ from each other with respect to some emotional reactions. Further on, one could make use of some well established results of color vision theory about particular phenomenal properties of certain colors to develop hypotheses about possible behavioral differences. For example, orange changes its phenomenal appearance dramatically under particular lighting conditions: it appears brown then. In contrast, turquoise does not change its appearance in this way. (Cf. Hardin (1996: 105 f.) or Nida-Rümelin (1998: 316 f.).) In the case of pseudonormal persons, it seems likely that orange light affects the visual system in a way turquoise light affects normal visual systems, and vice versa. *If* this difference is not compensated somehow, it seems likely that pseudonormal persons will report that turquoise and not orange changes its phenomenal character under the appropriate circumstances. However, in this case pseudonormal vision would not pose a problem for functionalism, since the differences in phenomenal experiences would be accompanied by differences in behavioral dispositions. It would of course be very interesting to consider pseudonormal vision in the light of Hardin's more general arguments against the possibility of inverted spectra, but that would require another paper.

[10] A defense of the Wittgensteinian point of view against the weaker inversion hypothesis would require a different kind of argument.

[11] Boynton (1979: 356). Here quoted from Nida-Rümelin (1996: 145); my italics.

[12] Damasio's (1994) research on emotions suggests that the body plays a crucial part in producing emotional reactions; if such reactions turned out to be of some importance for the sensation of colors, certain bodily states could perhaps have an influence on how colors are experienced.

[13] See for instance the relevant passages in Tye (1995) or Dretske (1995).

[14] Note that this interpretation of pseudonormal vision is not exposed to the criticism of possible functionalist analyses put forward by Nida-Rümelin (1996: 152 f.). In particular, this proposal does not imply a functional redefinition of morphologically defined G- and R-cones, nor does it "violate the widely accepted principle of supervenience for mental properties upon the *relevant* physiological properties" (1996: 153, my italics): the latter principle can be maintained if it is acknowledged that more physiological properties are *relevant* for color vision than those mentioned in figure 1. But this can hardly be denied, as I argued at the beginning of this section.

[15] Presently, neurophysiologists seem to have the impression that the search for sharp boundaries in the brain is not very promising. Apparently, empirical research has not yet led to the discovery of suitable places which could possibly be interpreted as boundaries of the kind in question.

[16] This paper was supported by a grant from the *Deutsche Forschungsgemeinschaft*.