

The Knowledge Argument, Abilities, and Metalinguistic Beliefs

Uwe Meyer
University of Osnabrück
Institute of Cognitive Science
D - 49069 Osnabrück, Germany
uwmeyer@uos.de
Draft

Published in: Erkenntnis 55 (2001), 325-347

<http://www.springerlink.com/content/mh55024n214v5213/fulltext.pdf>

Abstract

In this paper I discuss a variant of the knowledge argument which is based upon Frank Jackson's Mary thought experiment. Using this argument, Jackson tries to support the thesis that a purely physical - or, put generally: an objectively scientific - perspective upon the world excludes the important domain of 'phenomenal' facts, which are only accessible introspectively. Martine Nida-Rümelin has formulated the epistemological challenge behind the case of Mary especially clearly. I take her formulation of the problem as a starting-point and present a solution which is based solely on the concepts of capability and of metalinguistic beliefs. References to epiphenomenal facts, phenomenal knowledge etc. will be avoided completely. I specify my proposal against the backdrop of Burge's critical reflections about metalinguistic reinterpretation of expressions of belief and the externalist thesis held by Burge, Putnam and others that meanings and mental states are dependent upon the environment. My solution is then compared with Lewis' and Nemirow's ability objection. Finally I argue that the much discussed "knowing what it is like" has in its ordinary meaning nothing much to do with 'phenomenal knowledge' or knowledge of 'epiphenomenal' facts.

1. Jackson's Argument and Lewis' and Nemirow's Ability Objection

The origin of the whole discussion about the knowledge argument is - besides Nagel's bat¹ - the Mary-story in Jackson (1982):

Mary, a brilliant scientist, is able to find out all physical facts (in the widest sense) about colors and color-vision that there are to know, although she is imprisoned in a black-and-white room and has herself never seen colors. What will happen when she is released from her room one day and passes into a colored environment?

"Will she *learn* anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then it is inescapable that her previous knowledge was incomplete. But she had *all* the physical information. *Ergo* there is more to have than that, and Physicalism is false."²

This argument has several weaknesses. The most conspicuous (although seldom emphasized) is probably that it does not possess the argumentative form which is to be expected. Properly, it would be necessary to *show* that the - admittedly indistinct - ordinary methods or strategies of knowledge attribution *force* or at least make plausible the presumption that Mary acquires a knowledge of a special kind about facts which extend beyond the domain of the physical and are in this sense 'epiphenomenal' - a knowledge which following Nagel can be termed as "knowing what it is like".

Jackson clearly does not do this. Bring to mind, for instance, the phrase "It seems just obvious that she will learn something about the world and our visual experience of it". *Just obvious* this is doubtlessly only under certain philosophical assumptions. At heart, Jackson already has an epiphenomenal image of the Mary case and then describes this image with the help of epistemological terms. What he *ought* to show, however, is that ordinary and plausible methods of knowledge attribution compel us to accept the existence of epiphenomenal facts.³

Accordingly, several critics of Jackson are of the opinion that under philosophically unbiased scrutiny, Mary's case does not pose any true epistemological problems at all.

David Lewis and Laurence Nemirow hold the view that Mary does indeed learn something after her release; however this is not knowledge at all. What Mary does learn is rather a bundle of new abilities, a *know-how* instead of a *know that*⁴: She learns to imagine colors, to remember them, to recognize them, and to denote them correctly solely on the basis of their appearance when they are

¹Cf. Nagel (1974).

²Jackson (1982: 471).

³A few aspects of the argument are presented in somewhat more detail in Jackson (1986), but not in the proper form there either.

⁴Cf. Lewis (1983) and (1988), and Nemirow (1980) and (1990).

presented to her.⁵ This can be *described* by the phrase that Mary can know what it is like to see colors outside of her room only, but according to Lewis and Nemirow we should not conclude from this linguistic fact that any kind of propositional knowledge actually plays a part here. Instead, the 'knowledge' in question can be *identified* with the abilities to recognize colors, to imagine them, etc. In this way, Lewis and Nemirow do completely without epistemological concepts in their interpretation of the Mary case. They base their analysis wholly upon the concept of ability.

The suspicion that from a philosophically unbiased point of view it is not a matter of knowledge, but perhaps of certain abilities, certainly suggests itself as long as the impression exists that the knowledge 'argument' is basically just a kind of epistemological illustration of an image philosophically highly laden with assumptions. However, Martine Nida-Rümelin recently managed to work out a real epistemological problem based on a version of the Mary case which *as a problem* poses itself completely independently of any substantial philosophical assumptions. In my opinion, her considerations really do show that in the given context it is not possible to fully forgo an epistemic vocabulary or epistemological considerations; the bearing on abilities only is not sufficient. I do however not agree as much with Nida-Rümelin's method of solution as with her formulation of the problem. Actually, I agree with Lewis and Nemirow in that abilities play a fundamental role - if a slightly different one to the one these two authors think.

2. Marianna's Case: A Real Epistemological Problem

To begin with, let us have a closer look at Nida-Rümelin's argument. The starting point is a thought experiment, the central figure of which is Marianna: like Jackson's Mary, she grows up in a black-and-white room and does not ever see any colors. It is not presupposed here that Marianna is particularly well informed about the physiological basis of human color vision or any 'physical' aspects of colors. However, she has been told about the colors of certain objects: for instance, she has heard that the unclouded sky is blue. Before she is finally released from her black-and-white surroundings, that is before she is confronted with objects in their natural colors, she takes part in a kind of psychological test:

⁵Cf. Lewis (1988: 515). Lewis does not make the critical point clear here directly via the knowledge what it is like to see a color, but rather via the knowledge what it is like to taste vegemite (a kind of spice). It is clear, however, that the consideration is valid in the case of color too.

It is perhaps possible to imagine that Mary can utilize her utopian physical knowledge to influence her brain in such a way through neurosurgical intervention that she also acquires these abilities without actually seeing colors (cf. Lewis (1988: 516)). It was Daniel Dennett who emphasized that the utopian assumption of being completely informed about all physical aspects of color sight opens up quite a few surprising possibilities, cf. Dennett (1991: 398 ff.).

psychologists present her some monochromatic slides (blue, red, yellow, green) and ask her to point to the one showing the color of the sky. Since Marianna has heard that the color of the sky is particularly beautiful and since she is much impressed by the red slide, she chooses that one.⁶

According to Nida-Rümelin, we have good reasons for two contradictory belief attributions now. On the one hand, Marianna has heard from others that the sky is blue, and she would answer something like "The sky is blue" to appropriate questions. This supports the thesis that

(1) During the test, Marianna believes that the sky is blue.

On the other hand, she points to the red slide and declares that this is the color of the sky; usually, we would conclude from this that

(2) During the test, Marianna believes that the sky is red.⁷

However, she obviously cannot believe both propositions at the same time, at least *not in the same respect*.⁸ Accordingly, Nida-Rümelin proposes to make a difference between two respects in which Marianna believes something here. More exactly, she proposes to make a difference between phenomenal and a nonphenomenal beliefs about colors. We can then rewrite (1) and (2) as

(3) During the test, Marianna believes nonphenomenally that the sky is blue and

(4) During the test, Marianna believes phenomenally that the sky is red.⁹

As far as I see, Nida-Rümelin does not give any explicit definitions of the concepts of phenomenal and nonphenomenal belief. However, the basic idea seems to be that nonphenomenal beliefs about colors do not presuppose that the person in question has or had any immediate visual contact with colors. To have heard of colors and to be able to use color terms in a more or less satisfactory manner (for example, to be able to name the colors of certain objects on the basis of information gained by oral communication) is sufficient for nonphenomenal beliefs. On the other hand, phenomenal beliefs presuppose a kind of direct visual contact with colors.

Is this a non-physicalist solution to the epistemological problem in question? Nida-Rümelin emphasizes that it is in at least one respect: After all, it is essential

⁶Cf. Nida-Rümelin (1995: 221). Nida-Rümelin lays down two more details: Marianna is supposed to be normally sighted and to believe that she is normally sighted. These details are relevant mainly for the (alleged) problem of pseudonormal vision. I won't deal with that problem here. Cf. Nida Rümelin (1996) and Meyer (2000).

⁷ The belief attributions discussed in Nida-Rümelin (1995) are somewhat more complicated, but these complications are not relevant for our purpose. (1) and (2) correspond more or less to the versions Nida-Rümelin discusses in her (1993) book (cf. for example p. 11). I added the phrase "during the test"; it does not make any difference here, but it will be useful for later purposes.

⁸It is presupposed here that something blue cannot be red at the same time (and vice versa) for analytical reasons. Further, I make use of the concept of rational belief. Both presuppositions are unproblematic here. That Marianna seems to believe that the sky is blue in one and that the sky is red in another respect has certainly nothing to do with a lack of rationality.

⁹For purely technical reasons, Nida-Rümelin finally uses slightly different formulations, but this is not relevant for our purpose.

for phenomenal knowledge (or belief) that it can be gained only by immediate sensorial contact and not by any physical description whatsoever. So there is at least, according to Nida-Rümelin, a kind of knowledge which cannot be communicated by any physical, i. e. objective and scientific description. As far as physicalism entails the contrary, it is wrong.¹⁰ However, the conclusion that there is a kind of phenomenal knowledge which cannot be communicated by scientific and objective descriptions does not by its own commit us to the stronger non-physicalist thesis that there are phenomenal, non-physical *facts* which are the content of this knowledge. Nida-Rümelin is *inclined* to support this thesis as well, but she ultimately leaves the question unanswered.¹¹

However, what I want to argue for here, is that it is not even necessary to postulate the existence of phenomenal knowledge or belief in order to solve the epistemological problem posed by Nida-Rümelin. It is true that this problem cannot be solved if we refer to abilities alone. We must refer to epistemic concepts, too. But all we have to make use of is the well-known concept of *metalinguistic knowledge* and the method of *metalinguistic reinterpretation of belief sentences*.

3. A Metalinguistic Analysis

If one thinks about Marianna's situation for the first time, it is very likely that one instantly feels a strong inclination to assume that Marianna cannot be sufficiently aware of the meaning of color terms during the time of the experiment (this time being the most problematic). If not, how could she ever *say* that the sky is blue and at the same time choose the red slide as the one showing the color of the sky? Obviously, she thinks (wrongly) that red is the color which is called "blue" in English. This clearly indicates that she is not aware of the meaning of "blue" in English. If this is so, however, we cannot immediately infer the respective beliefs from what she *says* about colors and colored objects. Marianna may say something like "The sky is blue", but since she is not really aware of the meaning of "blue", we cannot conclude

(1) During the test, Marianna believes that the sky is blue
from that, but only

(5) During the test, Marianna believes that the color of the sky is called "blue"
in English.

In general, we can infer the belief that *p* from an utterance of "*p*" only if the speaker is sufficiently aware of what he is saying, of course. What we have here is a kind of *metalinguistic reinterpretation* of Marianna's utterance of "The sky is blue".

Since Marianna comes to the conclusion that a certain slide, namely the red one, shows the color of the sky, we can assume that

¹⁰Cf. Nida-Rümelin (1995: 237 ff.).

¹¹Cf. Nida-Rümelin (1993: 80 ff.) and (1995: 237 ff.).

(2) During the test, Marianna believes that the sky is red is true: after all, she is in direct visual contact with something red and she thinks that *this* is the color of the sky.¹² Note that it does not follow from (2) that Marianna would *say* that the sky is red. This would follow only if she were sufficiently aware of the meaning of "red" - but in fact she isn't: we would say that she is sufficiently aware of the meaning of "red" only if she were able to identify the color of the red slide as the one being called "red". But that is just what she actually cannot do. Rather, she obviously believes that the color of the slide is called "blue", because she thinks that the color of the slide is the same as the color of the sky and that the color of the sky is called "blue".

As a result, we have not to cope with the contradictory set

(1) During the test, Marianna believes that the sky is blue
and

(2) During the test, Marianna believes that the sky is red,
which is the basis for Nida-Rümelin's argument. Rather, we should describe the situation by the belief attributions

(5) During the test, Marianna believes that the color of the sky is called "blue"
in English

and

(2) During the test, Marianna believes that the sky is red,
which are fully compatible with each other. In this way, Nida-Rümelin's problem disappears.

In my opinion, the advantage of this analysis is that it is based solely on some elementary principles of belief attribution and on the concept of metalinguistic knowledge, which is, taken on its own, quite unproblematic: at least, it should be undisputed that there are *some* cases in which somebody is not aware of the meaning of a term and that in these cases it is not possible to attribute a belief immediately on the basis of a respective utterance. It is not necessary to introduce any concepts here which are not well known from other contexts: in particular, it is not necessary to introduce a new distinction between phenomenal and nonphenomenal knowledge.

However, in the philosophical discussion some important objections have been raised against an all too careless use of metalinguistic analyses. In the next section, I shall reconsider and elaborate my proposal against this background.

¹²Additionally, my assumptions about Marianna's beliefs during the test can be supported by the following argument: When Marianna finally leaves the artificial surroundings of the experiment and looks at the blue sky for the first time, she will be surprised and will notice that she was *wrong* about the color of the sky during the test. This can be interpreted suitably only if we assume that she does not believe that the sky is blue during the experiment (but only that the color of the sky is called "blue" in English): instead, during the test, she believes that the sky is red.

4. The Context Dependency of Linguistic Competence. Linguistic Competence and Abilities

The most influential criticism of an all too careless use of the method of metalinguistic reinterpretation is probably to be found in Burge (1979). According to Burge, a somewhat erroneous or incomplete understanding of the meaning of a concept does not necessarily lead to the conclusion that we cannot attribute beliefs containing this concept to the respective person:

"[...] it is important to see what an array of conceptual errors is common among us. And it is important to note that such errors do not always or automatically prevent attribution of mental content provided by the very terms that are incompletely understood or misapplied."¹³

In Burge (1979), this is made clear by some examples. For instance, someone can very well believe that he has arthritis in the thigh, although arthritis is *defined* as an inflammation of the joints. Further, many people believe erroneously that it belongs to the definition of a contract that it must be put down in writing; nevertheless, we do not hesitate to attribute beliefs about contracts to these people on the basis of their use of the word "contract".¹⁴ If we made use of the method of metalinguistic reinterpretation in all the many cases of a somewhat erroneous or incomplete understanding of a concept ("he believes that he has something in the thigh which is called 'arthritis' in English"), our common practice of belief attribution would become quite complicated or even impossible. The terms "arthritis" and "concept" have their ordinary meanings even if they are used by imperfectly informed laymen. This is because what a term means when it is used by a speaker does not depend on the speaker's conception of this term alone. Rather, it is also dependent on the manner the term is used by the linguistic community as a whole and by the experts in the respective areas. Burge refers to Putnam's concept of division of linguistic labor here.¹⁵ In order to refer to arthritis by the word "arthritis", it is not necessary to know the exact medical definition of this term: for a layman, it is enough to be a part of a large and more or less coherent and productive network of communicative structures in which experts knowing this definition play a certain role.

On the other hand, there are of course many cases in which metalinguistic reinterpretation is the appropriate method, as Burge himself is well aware:

"There are also examples of quite radical misunderstandings that sometimes generate reinterpretation. If a generally competent and reasonable speaker thinks that 'orangutan' applies to a fruit drink, we would be reluctant, and it would unquestionably be misleading, to take his words as revealing that he

¹³ Burge (1979: 542).

¹⁴ Cf. Burge (1979: 538 ff.)

¹⁵ Cf. Burge (1979: 564 (footnote 2)). This concept is to be found in Putnam (1975: 227 ff.), for instance. Of course, it is a very difficult question how this division of linguistic labor works in each single case. I shall not say much about it here. As far as natural kind terms are concerned, some more detailed considerations can be found in Meyer (1998: 206 ff. (ch. 16)).

thinks he has been drinking orangutans for breakfast for the last few weeks."¹⁶

The question when we should take a speaker's words seriously and infer the respective beliefs from them or when a metalinguistic reinterpretation is appropriate, depends on many different aspects:

"A person's overall linguistic competence, his allegiance and responsibility to communal standards, the degree, source, and type of misunderstanding, the purposes of the report - all affect the issue. From a theoretical point of view, it would be a mistake to try to assimilate the cases in one direction or another. We do not want to credit a two-year-old who memorizes 'e=mc²' with belief in relativity theory. But the patient's attitudes involving the notion of arthritis should not be assimilated to the foreigner's uncomprehending pronunciations."¹⁷

Can we stick to the solution developed in the last section if we take all this into account? I do think so. However, a more detailed analysis of Marianna's story is called for. In particular, it is necessary to distinguish three periods, namely Marianna's time in her black-and-white room, the period of the experiment in which she is shown the colored slides, but still has no opportunity of seeing objects in their natural colors, and the period after the experiment, when Marianna finally begins to live an ordinary life in our colored world.

As far as the first period is concerned, i. e. the period in the black-and-white room, I propose to take her utterances about colors seriously: in general, we should infer the respective beliefs from her utterances about colors and colored things. If, for example, she utters the sentence "The sky is blue" because she remembers that one of her trustworthy contacts talked about the wonderful blue sky every once in a while, we should in general attribute to her the belief that the sky is blue.

This may seem a bit surprising at first glance. Why suppose that Marianna knows what she is talking about when she utters something like "The sky is blue" in her black-and-white room, i. e. during a period in which she has never had any visual contact with colors? If we decide to judge the case against the backdrop of linguistic competence, metalinguistic reinterpretation etc., shouldn't we suppose that Marianna's knowledge about the meaning of "blue" is so incomplete that it is inappropriate to infer the respective belief from her utterance of "The sky is blue"?

I don't think so. The point is that, with respect to colors, no particular communicative or epistemological problems arise as long as Marianna stays in her room and is, for instance, never challenged to name colors on the basis of visual contact. Burge emphasized that there is no *absolute* answer to the question whether a speaker is sufficiently well informed about the meaning of the words he uses in an utterance to infer a respective belief from it. On the contrary, the an-

¹⁶ Burge (1979: 547).

¹⁷ Burge (1979: 548).

swer is dependent on many aspects of the context. In her particular situation, Marianna certainly lacks certain competencies with respect to colors and color terms which an *ordinary* speaker in an *ordinary* context is supposed to have; in particular, she is not able to look at a color and name it on the basis of its appearance. However, it seems that she does not need this competence either. Why should we, for example, require of Marianna to be able to name colors correctly on the basis of their appearance if she doesn't ever see any colors anyway? Under these circumstances, her inability cannot lead to any mistakes and confusions which could make us doubt her linguistic competence. Marianna knows colors only by hearsay, but so what? In her particular situation, we can nevertheless regard her as a competent speaker who does not make any serious mistakes because, in a certain respect, it is just the opportunity to make such mistakes she lacks. Color terms like "red" and "blue" have their ordinary meanings in Marianna's utterances, because these utterances can be part of a completely coherent communication with ordinary speakers, which can (in the sense of the hypothesis of the division of linguistic labor) be regarded as experts in the identification of colors. Accordingly, if Marianna utters something like "The sky is blue" in her black-and-white room, we can infer

(6) In her black-and-white room, Marianna believes that the sky is blue.

Immediately after her release, i. e. during the experiment, things are different. In Marianna's new surroundings, there are colored things, more exactly colored slides, for her to be seen. In this context, the ability to name colors on the basis of their appearance plays an important part. *People who live in colored surroundings and who can see colors* must be able to call something blue "blue" and something red "red" in order to count as a competent speaker with respect to color terms. If Marianna cannot do this, we cannot attribute to her the linguistic competence in question, and accordingly we cannot infer the respective beliefs from her utterances about colors. From her utterance of "The sky is blue", we can now only conclude that she believes that the color of the sky is called blue, that is:

(5) During the test, Marianna believes that the color of the sky is called "blue" in English.

In the last section, we saw that Nida-Rümelin's epistemological problem can be solved by this metalinguistic reinterpretation. Indeed, this kind of reinterpretation seems suitable here. Marianna's situation is quite different from Burge's arthritis-case, in which a metalinguistic reinterpretation seems inappropriate. The patient imagined by Burge does not know the exact medical definition of "arthritis", but he nevertheless knows enough about this illness that a more or less coherent communication about it is possible for him. To use a term of Putnam's: he has the *stereotyped* knowledge which makes him a sufficiently competent speaker (in most ordinary life situations, at least).¹⁸ On the other hand, Marianna lacks the *stereotyped ability* which is absolutely necessary for every nor-

¹⁸ On the concept of 'stereotype' in this context cf. Putnam (1975: 247 ff.)

mally sighted person in the given situation to count as a competent speaker: she cannot name colors on the basis of their appearance. If we still attributed to her a complete knowledge of the meanings of color terms, this would lead to quite a lot of inconsistencies - amongst others to just the inconsistency Nida-Rümelin refers to.

It should be stressed that the ability to name colors on the basis of their appearance plays a decisive part only if the speaker is normally sighted, in particular if he can see colors. It would certainly not be justified to presuppose *a priori* that color-blind persons can never know the meaning of a color term. In general, we can assume that an ordinary color-blind person knows the meaning "blue", for example, and accordingly we can attribute the belief to him that the sky is blue if he utters the sentence "The sky is blue". Under normal circumstances this will not lead to any inconsistencies, because color-blind persons themselves as well as their normally sighted interlocutors know that color-blind persons are excluded from certain moves in the language game from the outset: color-blind persons won't try to speculate about the colors of objects in their surroundings without any further hints, and no reasonable interlocutor will encourage them to do that. Situations which tend to produce inconsistencies are thus avoided. In a certain respect, this is again a case of division of linguistic labor. Normally sighted speakers who grew up in ordinary surroundings can count as experts in identifying colors here. Being uttered by color-blind people, color terms have their ordinary meanings because color-blind persons stand in a more or less coherent and productive communicative relationship to ordinary people counting as 'experts'. In this respect, color-blind persons resemble Marianna during her time in the black-and-white room. Indeed, it is partly this resemblance, together with the intuition that it is not justified to presuppose that color-blind persons can never know the meaning of a color term or have beliefs about the colors of objects, which makes me assume that we can attribute ordinary beliefs about colors to Marianna in her black-and-white room also.¹⁹

During the time of the experiment, however, Marianna - being a normally sighted person - can see the colors of the slides very well but assigns the wrong names to them: in effect, she assigns the term "blue" to the red slide. I can hardly imagine a case in which there would be more reason to suppose that someone is not really aware of the meaning of "blue".

Finally, when Marianna has learnt to name colors correctly some time after the experiment (perhaps because somebody taught her or she is now allowed to see 'objects' like the sky in their natural colors), things change for the last time. Now she meets the requirements which are made of a normally sighted, competent speaker in a colored environment with respect to his command of color terms,

¹⁹It was the critical remarks of an anonymous referee and of some participants of the Konstanz meeting of the *Forscherguppe Logik und Philosophie* (23.-24.03.2000), in particular of Andreas Kemmerling, on an earlier version of this paper which made it clear to me that I should be more explicit about this point.

and there is no reason not to infer the respective beliefs from her utterances in question. Therefore, we can assume:

(7) After the test and after an appropriate period of learning, Marianna believes that the sky is blue.

This analysis of Marianna's case seems to do justice to the objections which could be raised against it on the basis of Burge's considerations about the method of metalinguistic reinterpretation. Indeed, I think that my solution is in full accordance with what Burge, Putnam, Kripke and others so eagerly emphasize: the question which beliefs we should attribute to a person and which meaning we should give his words does not only depend on his 'internal states', but also on his surroundings. Meanings are not in the head, as Putnam says.²⁰ This is exactly what I presuppose when I assume that we can ascribe to Marianna the full linguistic competence with regard to color terms as long as she stays in her black-and-white room, but not in the colored environment in which the experiment takes place. The question when we can take a speaker's words seriously and infer the respective beliefs from them and when we must reinterpret his utterances is, as Burge rightly emphasizes, dependent on many aspects. In particular, the point is how well he can use the respective terms *in a given situation*. Finally, the criterion is the consistency and productivity of communication.

If we consider this analysis to be correct, must we say that Marianna acquires any relevant knowledge during the whole story? It seems not. Of course, many things have changed for Marianna. She can look at colors now, something she couldn't do in her black-and-white room; and she is now able to name colors on the basis of their appearance. But there seems to be no increase of *knowledge* in any relevant respect. As far as her knowledge of the meanings of color terms is concerned, many things have happened, of course; however, there is no moment at which Marianna knows *more* about these meanings than during the time in her black-and-white room. She had the appropriate knowledge when she lived in the 'undemanding' surroundings of her black-and-white room; she *lost* this knowledge when she came into the more demanding surroundings of the experiment, since she could see colors now but could not name them correctly; that means that at the moment she entered the colored surroundings of the experiment she had even *less* knowledge than during the time in her room; and finally she regains this knowledge when she learns to name colors correctly on the basis of their appearance, i. e. when she gains a certain ability. Of course, the way we interpret Marianna's utterances about colors changes in a completely analogue manner. However, it should be noted that we never take her utterances more seriously than during her time in the black-and-white room. She knows that the sky is blue at the beginning of the story as well as at the end, and again there is a certain period, i. e. the time of the experiment, at which she is in a worse position with respect to this belief than she is during the time in her room.

²⁰Cf. Putnam (1975: 227).

In sum, we have come to the following result. It is true that we need epistemic concepts to give an adequate analysis of Marianna's case. It is not enough to refer to abilities only. However, firstly the changes in what Marianna knows depend essentially on changes in what she is able to do, so that abilities remain fundamental; and secondly, our analysis does not presuppose that there is a moment after her time in the black-and-white room at which she *knows more* about colors than in the room in any relevant respect. So this solution for Nida-Rümelin's epistemological problem does not lead to the conclusion that there is some 'nonphysical' knowledge about colors Marianna can on principle not have as long as she is imprisoned in her room - whether we consider this nonphysical knowledge as knowledge of nonphysical facts (Jackson) or as a particular phenomenal knowledge (Nida-Rümelin).

5. Knowing What "Blue" Means, and Knowing What Blue Looks Like

The analysis I proposed in the last two sections essentially refers to the ability to name colors correctly on the basis of their appearance. As far as the epistemological problem of the Marianna-case is concerned, Marianna's decisive progress in the course of the story is her acquisition of this ability: what she gains after being released from her room is just this ability, but not any kind of knowledge. However, one could raise the objection that in general the ability to name something correctly on the basis of its appearance *is founded on a kind of knowledge*. For instance, take the ability to call a tiger a "tiger" correctly on the basis of its appearance. Obviously, we have to know some facts about tigers in order to be able to do this - e. g. what tigers look like, perhaps also how they behave, where they live, etc.; and we have to know that the animals described in such and such a way are called "tigers" in English, of course. It is quite obvious that this is real knowledge, which can be looked up in an encyclopedia or in a dictionary. Someone who knows that tigers are Asiatic big cats with a total body length in excess of 10 ft and a height of 3 ft, that a tiger's fur ranges from orange to brownish yellow with a white chest and belly and is covered with broken vertical black or dark brown stripes, that tigers live in colder forested regions of Asia and in some other areas etc., someone who knows all these facts knows what a tiger is; if he also knows that these animals are called "tigers" in English, he will probably be able to call tigers "tigers" correctly in English on the basis of their appearance - given that his visual abilities are normal, of course. It is quite obvious that to know what a tiger is and to know that tigers are called "tigers" in English is not simply the same: many Chinese people know what a tiger is, but probably only a minority of them knows that these animals are called "tiger" in English. However, this deficiency could easily be removed with the help of a dictionary. It seems that under normal circumstances two pieces of knowledge are sufficient for someone to be able to call a tiger a "tiger" correctly in English: he must

know what a tiger is and that the animals described in a certain way are called "tigers" in English.

Couldn't we now argue that under normal circumstances and given a normally sighted speaker there are also two bits of *knowledge* which are sufficient for the ability to call something blue "blue" correctly in English - namely the *knowledge what blue looks like* and the knowledge that this particular color is called "blue" in English? But if this is true, it seems that there is *still* a kind of knowledge Marianna lacked as long as she lived in her black-and-white room. Marianna was not able to name colors correctly during the period in her room; if a certain kind of knowledge is *sufficient* for this ability, it follows that she did not have this knowledge either. Obviously, she could not gain this knowledge until she left her room. One could argue that my analysis may be correct so far, but ignores that the decisive ability is based on a kind of knowledge. So Marianna still gained a kind of knowledge she could not have in her black-and-white room.²¹

This objection is not conclusive, however. It is true that under ordinary circumstances the ability to call blue things "blue" on the basis of their appearances is entailed by the knowledge what blue looks like and what this particular color is called in English. It follows that Marianna could not know what blue looks like or what this color is called in English during the period in her room, because it is presupposed that she could acquire the ability in question after her release only. Indeed, we should say that she did not know what blue looks like during the period in her room. However, in contrast to the knowledge of what a tiger is - or what a tiger looks like - the 'knowledge' what blue looks like is no *real* knowledge. The knowledge what a tiger looks like and the knowledge what blue looks like both make it possible to imagine the respective object, i.e. a tiger or the color blue, to identify it etc. But there is an important difference: the latter 'knowledge' cannot be acquired with the help of verbal descriptions, but *only* by an immediate visual contact with the object itself and by practice. One must look at colored objects and compare them to each other, listen to explanations like "*This* color is the same as *that*" etc. What we can acquire *only* by practice and in an immediate contact with the thing itself is generally considered as an ability, not as a kind of knowledge, however. In sum: What we can learn with the help of verbal descriptions is a kind of knowledge, which can possibly entail certain abilities. But what we can learn by practice only is an elementary ability which is not based on any real knowledge. Accordingly, the knowledge what blue looks like (in contrast to the knowledge what a tiger looks like) is not the *basis* for certain abilities: in fact, it is identical with them.²²

²¹The objection that an increase in abilities can be explained best by an increase in knowledge and that for this reason we cannot presuppose that Marianna acquires new abilities without assuming that she acquires some new knowledge is often raised against Lewis' and Nemirow's analysis. For example, see Lycan (1995: 247 f.), Nida-Rümelin (1995: 234 ff.), and Gertler (1999: 322 ff.).

²²Of course, we generally consider a sentence like "Orange is the color between red and yellow" as a possible answer to the question what orange looks like. But the fact that this sentence is an

So I admit that Marianna could learn what colors look like only after her release from the black-and-white room. But this is no problem, because the knowledge what a color looks like is identical with the abilities to imagine this color, to identify it etc. And it is obvious that *these abilities* could only be acquired outside the room: indeed, this is presupposed explicitly in Nida-Rümelin's variant of the thought experiment. To learn what colors look like, *that is* to learn to imagine colors, to identify them etc. is possible for Marianna only in a colored environment.

The decisive ability to name colors correctly can be acquired by Marianna only if additional information like "The first slide is red, the second slide is blue" etc. is given to her. So she actually has to rely on *information* here. However, this is no information she could not have had in her room for any deeper reasons. After all, it would have been possible to inform her before her release that the first slide will be blue and the second will be red in the experiment. In this case, only the knowledge what these colors look like, *that is* the abilities to imagine and remember them, would have been necessary to acquire the ability to name them correctly. This would be very similar to a situation in which Marianna would not have been led into the artificial environment of the experiment at first, but immediately into the natural environment of objects the colors of which she could tell us in her room already. So it is true that there is some real knowledge involved in the ability to name colors correctly on the basis of their appearance; but this is not a kind of knowledge Marianna could not have had during the time in her room.

In sum, a more exact analysis of the ability to name colors correctly does not lead to the result that there is some real knowledge Marianna could not have in her black-and-white room for fundamental reasons. Essentially²³, it was certain abilities she could not acquire.

Of course, this is along the same lines as Lewis' and Nemirow's arguments. However, there are some important differences in the details. In contrast to Lewis and Nemirow, I think that we cannot do without genuine epistemic concepts if

appropriate answer in some contexts should not tempt us to think that the knowledge what orange looks like is a real, ordinary kind of knowledge the content of which is the proposition expressed by the sentence in question. Under normal circumstances, we use this sentence to invite someone to imagine the color between red and yellow and to inform him that this color is called "orange". We would say that he knows what orange looks like only if he manages to imagine orange and to identify orange objects more or less reliably - i. e. if he has the decisive *abilities*. The information contained in "Orange is the color between red and yellow" is not the object of the knowledge what orange looks like, but only a possible way to gain this knowledge, as it were. This way begins with the knowledge what red and yellow look like, i. e. with the *abilities* to imagine this colors, to remember them, to identify them etc., which, finally, is a question of practice. Incidentally, it follows that Marianna can know that orange is the color between red and yellow in her black-and-white room without knowing what orange looks like in the sense which is relevant here. And the same is true of color-blind persons.

²³There are some other things she couldn't enjoy for trivial reasons: for example, she lacked the experience of looking at colors, of course.

we want to do the knowledge argument justice. This is the most important, but not the only difference. A second one concerns the question which kind of 'knowledge' can be identified with which ability. According to Lewis, we can identify the *knowing what it is like to look at something blue* with the ability to remember blue, to imagine this color, to recognize experiences of it and to name it correctly.²⁴ We have come to a different and more complex result here:

1. To know what blue looks like means to be able to imagine blue, remember blue, and identify blue; this "knowledge" is an ability because it can be acquired by practice only, but not by listening to any verbal descriptions.
2. To know the meaning of "blue" amounts to different things in different contexts. In ordinary, colored surroundings and for normally sighted persons who can see colors it means to be able to name blue correctly on the basis of its appearance. This presupposes the knowledge what blue looks like (in the sense of 1.). In black-and-white surroundings like Marianna's room²⁵ this presupposition does not hold: under such conditions, it is possible to know what "blue" means without knowing what blue looks like and without being able to name blue objects correctly on the basis of their appearance.

Remarkably, the much discussed "knowing what it is like" does not occur here. Indeed, I think that this phrase is of some importance in its actual, ordinary use but does not play an important part for Mary's or Marianna's case. I shall be a bit more explicit about this in the last section.

To sum up, we have arrived at the result that the knowledge argument is inconclusive. An exact analysis of Marianna's case has revealed that her situation can be described without postulating any kind of knowledge she cannot have inside, but only outside her black-and-white room. This analysis can easily be applied to Jackson's original case, too. So no clues arise about the existence of any non-physical facts about perceptions which can be known only on the basis of an immediate visual contact with colors. Jackson's original anti-physicalist thesis ultimately seems unfounded. Further, we need not introduce a particular phenomenal knowledge as it is proposed by Nida-Rümelin.²⁶ It is just ordinary knowledge about colored objects, knowledge about the meanings of color terms, and some abilities that are involved here.

²⁴Cf. Lewis (1988: 515). As mentioned before, Lewis does not refer to the knowledge what it is like to see something blue but to the knowledge what it is like to taste vegemite here.

²⁵Or for color-blind persons, cf. section 4 above.

²⁶And finally it is not necessary to postulate a kind of knowledge which arises from a particular introspective access to facts which are in principle accessible to objective and scientific methods either. For the sake of brevity, I shall not deal with this rather prominent analysis of the knowledge argument. Cf. Churchland (1995: ch. 8), Loar (1990), Tye (1995: ch. 6, 161 ff.), Lycan (1995), and many others.

6. Knowing What It Is Like

In this last section, I'd like to make some short remarks on the much discussed *knowing what it is like*. This knowledge is interesting enough on its own, but in our context it is particularly important because some philosophers argue that Marianna or Mary can know what it is like to see colors only after her release and that we must *for syntactical and semantical reasons* assume that the *knowing what it is like* is a real kind of knowledge (not an ability).²⁷ The object of this knowledge is supposed to be (in a way non-physical) phenomenal information which is 'subjective' in the sense that it cannot be communicated by verbal descriptions: the only access to this kind of information is by immediate experience.²⁸

However, this seems wrong to me. To see that, it is not even necessary to care about the rather sophisticated details of the syntactical discussion. All we have to do is to look at the actual way the expression "knowing what it is like" is used in ordinary contexts. I think that this expression has two related and quite important meanings in the ordinary language - but it never expresses an only introspectively accessible, subjective proposition or anything like that.

In its first sense, *knowing what it is like* refers to judgements or valuations, often of a more or less emotional kind. For example, what is it like to be free after having served many years in prison? Of course, it is good to be able to walk outside as long as you want. It is good not to be a part of the violent hierarchy of a prison any longer. On the other hand, it is very difficult to get along in the world outside. Many things have changed, perhaps you have lost all your friends during the long time you were imprisoned; you have forgotten how to act on your own, and for a man with a previous conviction it is very difficult to find a job. This could be a completely ordinary answer, and everybody can understand it, even if they have never been a prisoner. In this sense, everybody can know what it is like to be free after having served a long prison sentence, although not everybody has had such an experience.

Answers to *what is it like*-questions are often of practical and moral significance. What is it like for a chicken to live its life on a small metal grating in a chicken farm? Is this question unanswerable because we can never have any access to chicken-qualia which can only be experienced introspectively or because chicken have some phenomenal knowledge we can never acquire? It seems not. On the

²⁷Cf. Lycan (1995: 244 f.), for example. Of course, there are some other standard objections against Lewis' and Nemirow's interpretation of Mary's case, cf. Levin (1990), or Lycan (1995: 244 ff.) for a survey. A new objection was raised by Tye (forthcoming) recently. It would be interesting to examine these objections with respect to the analysis proposed here. However, this would require another paper.

²⁸Cf. Lycan (1995: 254). What, for example, is the phenomenal information Mary gets access to when she leaves her room? Well, "that actually to experience red is like - ploiku! ..." - where "ploiku" is supposed to be a morpheme of Mary's private language which can, strictly speaking, not be translated into standard English or any public language. Really: ploiku!

contrary, the answer is obvious: to live such a life is tormenting for an animal which would spend its time in nature with running around, scratching, and pecking. You needn't be a chicken to judge this.

And what is it like to see blue? It depends. In some respect, it is nothing to write home about for the most of us. After all, we frequently see something blue in our surroundings. In another respect it is often said that blue gives rise to a feeling of coolness and distance, but also of freedom. I personally like blue; it would be more pleasant for me to be in a room with a blue carpet than in a room with an orange one. And for Mary and Marianna it would of course be very exciting to see something blue for the first time; after a while, they probably would become used to it, and it would not be much more exciting for them than it is for us.

In this sense, Marianna can know perfectly well what it is like - for someone particular or in general - to see something blue even before her release. After all, I could tell her for instance that to see something blue is more pleasant for me than to see something orange. Or perhaps it would be possible for her to find out that to see something blue generally gives rise to a certain range of feelings. There seems to be no reason to assume that she cannot have such beliefs. A color-blind person can know these sorts of things about colors too, and someone who has never been in serious pain can of course know that it is very bad to be in serious pain - and that we therefore should do everything to prevent human beings (and animals) from suffering it.

In all these cases of *knowing what it is like* it is quite clear that the content of this knowledge does indeed go beyond the realm of the 'physical' sciences. However, this kind of knowledge does not refer to any only introspectively accessible 'phenomenal' propositions; nor is it a particular 'phenomenal' kind of knowledge. In fact, it is about more or less emotional *evaluations*, which are in many cases of some moral significance. That this does not belong to the realm of the objective, 'physical' sciences is true, of course, but it is not the latest news: on the contrary, it is the prevailing doctrine.

There is, however, a second sense of *knowing what it is like*. In this second sense we sometimes say of a person who has never been in serious pain that he does not *really* know what it is like to be in serious pain *even* if he does know it in the sense discussed above - i. e. even if he is well informed about all emotional and moral aspects of pain anyone can think of. We can make a distinction here between *theoretical* and *practical* knowledge, the latter arising only from personal experience. The point seems to be that in the case of merely theoretical knowledge the respective person *knows* that serious pain gives rise to extremely negative emotional valuations but does not *feel* the emotions in question when he thinks about serious pain or is confronted with someone in serious pain; or at least he does not feel these emotions as intensively as someone who has had experiences of pain and all its emotional concomitants. In persons who have had such experiences, the respective emotions are *coupled to* thoughts about pain and situations in which someone suffers from pain. Someone who has had expe-

periences of pain will feel certain negative emotions when he is confronted with it in any way: when he recognizes that a person is in pain, for example, or even when he merely thinks about pain. Someone who only knows about pain in a theoretical way will probably not feel these emotions in situations of this kind.

I shall not maintain here that *knowing what it is like* in its practical sense can be acquired by personal experience *only*, though this seems to be the most effective way. For example, a doctor probably can acquire this practical knowledge, i. e. the ability to feel sympathy or compassion in the strictest sense of the word, by frequently observing the close connection between physical pain and emotional grief *in practice* (as it were), that is in his patients. On the other hand, reading textbooks will probably not be enough. The practical *knowing what it is like* to be in pain seems to turn out more or less as the disposition to feel real sympathy with people suffering from pain. And this disposition, the *coupling* between the perception of a situation and the rise of the suitable emotions, as it were, can be acquired by an immediate contact to the world only. The same seems to be true of positive experiences and emotions.²⁹

In principle, there is also some practical *knowing what it is like* to see colors: Those who have never seen the beautiful colors of a sunrise do not ('really') know what this is like, i. e. there will be no particular emotion connected with thoughts about sunrises in them, or at least they will not be as engaged emotionally as people who have had experiences of beautiful sunrises.

Such practical knowledge hardly plays a particularly important part as far as colors are concerned, but in many cases the difference between theoretical and practical *knowing what it is like* has substantial consequences. For example, someone who knows what it is like to be in serious pain in the practical sense will in general react to other people's pain more spontaneously and 'instinctively' than somebody who has the respective theoretical knowledge only - even if the latter one has the best intentions in principle. Finally, real sympathy is a stronger motive than theoretical reasons. However, in those cases in which we cannot empathize with a being, the theoretical *knowing what it is like* is of great significance. The more the being in question differs from us, the more limited is our empathy, of course. As far as other human beings with similar experiences are concerned, it is comparatively easy for us to feel ourselves into them. But when it comes to chickens or even fishes, things get rather difficult. I can know what it is like for a human being to be locked up in a very confined space, in the

²⁹In a somewhat different context, R. de Sousa emphasized that the connection between certain situations with suitable emotions must be learned in a process of socialization. Cf. his considerations about "paradigm scenarios" in (1987: 181 ff.). According to de Sousa, the study of literature can play an important part in this learning process, in particular as far as adults are concerned. Note that this would not contradict my thesis: de Sousa does not mean textbooks here, but novels, dramas etc., in which the connections are not described theoretically but presented in a concrete and exemplary manner. Indeed, it was a central thesis of the authors of the Enlightenment that it is one of the pedagogical purposes of literature to teach the readers to develop the suitable emotional reactions by confronting them with exemplary destinies.

practical sense; what this is like for a chicken I probably can know in the theoretical sense only, i. e. after having observed the behavior of chickens in their natural environment. However, in a theoretical sense it becomes clear rather quickly that it is quite bad for a chicken to be locked up in a very confined space, too. This theoretical knowledge *commits* us morally to prevent chickens from such a fate if possible. That our interest in cheap eggs in actuality prevails over this moral obligation shows that a merely theoretical *knowing what it is like* is a rather weak motive in many cases.

I do not want to continue with these more or less moral considerations here.³⁰ What is important in our context is that apart from the theoretical sense there is a *practical* sense of *knowing what it is like* which is not contained in any 'physical' theory of the world and which can be acquired by a (more or less) immediate contact with things themselves only. At first glance, this could look like a late triumph of anti-physicalism: after all, that doctrine postulates a knowledge of this kind. However, a closer examination quickly reveals this as a misunderstanding. Again, such a practical 'knowledge' is not really knowledge, but a basic disposition or ability: the disposition to react to certain situations with adequate emotions or the ability to feel oneself into the emotional position of other sentient beings. So finally it turns out that a closer look at the actual meaning of "knowing what it is like" supports anti-physicalist conclusions just as little as a detailed analysis of Mary's or Marianna's case.³¹

³⁰Cf. Lenzen (1999: 294 ff.) for some more considerations about *knowing what it is like* in moral contexts.

³¹I am grateful to Andreas Kemmerling, Wolfgang Lenzen, Martine Nida-Rümelin, an anonymous referee, and the participants of the Konstanz meeting of the *Forschergruppe Logik und Philosophie* (23.-24.03.2000) for helpful comments. This work was supported by grant ME 1622/1-2 from the *Deutsche Forschungsgemeinschaft*. Ulf Heinecke helped with linguistic corrections.

References

- Burge, T.: 1979, 'Individualism and the Mental', *Midwest Studies in Philosophy*, Vol. IV, Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein (eds.), University of Minnesota Press, Minneapolis, 73-121. Here quoted from Rosental, D. M. (ed.): 1991, *The Nature of Mind*, Oxford University Press, New York, Oxford, 536-567.
- Churchland, P. M.: 1995, *The Engine of Reason, the Seat of the Soul*, Massachusetts Institute of Technology.
- Dennett, D. C.: 1991, *Consciousness Explained*, Little, Brown, and Company, Boston, Toronto, London.
- de Sousa, R.: 1987, *The Rationality of Emotion*, The MIT Press, Cambridge/Mass., London.
- Gertler, B.: 1999, 'A Defense of the Knowledge Argument', *Philosophical Studies* 93, 317-336.
- Jackson, F.: 1982, 'Epiphenomenal Qualia', *Philosophical Quarterly* 32, 127-36. Here quoted from Lycan, W. G. (ed.): 1990, *Mind and Cognition*, Basil Blackwell, Cambridge/Mass., 469-477.
- Jackson, F.: 1986, 'What Mary Didn't Know', *Journal of Philosophy* 83, 291-295.
- Lenzen, W.: 1999, *Liebe, Leben, Tod. Eine moralphilosophische Studie*, Reclam, Stuttgart.
- Levin, J.: 1990, 'Could Love Be Like a Heatwave? Physicalism and the Subjective Character of Experience', Lycan, W. G. (ed.): 1990, *Mind and Cognition*, Basil Blackwell, Cambridge/Mass., 478-490.
- Lewis, D.: 1983, 'Postscript to "Mad Pain and Martian Pain"', *Philosophical Papers*, Vol. 1, Oxford University Press, New York, Oxford, 130-132.
- Lewis, D.: 1988, 'What Experience Teaches', *Proceedings of the Russellian Society*, University of Sydney. Here quoted from Lycan, W. G. (ed.): 1990, *Mind and Cognition*, Basil Blackwell, Cambridge/Mass., 499-519.
- Loar, B.: 1990, 'Phenomenal States', J. Tomberlin (ed.): *Philosophical Perspectives*, Vol. 4, Ridgeview Publishing Co., Atascadero, 81-108.
- Lycan, W. G.: 1995, 'A Limited Defence of Phenomenal Information', Metzinger, T. (ed.), *Conscious Experience*, Imprint Academic/Schöningh, Paderborn, 243-258.
- Meyer, U.: 1998, *Glaube und Notwendigkeit. Eine Untersuchung zur Sprachphilosophie, zur Erkenntnistheorie und zur Philosophie des Geistes*, Schöningh, Paderborn, München, Wien, Zürich.
- Meyer, U.: 2000 (forthcoming), 'Do Pseudonormal Persons Have Inverted Qualia? Scientific Hypotheses and Philosophical Interpretations', *Facta Philosophica*.
- Nagel, T.: 1974, 'What Is It Like To Be a Bat?', *The Philosophical Review* 83, 435-450.

- Nemirow, L. E.: 1980, 'Review of Nagel's Mortal Questions', *Philosophical Review* 89, 475-476.
- Nemirow, L. E.: 1990, 'Physicalism and the Cognitive Role of Acquaintance', Lycan, W. G. (Ed.): *Mind and Cognition*, Blackwell Publishers, Cambridge/Mass.
- Nemirow, L. E.: 1995, 'Understanding Rules', *The Journal of Philosophy* 92, 28-43.
- Nida-Rümelin, M.: 1993, *Farben und phänomenales Wissen*, VWGÖ, Wien.
- Nida-Rümelin, M.: 1995, 'What Mary Couldn't Know: Belief about Phenomenal States', Metzinger, T. (ed.), *Conscious Experience*, Imprint Academic/Schöningh, Paderborn, 219-241.
- Nida-Rümelin, M.: 1996, 'Pseudonormal Vision - An Actual Case of Qualia Inversion?', *Philosophical Studies* 82, 145-157.
- Putnam, H.: 1975, 'The Meaning of 'Meaning'', Gunderson, K. (ed.): *Language, Mind, and Knowledge*, University of Minnesota Press, Minneapolis. Here quoted from Putnam, H.: 1975, *Mind, Language and Reality. Philosophical Papers, Vol. 2*, Cambridge University Press, Cambridge, London, New York, Melbourne, 215-271.
- Tye, M.: 1995, *Ten Problems of Consciousness*, The MIT Press, Cambridge/Mass., London.
- Tye, M.: forthcoming, 'Knowing What It Is Like: the Ability Hypothesis and the Knowledge Argument'.