[32] Tarski's contention that the English truth-predicate is "contradictory" could raise some doubts in this regard. This is another reason not to follow him on this matter as I suggested that we should not in section II above.
My disagreement with Etchemendy on the significance of Tarskian truth definitions accounts also for my disagreement with him on his dim evaluation of Tarskian, model-theoretic, analyses of the logical properties – *logical truth* and *logical consequence* – in his *The Concept of Logical Consequence*, Cambridge, Mass.: Harvard University Press, 1990. See my "A Defense of the Model-Theoretic Account of the Logical Properties," *Notre-Dame Journal of Formal Logic*, 34 (1993): 107–131.
[33] Scott Soames, op. cit., 419f.
[34] See Hartry Field, "The Deflationary Conception of Truth," in G. MacDonald and C. Wright (eds.), *Fact, Science and Morality*, Oxford: Basil Blackwell, 1986.
[35] Oxford: Basil Blackwell, 1990.
[36] Actually, Paul Horwich told me in personal conversation that he intended the axioms as specified by putting the same sentence in both 'p' positions. This is in fact to say that the axioms are not propositions, but propositions "under a linguistic guise."
[37] I owe some of the points in the last two paragraphs to my colleague, Ignacio Jané.

Departamento de Lógica, Historia y Filosofía de la Ciencia,
Universidad de Barcelona
08028 Barcelona
Spain

MARTINE NIDA-RÜMELIN

PSEUDONORMAL VISION

An Actual Case of Qualia Inversion?

(Received in revised form 9 October 1993)

1. INTRODUCTION

Is it possible that a person who behaves just like you and me in normal life situations and applies colour words to objects just as we do and makes the same colour discriminations and colour similarity judgements that we make, see green where we see red and red where we see green? Many philosophers assert that the description of such a case is somehow incoherent. Often the motivation for this assertion is "that they suspect that admitting that claim [the possibility of such a case] will put one on a slippery slope which will eventually land one in skepticism about other minds".[1]

Among philosophers, however, it does not seem to be common knowledge that there is scientific evidence for the existence of such cases. Theories about the physiological basis of colour vision deficiencies together with theories about the genetics of colour vision deficiencies lead to the prediction that some people are 'pseudonormal' (according to an estimation of Piantanida (1974) this occurs in around 14 of 10 000 males).[2] Pseudonormal people "would be expected to have normal colour vision except that the sensations of red and green would be reversed – something that would be difficult, if not impossible, to prove."[3]

Any philosophical theory of mind or more specifically about colour, colour appearances or colour concepts should meet the following plausible prima facie constraint: *No hypotheses accepted or seriously considered in colour vision science should be regarded according to a philosophical theory to be either incoherent or unstatable or false.* Therefore – regardless of whether the hypothesis of the existence of pseudonormal people is correct – the mere fact that the hypothesis is seriously considered in colour vision science, is philo-

Philosophical Studies **82:** 145–157, 1996.

sophically relevant. Central claims of colour vision science when combined with specific empirical assumptions lead to the prediction that there are red-green-inverted people. Therefore any philosophical theory which excludes such a case does not meet the above formulated constraint. The failure to meet this prima facie constraint does not in itself justify the rejection of a philosophical proposal, but it does represent a serious objection. This kind of criticism will be advanced against some widely held philosophical proposals in the present paper. But let me begin with a short sketch of the relevant parts of colour vision science.

2. PSEUDONORMAL VISION. THE SCIENTIFIC BACKGROUND

There are three types of photoreceptors on the retina that play a central role in human colour vision (B-, G- and R-cones). They are morphologically distinguishable, they play different roles in colour information processing and they normally contain three chemically different photopigments. For each cone type there is a characteristic function (the so-called sensitivity curve) which describes how the level of stimulation caused by monochromatic light in a cone of the given type depends on the wavelengths of the light at a given intensity level. It is assumed that the sensitivity curves are determined by the absorption spectra of the pigments contained in the receptors. The expected level of stimulation of a cone caused by non-monochromatic light (which is the normal case) can be calculated on the basis of the sensitivity curve characteristic for its type. When light reaches a given area on the retina, then some neural mechanism will calculate the average stimulation of the cones in the area of any of the three types. The average stimulation of the three cone types is then compared and information about the results is carried to the brain by two neural channels, the r-g-channel (responsible for red- and green-sensations) and the y-b-channel (responsible for blue- and yellow-sensations). If (b, g, r) represents the average stimulation of the B-, G- and R-cones in the area at issue then how the channel states depend on the average stimulation (b, g, r) of the three cone types can be represented (according to a simple model of so-called opponent process theory presented in Boynton (1979)) by the following two functions.

(1) $C1((b, g, r)) = r - g$

(2) $C2((b, g, r)) = g + r - b$

It is assumed that the amount of greenness, yellowness, blueness and redness experienced by an observer in a concrete case can be predicted on the basis of the values of C1 and C2. In case C1 ((b, g, r)) = O, the corresponding object will appear neither greenish nor reddish to the person. At the zero-point of the second function, there will be no blue or yellow component in the perceived colour. For positive values of C1, the person does not sense any greenness and the amount of redness increases with the distance from zero. With negative values of C1, the person does not sense any redness, and the amount of greenness increases with the distance from zero. Analogously yellow-sensations are correlated with values of C2 greater than zero and blueness-sensations with values of C2 smaller than zero.

According to the prevailing theory about red-green blind vision these people differ from normal people in the following respect: their G-cones and R-cones contain the same photopigment. Therefore the average stimulation of their R- and G-cones will be equal for any light stimulus. The value of C1 consequently will always be zero and it follows from the theory that nothing will appear reddish or greenish to the subject. One group of red-green-blind people (so-called protanopes) have the photopigment normally contained in the G-cones not only in their G-cones, but also in their R-cones. For the other group of red-green blind people the reverse is true: their G-cones and their R-cones both contain the photopigment normally contained in R-cones. According to a widely accepted model of the inheritance of colour vision defects, both genes, the one that causes production of the G-cone photopigment in R-cones and the one that causes production of the R-cone photopigment in G-cones, may be active simultaneously in one single individual. In these cases the photopigments of the two cone types at issue are simply exchanged. The result should be a person which does not have any obvious colour vision defect. These people are called pseudonormal since they appear to be normal but really are not. To any light stimulus their R-cones react like normally filled G-cones and their G-cones react like normally filled R-cones. The reversed filling of cones with

photopigments only affects the causal interconnections between external stimuli and cone type activation. It does not, however, affect the causal interconnections between cone type activation and the states of the two chromatic channels. This second causal dependency is therefore assumed not to be altered in pseudonormal people. It follows that any light stimulus which causes the r-g-channel of a normal person to have the value y, will cause the r-g-channel of a pseudonormal person to have the value -y. If y corresponds to a reddish component in the perceived colour, then -y corresponds to a greenish component in the perceived colour (and vice versa). It therefore follows from received scientific theory about human colour vision, that pseudonormal people, if they exist, are red-green-inverted in the following sense: things that appear reddish to normal people to a certain degree, appear greenish to pseudonormal people to roughly the same degree (and vice versa) while the perception of yellowish or bluish components remains unaffected.[4]

3. PHILOSOPHICAL CONSEQUENCES

3.1. *A Problem for Wittgensteinians*

Let us call an N-case a case where a person P is red-green-inverted and yet there is no behavioural difference between P and normally sighted people detectable in normal life situations that would give any reason to suspect that P's colour perceptions differ from those of normal people. Some of those philosophers who are influenced by Wittgensteinian ideas think that the possibility of an N-case can be excluded without empirical research on the basis of philosophical considerations alone. They would subscribe to the following view: Ripe 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. These rules do not require physiological examination of someone's visual system. Pseudonormal people are expected by colour vision science to behave roughly like normal people do in colour discrimination and colour judgement and therefore the conditions meant by Wittgensteinians for an appropriate ascription of normal colour perception are certainly fulfilled. So it seems that the Wittgensteinian must deny that pseudonormal people

are red-green inverted and finds himself in conflict with what colour vision science asserts.

The Wittgensteinian however might defend his view claiming that the rules governing the use of colour appearance concepts in normal language are different from those governing scientific usage of these terms. He might then adopt one of the two following slightly different strategies: a) he might say that philosophy is concerned with everyday language and therefore need not care about how colour vision science describes the phenomenology of pseudonormal vision or b) he might admit that given the results of colour vision science sketched above we have reason to change the rules of the game and adopt the view that pseudonormal people are red-green inverted. In order to argue against both defence strategies it is necessary to show that colour vision science when using colour appearance terms does not introduce new concepts but rather uses these terms in their normal way. This indeed seems quite obvious, but to argue for this claim is a more complicated task which cannot be completed in the present paper.[5]

3.2. *Pseudonormal Vision and Functionalism*

It has been objected to functionalism that there could be what I will call an F-case. An F-case is a case where there is no relevant functional difference between a person P and normally sighted people although P is red-green-inverted.[6] Before we can begin discussing whether pseudonormals represent an F-case we need to distinguish different senses of "functional difference" and thereby different versions of functionalism.

Conceptual functionalism claims that the meaning of mental terms may be analysed in functionalist terminology. According to conceptual functionalism to see something as red means to be in a state which plays a specific causal role. This causal role, according to conceptual functionalism, can be specified by reference to a) typical causes of the state and b) typical causal influence of the state at issue upon other mental states. The proponent of conceptual functionalism therefore must deny the possibility of what I will call an F1-case: An F1-case would be realized if for a red-green-inverted person P something like the following two conditions hold: a) P does not differ from normals with respect to colour naming and colour

discrimination behaviour and b) if there is a specific difference in the role red- and green-sensations play in connection with emotions, other modes of perception, space perception and the like, then these roles in the case of person P are reversed too. For N-cases it was required that the difference between P and normally sighted people could not be detected in normal life. It is required in addition for F1-cases that P will behave like a normal person even in sophisticated psychological and psychophysiological experiments. To the proponent of conceptual functionalism, we may ascribe the view that F1-cases are incoherent. To reject conceptual functionalism it is not necessary to show that pseudonormal people represent F1-cases. It suffices to argue that according to received colour vision science the question whether they do represent F1-cases or whether they do not needs to be settled by empirical research. This is enough since no hypotheses seriously considered in scientific theory should be regarded incoherent by any philosophical proposal. It has already been shown that pseudonormals, if they exist, are red-green inverted according to scientific theory. Whether they do represent an F1-case therefore only depends on the answer to the following question: Are there differences between red-sensations and green-sensations with respect to their causal influence upon other mental states which are innate and will not be overridden by learning processes? If the answer is 'yes', then pseudonormal people cannot represent F1-cases and could be detected by sophisticated psychological experiments. If the answer is 'no', then pseudonormal people could not be detected without direct investigation of their retina and they would represent F1-cases. Obviously the question needs to be settled by empirical research, and conceptual functionalism thus violates the above formulated prima facie constraint.[7]

Psychofunctionalism does not claim to give an analysis of the meaning of mental terms. Rather it proposes to accept the *empirical* hypothesis that mental terms will turn out to refer to functionally definable internal states. Colour vision science asserts that red sensations occur when the relevant r-g channel is in a specific type of state which is represented by positive values of C1. Let us call this type of state 'positive r-g channel state', analogously I will talk of negative r-g channel states. The question of whether psychofunctionalism violates the above formulated prima facie constraint then

depends on whether the difference between positive and negative r-g channel states is a functional difference in the sense of psychofunctionalism. At first sight it seems that it is not: positive and negative r-g channel states can be distinguished by reference to their causes. Positive states are caused by a predominance of R-cone activity, while negative states are caused by a predominance of G-cone activity. But this is a functionally describable difference only if the two types of cones can be functionally defined. The most obvious way to define receptor types in the present context is by reference to the way they react to light stimuli. This strategy however is not available here. R-cones can be filled with the photopigment normally contained in G-cones and thereby be caused to behave like G-cones. But, as the hypothesis of pseudonormal vision shows, colour vision science explicitly denies that a G-cone filled with the wrong pigment thereby ceases to be a G-cone. Of course there is a difference in causal role between G-cones and R-cones: They have different influences upon the channel states. But this is what we started with. So the difference between positive and negative r-g channel states cannot be functionally specified by reference to the way these states are caused. Still, the psychofunctionalist may hope that the two channel states will turn out to play different functional roles on higher levels of information processing.

I have characterized psychofunctionalism by the empirical hypothesis that terms for mental states will turn out to refer to functionally definable states. A somehow stronger claim is however in the spirit of psychofunctionalism. Those who subscribe to some kind of psychofunctionalism certainly would have expected any theory of colour information processing to be a functional theory from the outset. This would mean that the central notions of colour vision science at any of its historical stages should be explicable in functionalist terminology. This stronger claim, however, is quite obviously wrong.[8]

3.3. *Fixing the Reference of Physiological Concepts*

The *real* story about the development of colour vision science seems to be this: It is a central assumption of colour vision science which has been accepted from the very beginning of this empirical discipline and has turned out fruitful that for any of the four phenomenally

basic hues there must be some specific physiological process responsible for the occurrence of that colour sensation. (The assumption is hold true for the whole range of sighted people independently of their specific kind of – normal or abnormal – colour vision). One first step in the development of modern colour vision science was to *postulate* the existence of one type of physiological process responsible for every basic hue sensation and to assume that any of these four processes allows for degrees which are correlated with the corresponding amount of f-ness (where f is a basic hue) in the phenomenally given colour. *Thus the reference of physiological concepts was fixed in colour vision science by definite descriptions formulated using phenomenal concepts* (e.g. "the process p such that the 'degree of p' is correlated with the amount of redness"). It was assumed as a working hypothesis that these phenomenal descriptions are successful in picking out specific physiological types.

If this description is correct, then phenomenal concepts used in their everyday meaning did play and still do play an essential role in the development of scientific terminology. If this is true, then the psychofunctionalist who wishes to uphold what I called his stronger claim, needs to show that our phenomenal concepts really are functional concepts. He thus has to support, in addition, some kind of conceptual functionalism. Conceptual functionalist, however, has already been shown to be inadequate.[9]

4. TWO OBJECTIONS

Here is a possible objection that needs to be discussed: Someone might propose to redefine R-cones, G-cones and B-cones in terms of their corresponding spectral sensitivity curves. This indeed would cause the argument to break down. We then could not say of pseudonormal people that their cones contain the 'wrong' pigment, since, by containing erythrolabe and thereby a specific spectral sensitivity curve, a receptor *by definition* becomes an R-cone. This definition of cone-types, combined with a definition of the relevant states of the r-g-channel according to its causal relations to the cone-types, leads to the conclusion that in normal subjects and in pseudonormal subjects the same external conditions cause the same r-g-states. It would follow that red things appear red to pseudo-

normal people just as they do to normal ones. The philosopher proposing this redefinition might make his view still more difficult to attack by adding: My position does not need the assumption that the proposed definition is more adequate than a morphological individuation of cone types. It probably is a matter of practical convenience which definition should be preferred. Since it depends on what definition we choose whether opponent process theory predicts normal vision or inverted vision for pseudonormal subjects, the question whether an object appears red or green to a pseudonormal person turns out to be decidable by convention. It is then not a factual question about what really is the case. This result – the opponent might go on – is almost as good or even better than genuine impossibility of qualia inversion.

This counterargument can be met in two ways: First, redefining receptor types in the way proposed would by definition exclude specific cases of *acquired* red-green-inversion which seems quite unacceptable.[10] Second, the proposal violates the widely accepted principle of supervenience for mental properties upon the relevant physiological properties: Let us for the moment accept that the relevant states of the r-g-channel can be functionally defined in the way proposed, and let us call states represented by positive C1-values, positive states of the r-g-channel and the same for negative values. Since the neural hardware is not affected by exchanging photopigments, we must assume that the physiological state produced by a specific pattern of stimulation of concrete photoreceptors in a given person is the same regardless of whether the photopigments are reversed. The proposal, therefore, entails that the same physiological state that realizes a positive r-g-state, given normal distribution of photopigments, realizes a negative r-g-state, given pseudonormal distribution of photopigments. So the proposed definition, combined with opponent process theory, entails the prediction that the *very same* physiological state will lead to a red-sensation in the one case and to a green-sensation in the other. Since the only difference between the two cases lies in the way the physiological state is *caused* (by different patterns of light stimuli) and since the brain does not have any access to this information, this would seem rather mysterious.

David Lewis defended functionalism against the so-called Inverted Qualia Argument claiming that "object o looks red to person x" is ambiguous, needs to be relativized to a population P and means something like the following: "object o produces in x a state, which in people of population P plays the role of red-perceptions" where 'the role of red-perceptions' is assumed to be explicable in functional terms.[11] On this account the assumption that green things look red to pseudonormal people (in the sense in which it is true) would *mean*: when looking at grass the brain of pseudonormal people is in a physiological state which occupies the role of seeing something red in normally sighted people.[12] Lewis' proposal, however, yields an inadequate interpretation of the following central assumption in colour vision science: there is a specific physiological state which is responsible for red sensations in general (whoever is in that state has an experience of red and vice versa). This assumption is supposed to be true for all human beings in a non-trivial sense. This basic assumption, therefore, should not follow from the following 'weaker' claim: There is a physiological state (or process) which occupies a specific functional role F in normal subjects. On Lewis account, however, it does.[13] This argument does not in itself show that the proposal violates the above formulated constraint for philosophical theories, but it does prove the violation of another plausible necessary condition for an adequate philosophical theory: *If a hypothesis H which is accepted or seriously considered in some well-established scientific theory contains a concept C and if the philosophical theory proposes a definition of C, then replacing C by the proposed definiens should not change the empirical content of H.*

5. FINAL REMARK

The two constraints used in this paper only provide *prima facie* reasons for rejecting a given philosophical proposal. They may be overridden by philosophical considerations in some cases even if the scientific theory is empirically well-established. However, in such a case, the philosopher who wishes to reject scientific terminology, should be able to argue convincingly that the theory can be replaced by an alternative one, which does conform to the philosopher's intuitions and is yet in some relevant sense empirically equivalent to the original one.[14]

NOTES

[1] Sidney Shoemaker "The Inverted Spectrum", *The Journal of Philosophy* 79 (1982): 357–382, p. 364.

[2] See T. P. Piantanida "A replacement model of X-linked recessive colour vision defects." *Annals of Human Genetics* 37 (1974): 393–404 and Robert M. Boynton "Human Color Vision", New York et al. 1979, Holt Rinehart and Winston, p. 351–358.

[3] Boynton in "Human Color Vision" op. cit., p. 356.

[4] In an earlier paper and in my dissertation I discussed the case of photopigment exchange between R- and G-cones as an empirically possible but only imaginary case for which colour vision science would have to predict red-green inverted vision (see my "Irreduzibel mentale Prädikate in physiologischen Theorien der Farbwahrnehmung", *Berichte des Internationalen Wittgensteinsymposiums* 1988, Wien 1989 S. 59–62 and my "Farben und phänomenales Wissen", Conceptus Studienband 9, Wien 1993: Academia Verlag (St. Augustin)). Three years later I discovered that the imaginary case had actually appeared as a serious hypothesis in scientific literature. As far as I know pseudonormal vision has not yet been discussed in philosophical literature which might be due to the fact that the hypothesis can only be found in chapters or articles about the inheritance of colour vision deficiencies which philosophers might tend to skip.

[5] If pseudonormal people exist, then normal subjects are systematically wrong about the colour experiences of these people as long as they believe them to be normal. Both strategies discussed above would commit the Wittgensteinian to the view that prior to the development of modern colour vision science there was no such error.

[6] See e.g. Ned Block and Jerry Fodor, "What Mental States Are Not", *Philosophical Review* 81 (1972): 158–182, pp. 172–174.

[7] It violates the constraint in the following way: A hypothesis which according to colour vision science needs to be settled by empirical research (the hypothesis that there are F1-cases) is incoherent according to conceptual functionalism. The case shows that conceptual functionalism violates a further plausible prima facie constraint: No claim should be conceptually true according to a philosophical theory if it has to be settled by empirical research according to colour vision science. (Conceptual functionalism violates this further constraint with respect to the hypothesis that there are no F1-cases).

The argument against conceptual functionalism may also be put this way: According to received scientific theory (according to central claims of colour vision science plus the hypothesis of pseudonormal vision) there are F1-cases iff there are no innate differences in our reactions to red and green. Our colour concepts do not suffice to tell us that there are such innate differences. Therefore, contrary to conceptual functionalism, the existence of F1-cases cannot be excluded by conceptual considerations either. (I am grateful for a comment by Ned Block, which brought me to see this alternative way of making this point).

[8] It has been pointed out to me independently by Janet Levin and by Ned Block that the functionalist might reply claiming that normals and pseudonormals looking at a red thing simply have different *physiological realizations* of the same sensory quality. Of course, this is what some functionalists would like to say about the case. My point is that this description of pseudonormal vision (we would have to

say e.g. that red things look red to them) is in conflict with the way the case is described in color vision science. This reply, therefore, does not meet the above formulated prima facie constraint for philosophical theories: According to this account certain hypotheses accepted in colour vision science turn out false.

[9] It is just a historical observation which is not in need of any philosophical argument that phenomenal concepts ("sensations of blue", "sensations of yellow" etc.) were used to pick out physiological types in the way roughly described in the text. For my argument I do not need the stronger claim that these concepts as used in these contexts cannot be interpreted in a behaviourist or functionalist manner (although I am certain they cannot). My point is, rather, that given the above historical observation, the stronger claim of the psychofunctionalist fails unless it is combined with some kind of conceptual functionalism.

[10] Assume that someone's R-cones and G-cones start to produce the wrong photopigments at some point in his adult life. Colour vision science predicts such a person will experience and report a radical change in his colour perception. Who accepts the proposed redefinition of cone types and subscribes to opponent process theory, however, would have to insist that no such change has taken place: Those individual receptors that were R-cones before the inversion of photopigment distribution in the retina of the person at issue, *turned into G-cones* according to the proposed redefinition. Thus, green objects cause a predominance of G-cone-activity before the inversion *and* after the inversion. Therefore, the channel state produced by green things is a negative r-g-channel state before *and* after the change. So, according to the proposed redefinition, acquired photopigment inversion could not result in any change in the colour perceived by the subject.

[11] See David Lewis "Mad Pain and Martian Pain", in *Philosophy of Psychology*, Vol. I, ed. Ned Block (Cambridge: Harvard University Press, 1980): 216–222, p. 200.

[12] Assuming that pseudonormal and normal people are functionally equivalent in the relevant sense, on Lewis' account the following further assumptions hold: green things look green to pseudonormal people relative to the group of pseudonormal people, green things look red to normal people relative to the group of pseudonormal people, green things look green to normal people relative to their own group. Lewis's proposal, of course, should not be confused with the view that pseudonormal people and normal people simply refer to different subjective qualities when they use colour appearance concepts.

[13] The argument can be formulated more precisely:

(A) $\exists s \forall x (\langle s,x\rangle \epsilon \alpha \leftrightarrow R(x))$

(B) $\exists s \forall x (\langle s,x\rangle \epsilon \alpha \leftrightarrow \langle \text{ the } s' FR(s',P*),x\rangle \epsilon \alpha)$

(C) there is exactly one s such that FR(s,P*)

$\langle s,x\rangle \epsilon \alpha$: the brain of the person x is in the physiological state s

R(x) : x has a sensation of red

FR(s,P) : the state s occupies the functional role of seeing something as red in population P

P* : population of normally sighted people

The s ϕ[s] : the state s which satisfies ϕ. Quantifiers followed by s or s′ quantify over physiological states, quantifiers followed by x quantify over people.

(A) is the basic assumption at issue. (B) is the account of Lewis for this assumption (R(x) is replaced by the proposed definiens). (B), however, logically follows from (C) and therefore cannot be equivalent to (A) as meant in colour vision science.

[14] I have benefitted from discussions on this topic with Max Drömmer, Andreas Kemmerling, Martin Rechenauer and Wolfgang Spohn. I am very grateful to Ned Block for detailed criticisms of an earlier version of the paper. The work was supported by the grant Nr. Sp. 279/4-1 from the *Deutsche Forschungsgemeinschaft*. Special thanks is due to Edith Vanghelof who helped with linguistic corrections.

Institut für Philosophie
Logik und Wissenschaftstheorie der
Ludwig-Maximilian-Universität München
Ludwigstr. 31
80539 Munchen
Germany