

**teorema**

Vol. XXX/3, 2011, pp. 113-133

ISSN: 0210-1602

[BIBLID 0210-1602 (2011) 30:3; pp. 113-133]

**Phenomenal Concepts, Color Experience, and Mary's Puzzle<sup>1</sup>**

Diana I. Pérez

## RESUMEN

El objetivo del trabajo es analizar la relación entre la experiencia fenoménica y nuestra conceptualización folk de ella. Me concentraré en la estrategia de los conceptos fenoménicos como respuesta al rompecabezas de María. En la primera parte presento el argumento de María y la estrategia de los conceptos fenoménicos. En la segunda parte explico cuáles serían los requisitos que los conceptos fenoménicos deben satisfacer para responder al problema de María. En la tercera parte presento varias teorías acerca de los conceptos fenoménicos y muestro las dificultades que presentan. Finalmente, desarrollo mi propia teoría de los conceptos fenoménicos. Mi tesis es que los conceptos fenoménicos son complejos, ya que sus condiciones de posesión dependen de la maestría de otros conceptos, algunos muy sofisticados como la distinción apariencia-realidad (que pertenece a nuestra teoría de la mente), y los conceptos de color (al menos en el caso de los conceptos necesarios para dar cuenta del caso de María). Y estos últimos conceptos tienen peculiares condiciones de posesión: incluyen el uso de capacidades reconocitivas no conceptuales.

**PALABRAS CLAVE:** *argumento del conocimiento; conceptos fenoménicos; conceptos de color; teoría de la mente.*

## ABSTRACT

The aim of this paper is to analyze the relationship between phenomenal experience and our folk conceptualization of it. I will focus on the phenomenal concept strategy as an answer to Mary's puzzle. In the first part I present Mary's argument and the phenomenal concept strategy. In the second part I explain the requirements phenomenal concepts should satisfy in order to solve Mary's puzzle. In the third part I present various accounts of what a phenomenal concept is, and I show the difficulties each of them have. Finally, I develop my own account of phenomenal concepts. My thesis claims that phenomenal concepts are complex concepts whose possession conditions depend upon the mastery of many other concepts, in fact, quite complex concepts such as the distinction between appearance and reality (which belongs to our theory of mind system), and color concepts (at least in the case of the phenomenal concepts needed in order to account for Mary's case). And these later concepts are concepts that have special possession conditions: they include the deployment of nonconceptual recognitional capacities.

**KEYWORDS:** *Knowledge Argument; Phenomenal Concepts; Color Concepts; Theory of Mind.*

To begin by teaching someone “That looks red” makes no sense. For he must say that spontaneously once he has learnt what red means...

*Zettel*, § 418

Why doesn't one teach a child the language-game “it looks red to me” from the first? Because it is not yet able to understand the rather fine distinction between seeming and being?

*Zettel*, § 422

How do I know that this colour is red? -It would be an answer to say: “I have learnt English”

*Philosophical Investigations*, § 381

## INTRODUCTION

Nobody has said that what is more basic and primitive phylogenetically and ontogenetically should be easier to conceptualize. In fact, I will develop just the opposite idea in this paper: a fitting description of simpler phenomena needs more complex and sophisticated conceptual resources than complex ones. At this time, my target will be phenomenal experience and our folk conceptualization of it.<sup>2</sup> In order to develop my ideas I will analyze the discussion around phenomenal concepts,<sup>3</sup> the concepts postulated by the physicalist in order to refute the most famous arguments for qualia. In this paper I will focus my attention on one single argument: Jackson's Mary. But my aim here is not mainly to argue for physicalism (or against dualism) but to explore the way in which we conceptualize our own minds, in this case, how we conceptualize our subjective experience of colors, assuming a physicalist framework as the defenders of the phenomenal concept strategy do. The analysis of Jackson's arguments and the answers given invoking phenomenal concepts are just a philosophical resource useful to state my own position. I will show the limitations of Jackson's argument and, in so doing, I will develop an alternative way to solve Mary's puzzle focusing on our folk conceptual resources. I claim that a proper understanding of our folk psychological capacities would be the key to answering these anti-physicalist arguments.

The structure of the paper is the following. In the first part I present Jackson's argument and the phenomenal concept strategy, one of the most popular physicalist answers to this argument. In the second part I explain the requirements phenomenal concepts should satisfy in order to solve Mary's

puzzle. In the third part I present different accounts of what a phenomenal concept is, and I show the difficulties each of them have. Finally, I develop my own account of phenomenal concepts. I claim that phenomenal concepts are not a pure label for our phenomenal experiences, as many defenders of these notions are led to hold. Rather, they are complex concepts whose possession conditions depend upon the mastery of many other concepts, in fact, quite complex concepts such as the distinction between appearance and reality (which is part of our theory of mind system), and color concepts (at least in the case of the phenomenal concepts needed in order to account for Mary's case). And these latter concepts, i.e. color concepts, are recognitional concepts that is, concepts that have special possession conditions: they include the deployment of nonconceptual recognitional capacities.<sup>4</sup> This peculiarity of color concepts is what explains the puzzle produced by the case of Mary.

### I. MARY'S PUZZLE

According to Jackson (1982), (1986) Mary has spent her entire life in a black-and-white room without the possibility of seeing any color. She becomes obsessed with the problem of the perception of colors and decides to study all there is to know about the neurophysiology of color perception. Hypothetically, Mary lives in a not too distant future when physics is a complete science and has the conceptual resources to describe all physical facts down to the smallest detail. But, unexpectedly, Mary is released from her black-and-white room and she perceives a ripe tomato for the very first time. Does she learn anything? Does she come to know something that she did not know beforehand? Jackson's intuition (and almost everyone else's)<sup>5</sup> is that she really learns something. But since she already knew all the physical facts, she has to learn a non-physical fact. Hence, physicalism is false.

Before examining the answer to this puzzle, I want to make some remarks about the thought experiment. They concern Mary's situation. The conditions under which the experiment is set out presuppose that:

1. Mary has normal vision, and is only deprived of colored experiences. But she can see black and white (and grey in different hues) things. Hence she knows what it is like to experience grey (or black or white), and she can have thoughts about these experiences.
2. According to Jackson 1982, Mary uses the words "red" and "blue" exactly like us. But, let me remark that there is a difference; she has no color experiences and hence she cannot apply color terms based on her visual experiences.

3. According to Jackson 1982, Mary learns two different things when she experiences the tomato for the first time: (1) she learns something about the world (she learns what red is), and (2) she learns something about our visual experience of the world (she learns what it is like to experience red). In the literature surrounding Jackson's Mary, authors stress the second fact and fail to acknowledge the first. Yet it is important that Mary gains knowledge both from the outside world and from within herself. I will return to this point below.

One of the more widespread physicalist answers to Jackson's argument is the phenomenal concept strategy or perspectival strategy developed, among others, by Loar (1990), Lycan (2002), Papineau (2002), (2007), Tye (2003), Levin (2007), Balog (2002), Aydede and Güzeldere (2005) and MacDonald (2004). According to these supporters of the phenomenal concept strategy, there are special kinds of concepts – pure phenomenal concepts –, which cannot be possessed before having the experiences referred to by them. If experiences are physical states, as the physicalist claims, then phenomenal concepts are concepts that refer to physical properties from a different mode of presentation.<sup>6</sup> But these concepts have a peculiar nature according to which only those who have undergone the specific experiences are able to grasp them. Given that Mary cannot possess them before being released from her black-and-white room (because of the very nature of these special kinds of concepts) and that they refer to physical properties, Jackson's intuition that Mary learns something holds without abandoning physicalism. Mary learns something because she acquires new concepts and she can thus entertain new propositions in her mind. However, this new epistemic situation does not have any ontological consequences.

In this paper I discuss this type of response to Mary's puzzle. I defend the idea that there is a grain of truth in the claim that we need a special kind of concept, the so-called "phenomenal concepts", to solve Mary's puzzle. I agree with Jackson's idea that Mary learns something when released. This new knowledge involves new thoughts that Mary can think, and that she could not think before. So, if she can think new thoughts, then the only plausible explanation is that she acquired new concepts. The opposite option would be to hold that what happens to Mary is that, when released she can apply old concepts to new facts. But given the fact that I do not concede that there are new facts to be learned, I hold that she acquires new concepts for the old facts.

The problem we have to face now is that phenomenal concepts are difficult to characterize; there are many different ways of understanding them. In the next section I consider the conditions these concepts should satisfy in order to solve Mary's puzzle.

## II. PHENOMENAL CONCEPTS AS A PHYSICALIST ANSWER TO MARY'S PUZZLE

The central idea behind the phenomenal concept strategy is to separate the conceptual issue from the ontological one. When Mary is released she learns something she did not know before, because she acquires new concepts (about old properties), and not because there is some obscure and physically inaccessible phenomenal fact that Mary now has access to. These new concepts are the so-called "phenomenal concepts".

However, many disagreements exist among the defenders of the phenomenal concept strategy in answer to Mary's argument. In the first place no agreement has been reached about the extension of the notion "phenomenal concept". Almost everyone includes examples such as SENSATION OF RED or EXPERIENCE OF RED [Loar (1990), Aydede and Guzeldere (2005), Tye (2003)].<sup>7</sup> But, some also include concepts such as RED applied to external things [MacDonald (2004), p. 507] in the list of phenomenal concepts.<sup>8</sup> Others offer a more subtle distinction between different phenomenal concepts; for example, Stoljar (2005) claims that RED SENSATION, SENSATION THAT ONE GETS FROM LOOKING AT RED THINGS and SENSATION THAT REPRESENTS THINGS AS RED are three different concepts. And Chalmers (2002), pp. 223-229, proposes distinguishing four different phenomenal concepts included in beliefs such as Mary's when she says to herself *I think I am having an experience of such-and-such*. These concepts are:

- (1) The community relational concept or RED<sub>C</sub> = the phenomenal quality typically caused in normal subjects within my community by paradigmatic red things.
- (2) The individual relational concept or RED<sub>I</sub> = the phenomenal quality typically caused in me by paradigmatic red things.
- (3) The demonstrative concept that we can call E, referred to by expressions such as "this quality" or "this sort of experience".
- (4) A pure phenomenal concept, called R, one that does not pick out our phenomenal redness relationally, but rather picks it out directly in terms of its intrinsic phenomenal nature.

According to Chalmers, these four concepts refer to the same phenomenal quality Q, but (4) is the one involved in the solution to Mary's puzzle.<sup>9</sup>

Nevertheless, the extension of "phenomenal concept" is not the only problem. On the contrary, there is a second difficulty concerning the English expressions corresponding to these concepts.<sup>10</sup> Chalmers (2002), p. 240 claims that we do not have public language expressions that distinctively ex-

press the content of direct phenomenal concepts (that is why he proposes “RED<sub>C</sub>”, “RED<sub>I</sub>”, “E” and “R” as the linguistic labels of the different phenomenal concepts he distinguishes). In fact, he acknowledges that there are four concepts involved in the belief expressed, in English terms, as “I am having a sensation of red”. So, we have no one-to-one correspondence between public language and concepts (at least in the case of phenomenal concepts). But, not everyone is as explicit as Chalmers. And there is nothing in Loar (1990), MacDonald (2004), Aydede and Güzeldere (2005), Tye (2003) or in any other author I have read which suggests that they think that more than one phenomenal concept exists behind the English expression “sensation of red”. Even Stoljar (2005), who adopts the same identity criterion for concepts as Chalmers, chooses different English expressions to express the different concepts, although all of them are complex ones. In any case, everyone agrees that there is no single word in English used to express a phenomenal concept.<sup>11</sup> All the examples include at least two English terms, a color term (“red”, for example) and a psychological term (such as, “sensation” or “experience”). My proposal will be that this complex character of English expressions reflects the complex character of the phenomenal concept involved.

Leaving aside these differences among the defenders of the phenomenal concept strategy, everyone would agree, in my opinion, on the adequacy conditions that phenomenal concepts should fulfill in order to solve Mary’s puzzle. These adequacy conditions are the following.

1. Phenomenal concepts are concepts we appeal to in everyday reflecting on human mental lives. When analyzing Mary’s case we are concerned with the usual way in which ordinary people think about their own (and others’) thoughts. There is no need to appeal to scientific or philosophical special vocabulary to understand our thoughts.
2. Phenomenal concepts are perspectival. By perspectival I understand that they involve essentially the first person point of view and they are not purely theoretical. In my opinion this is not a specific feature of phenomenal concepts but a general feature of all the concepts we use in order to think about our and others’ minds. As is well known, all psychological concepts have different criteria of application in the first person case than in the third person.
3. Phenomenal concepts refer to physical properties. This requirement is applicable only to those defenders of phenomenal concepts who adopt this strategy in order to give a physicalist answer to Mary’s puzzle (as I am trying to do). If someone – like Chalmers 2002 – thinks that there are phenomenal concepts but adopts dualism, this requirement does not apply to him.

4. Phenomenal concepts are explainable in a way compatible with physicalism. The idea I am attempting to capture with this requirement is that the explanation given for these concepts cannot depend upon the postulation of non-physical entities like qualia. Otherwise the explanation of phenomenal concepts will not be appropriate to defend a physicalist view.
5. Phenomenal concepts should be conceptually irreducible to physical concepts. This requirement is at the core of the intuition that Mary learns something new when released. If phenomenal concepts were derivable from physical ones, then Mary would have been able to entertain thoughts with such phenomenal concepts while captive, because she already had all the physical concepts and theories about the physical processes going on.
6. Phenomenal concepts are recognitional concepts. The idea is that experience is constitutive of the possession conditions of this kind of concept. In other words, without experience the thinker cannot entertain thoughts including this specific kind of concept. In my opinion this experience can be understood as a set of physical nonconceptual states.

### III. ACCOUNTS OF PHENOMENAL CONCEPTS AND THEIR DIFFICULTIES

Despite the fact that defenders of the phenomenal concept strategy generally agree to the above adequacy conditions, there are many varied theories of phenomenal concepts. In this section, I divide the most common theories into two major groups. The first group treats phenomenal concepts as atomic concepts, as picking out qualitative properties [Chalmers (2002)], or as atomic labels of nonconceptual representations, i.e. as specific kinds of physical entities [Tye (2003)]. The second major group treats phenomenal concepts as more complex concepts, using one of the following three sub-strategies. The first sub-strategy, whose adherents include Loar (1990), Lycan (2002), and Levin (2007), treat phenomenal concepts as a kind of demonstrative concept with the general form “That kind (of experience)”. The second sub-strategy includes “constitutivists” like Balog (2002), and “quotational” adherents like Papineau (2002). Constitutivists, as the name would suggest, hold that experiences are constitutive of phenomenal concepts. Papineau treats phenomenal concepts as “compound terms, formed by entering some state of perceptual classification or re-creation into the frame provided by a general experience operator ‘The experience: —’” [Papineau (2002), p. 116]. The third sub-strategy is my own.<sup>12</sup> In my view, we can conceive of phenomenal concepts as complex, constituted by a color concept and a psychological concept (in the case of the phenomenal concepts needed in order to solve Mary’s

puzzle). I will also remark on some differences between these kinds of concepts and other phenomenal concepts (such as PAIN), and I will show that Jackson's puzzle can not be generalized to every kind of phenomenal concept.

I begin with the atomic view of phenomenal concepts. Chalmers' position is perhaps the most problematic. I have argued elsewhere against this view [Pérez (2005),( 2010)]. But we can leave aside this proposal in this context because it is not intended to fit a physicalist world view, because his account explicitly denies requirement 3. and 4.

Tye (2003), for his part, defends a physicalist position. According to Tye (2003), Mary's beliefs outside her black-and-white room are beliefs composed by phenomenal concepts that are mental representations with no structure, either definitional or of another kind, and with no associated reference fixing descriptions. Tye offers the following argument to prove this point: if phenomenal concepts were structured, their parts/constituents should either be phenomenal or physical. If they were phenomenal, then there would be a vicious circle, and, according to Tye, physicalism would turn out to be false, that is, the requirement 4. would be false. But if they were physical, the requirement of conceptual independence of phenomenal concepts would be violated, and, hence, the requirement 5. would be false.<sup>13</sup> Tye concludes, "Phenomenal concepts refer directly. They have no associated reference-fixers, no descriptive content at all. For these concepts, the reference is presented without the assistance of associated features distinct from the referent which the thinker a priori associates with it" [pp. 5-6]. But in the case of phenomenal beliefs (as in any other kind of belief) not only the content (that is, reference, in Tye's case because he is an externalist about content) but also something else (the functional role) is involved in the identity of thought types. The functional role of the same content –experience of red– is different when it is given a physical description of the phenomenal state –say, the experience of red described with the theoretical concepts Mary possesses while in her room– from the one that Mary gives using phenomenal concepts on being released.<sup>14</sup>

Tye (2003) view has two main flaws. First, it could be argued that the entities postulated by Tye are not really concepts. This is because, as Tye (2009), p. 54 remarks, this account like any other externalist account lack fine grain individuation conditions. And if fineness of grain is necessary for being a concept, these are not in fact concepts. Furthermore, these concepts are not linked to the right kind of abilities [Prinz and Clark (2004)]. Let me explain. Let us suppose that Mary is released and while experiencing something red for the first time, she automatically learns the new concept SENSATION OF RED. Does she have the ability to apply her new concept to other cases, to imagine red sensations, and so on *ipso facto*? This kind of ability, like any other ability, seems to be the kind of thing that we acquire over a period of time and training, not just instantaneously, upon being exposed once to a certain stimulus. So, if a concept is something linked to a cer-



tain kind of ability, Tye's phenomenal concepts are, again, not concepts. So, there are various reasons to call into question the idea that these phenomenal concepts are concepts in their own right. Requirement 1.- is, hence, violated.

In the second place Tye (2009), criticizing his own previous view, argues that this account of phenomenal concepts does not solve Mary's puzzle in an adequate way from a physicalist point of view. The view violates requirement 4. His point is that if physicalism is true, Mary does not learn anything new because she only adds a label to a property she already knew by physical description while in her black-and-white room. I would like to press this idea a little bit further: If the only thing Mary gets while out of the room is a new label of an old property, what explanation can we give in order to prove that she could not have learned that very same label inside the room, before her liberation? If she can know the physical property by means of theoretical concepts, then she could have decided to label it in whatever way she wanted, including the "phenomenal concept" label "R", while still inside the room. And if not, there is no explanation of why she could not.

The second group of theories of phenomenal concepts understands them as being complex. According to Loar (1990), Lycan (2002), Levin (2007) among others, they are a kind of demonstrative concept, with the general form "That kind (of experience)". But this strategy also has many problems. First of all, just because we often use a demonstrative expression in public language to express the kind of concept in question, it does not follow that the concept itself must be demonstrative [Abath (2005), Papineau (2007)]. In the second place, phenomenal concepts are not like demonstratives because they lack "character" in Kaplan's sense, and are without contextual meaning insofar as their content is the same in all contexts [Papineau (2007)]. In the third place, as in the case of Tye's (2003) proposal, a demonstrative account of phenomenal concepts cannot explain the new abilities Mary acquires when released. After seeing the tomato, she can recreate the experience of it in her imagination, and classify other experiences as being of the same or different type. Her new abilities are not explained by a demonstrative account for the very reasons I stated above. Finally, a single color experience might be picked out by a demonstrative phenomenal concept, but many different phenomenal concepts could be involved in it: the concept EXPERIENCE OF RED, the concept THIS PARTICULAR SHADE OF RED, the concept EXPERIENCE OF SCARLET, or the concept EXPERIENCE OF COLORED THINGS could all be at play. The demonstrative account of phenomenal concepts also lacks fineness of grain.<sup>15</sup>

Only the constitutivist view of phenomenal concepts remains for our examination. Papineau (2002) defends a "quotational" view of these concepts. He says phenomenal concepts are "compound terms, formed by entering some state of perceptual classification or re-creation into the frame provided by a general experience operator 'The experience: —'" [Papineau (2002), p. 116].

Balog (2002) holds something similar; the phenomenal state, or the recreation of it in the imagination, is constitutive of the phenomenal concept. These constitutive accounts also suffer from many difficulties.

In the first place, I think Crane's (2005) point against Papineau's explanation of phenomenal concepts is right. He holds: "[...] it seems to me incredible that when we think about, say, pain, one must, as a necessary part of that very act of thinking, have an experience which *in any way* resembles pain" [Crane (2005), p. 156]. It seems incredible to me, too [and also Tye (2009), p. 47]. There is nothing painful in pain thoughts. What's more, the thought "I am not having this experience E now" would turn out to be a contradiction on the constitutivist account. I would have to have the experience E in order to think about it, even if I am thinking that I do not have E.

Tye (2009) offers a couple of additional objections to the constitutivist view. First, he points out that if I see something with a particular hue of red, then I have the experience of red, and of scarlet, of color, etc. And yet, I cannot individuate the different concepts EXPERIENCE OF RED, EXPERIENCE OF SCARLET, EXPERIENCE OF COLOR, because there is no difference between them. The very same experience constitutes them all. We have only one experience but many different phenomenal concepts. This version of phenomenal concepts, again, lacks proper fineness of grain. Second, there is a problem with the general concept PHENOMENAL PROPERTY. Do we have to activate all our phenomenal states in order to think in general about them? Which particular experience do we have to activate in order to have this general thought? And if I think about my phenomenal color experiences in general, do I have to activate all my past color experiences? It is on the basis of these difficulties that I reject the constitutivist view.

#### IV. MY PROPOSAL ON PHENOMENAL CONCEPTS FOR MARY'S CASE

I have tried to show that the characterizations of phenomenal concepts that are usually given to solve Mary's puzzle are flawed. But, denying that a new concept is involved in the solution to Mary's puzzle seems to leave us without an explanation of the intuition about Mary's epistemic gain when released.<sup>16</sup> If she learns something, if she knows something that she did not know beforehand (though she already knew all the physical facts), either there is a new fact (a phenomenal fact) to know (and hence, physicalism is false) or there is a new concept (a phenomenal concept) for an old property (a physical property) she already knew under a physical description, and physicalism holds.

So, it seems that the solution to Mary's puzzle is to give an account of the concept SENSATION OF RED, which fulfills the constraints proposed for phenomenal concepts and preserves the persistence and connection to the rest

of our conceptual life needed for concepts, while including recognitional capacities in their grounds in order to respect the intuition that there is something we cannot know without experience –something we cannot know only through theoretical means.<sup>17</sup>

Before detailing my proposal, I want to make two more remarks about Mary's situation:

- (1) If Mary really learns two things when released (as Jackson 1982 acknowledges and I highlighted in the first section), then Mary should learn two new kinds of concepts, instead of only one. She learns our *color concepts* (such as RED), those that allow her to gain more knowledge about the world, and *phenomenal concepts* (such as SENSATION OF RED), those that allow her to gain more knowledge about the human mind.
- (2) Before she was released, however, Mary already possessed *psychological concepts* (i.e. all the concepts we use in order to think about our own minds) because, *ex hypothesi*, the only difference between Mary and us is that she did not experience colors. Remember, as I stated in the first section, that she has her “theory of mind” intact, and she can think about what she believes and desires, and what it is like to taste chocolate, smell a rose, hear music, and also what it is like to experience black, white or grey. And she also has the ability to distinguish appearance from reality.

Given these remarks, my proposal is as follows: Mary in her black-and-white room does not have the concept EXPERIENCE OF RED, as everybody accepts, insofar as she does not have our concept RED. And when released, she begins to understand our concept RED and, it is only when she combines this new concept with the psychological concept of EXPERIENCE, or SENSATION, a concept she already had, that she also acquires the new concept SENSATION OF RED [See above *Zettel*, § 418].

It is important to note that this explanation does not work for concepts of bodily sensations such as PAIN. Because in this case we cannot say that there are two things that injured Mary would learn when she is in pain for the first time. In the case of colors, as I said before, there are two things to be learnt: (1) something about the world, what red is, in the end the contribution that the world makes in order that a normal subject can experience that color, and (2) something about the experience of color that the subject has, something about the subject. This double life of colors, as something in the object but also in the subject is what explains that there have to be certain normal conditions satisfied in order to perceive certain things as red instead of white. So, in the case of colors there exists the possibility of wondering whether

something that looks like white to us is really (under normal conditions) red. But this doubt cannot be raised in the case of pain. If it seems to me that I am in pain, then I am in pain, and *vice versa*. Hence my proposal cannot be generalized to all the phenomenal concepts we have, because it does not work for the concepts of bodily sensations.<sup>18</sup>

But it does not work for some other cases either, where the object of experience is not a color, but a physical object, for example. To see this point, we can consider another Mary, changing the impossibility to see colors for other impossibilities, such as the impossibility to see a specific kind of physical object, for example, elephants, or a specific object such as the Eiffel Tower, or a specific person such as the Pope.<sup>19</sup> It seems clear to me that there is something special about colors that prevents us from posing a similar problem when we consider another Mary, deprived of visual experiences the experience of seeing an elephant, i.e. in the case of someone who has never seen one not even in pictures, photos, or TV. A person deprived of visual experiences of elephants but knowing everything there is to know about an elephants' physicality, could recognize an elephant when in front of one of them for the first time in her life, and moreover, in my opinion this person can have elephant-experiences before being released of his deprived-of-elephants room because she can imagine an elephant using the physical information she acquires. And something similar occurs *mutatis mutandi* in the case of the Eiffel Tower or the Pope. But Mary cannot imagine the peculiar hue the color red has before experiencing red for the first time. There is something special about colors in Jackson's argument that I believe has been ignored in the discussions of phenomenal concepts. The conclusion of these alternative thought experiments is that Jackson's argument only works easily and intuitively with colors, which are peculiar entities in their own right.<sup>20</sup> But it does not work in the case of ordinary physical objects. I want to make it perfectly clear that this is not so much a limitation in my account, as it is a limitation in the scope of Jackson's original argument.

My thesis depends upon the idea that Mary did not have our concept RED (i.e. the concept RED we, normal adult human beings, have) while in her black-and-white room. In effect, some concepts, for example, color concepts are grounded in perceptual experiences and, hence, people cannot possess a color concept without having had the corresponding color experience at least once in their lives. RED is, without a doubt, a recognitional concept. There is a sense in which a person born blind will not have *our* color concepts (That is why I believe that MacDonald (2004) takes the right line when she analyzes Mary's case in the light of Molyneux' problem).

According to Jackson, one of the conditions of Mary's experiment is that Mary uses correctly the words "red" and "blue". But as I said above, this is only partially true. It is true that, while in her room, she uses the words correctly. There she is only able say general things about colors, such as "blood

is red” or “the sky is blue”. She can only make what can be called theoretical assertions about colors. But in a deeper sense she still cannot use color words correctly, because when released she cannot apply them correctly. In order to see my point, let us consider yet another Mary. Suppose now that this Mary is released in front of a white wall with a number of colored squares, and everything else around her is still black-and-white.<sup>21</sup> Assuming that she has no special devices in order to look inside her own brain, i.e. if we consider only the personal level abilities that Mary deploys, it becomes clear that she will not be able to tell us the color of each of the squares. All she knows about colors, her complete theoretical understanding of color concepts, is not enough. She does not know which color terms apply to which squares. She might be able to group all the red squares and the yellow squares, but she will not be able to say which of the sets of similar squares is the red one [see also Nordby (2007)]. Clearly then, captive Mary does not possess the same color concepts as we do, and the problems Mary would have with the proper use of color terms is a result of this fact. A person born blind can have some beliefs, such as, tomatoes are red or grass is green, but in my opinion we should say that such a person only has a partial understanding of the color concepts involved in these beliefs, in a similar way to which Burge’s Bert had a partial understanding of arthritis.<sup>22</sup>

We must know many things to have a full, working understanding of a concept. In the case of colors concepts, we must possess different abilities in order to possess them: we should have certain inferential/theoretical capacities, some general knowledge about the world, such as the belief that lemons are yellow, or that the bullfighter’s cape is red, and we also have to possess some nonconceptual (perceptual) discriminative capacities. If I am right, we need to know the following things in order to possess a color concept like RED:

- (1) We must know how to discriminate red things from non-red things. This is what a man born blind and Mary in her black-and-white room cannot do.<sup>23</sup> This is a cognitive ability based on our perceptual capacities and it is dependent upon the fact that we have nonconceptual experiences.<sup>24</sup>
- (2) We should know what objects are paradigmatic cases of RED; for example, we should have some beliefs such as the beliefs that ripe tomatoes, blood, and the bullfighter’s cape are red.<sup>25</sup>
- (3) We should know that under normal circumstances (with normal eyesight and under normal illumination) all human beings have the same color experience while the paradigmatic objects are present, at a reasonable distance, with the right light, and so on.<sup>26</sup>

The above conditions are for a full, working understanding of the concept RED.<sup>27</sup> If any of these conditions are not present, we cannot have a full, working understanding of the concept. On my analysis, Jackson's Mary fails to satisfy (1) but fulfills requirements (2) and (3). And therefore, she has only a partial understanding of RED. Given that (1) is constitutive of our concept RED and that it depends upon having undergone the corresponding experiences, RED is a recognitional concept after all.<sup>28</sup> Note that Mary's case is not exactly like Bert's. Bert lacked scientific knowledge about the concept ARTHRITIS and he had to trust a certain expert community in order to acquire a full understanding of the concept. Mary, on the other hand, needs no information from any expert whatsoever. She lacked the usual experience that, in normal cases, is present when we learn color concepts. This is why Jackson's Mary cannot be but a thought experiment. There is no way in which we can learn the color concepts and correctly use the color words in all contexts before having color experiences, as the case of the additional Mary faced with only red colored squares in her black-and-white room shows. In my opinion this is exactly what is meant by the idea that a given concept is "recognitional".

Now, let us suppose that, as I suggested above, phenomenal concepts like RED SENSATION are complex concepts produced by the combination of a color concept and a psychological concept. It seems natural to suppose that the special characteristics of each of the components would be inherited by the complex concept. So if RED is one of the components of SENSATION OF RED and RED is a recognitional concept, then SENSATION OF RED should be so, too.<sup>29</sup>

And if (1) is, as I believe, one of the possession conditions of RED, then Mary before being released, Molyneux's blind man, and the achromatopsic philosopher, will not possess our concept RED, and therefore, any concept constituted by it. There is a sense, though, in which Mary learns a new concept when released, our concept RED and all the concepts depending upon it, in particular SENSATION OF RED.<sup>30</sup> And this complex concept fulfills all the requirements I mentioned in Section II for phenomenal concepts:

1. They are used when we reflect upon our mental lives,
2. They are perspectival, i.e. they essentially involve a first person point of view, because they are composed of a psychological concept and all psychological concepts are perspectival,<sup>31</sup>
3. If physicalism is true, they might refer to a physical property (at least, nothing in my account prevents it),
4. They can be explained in a way acceptable to the physicalist, if psychological concepts and color concepts in general are similarly explained (a point which is not questioned in Jackson's argument),

5. They are conceptually irreducible to physical concepts, because of (1): there is a perceptual nonconceptual capacity involved which should be exercised at least once in order to possess these concept, and
6. They are recognitional, because RED is.

*Departamento de Filosofía  
Universidad de Buenos Aires-CONICET  
Bulnes 642 (CP. 1176), Buenos Aires, Argentina  
E-mail: dperez@filo.uba.ar*

<sup>1</sup> I presented these ideas in various places: the APA Conference, Special Session on Contemporary Philosophy of Mind in Latin America and Spain, hosted by the APA's Committee on International Cooperation, Chicago, April 17-20 2008; the Philosophy Department at Kansas State University, April, 21th 2008; the V Encontro internacional de filosofia da mente, Joao Pessoa, Brazil, May, 6-9th 2009; the Instituto de Investigaciones Filosóficas UNAM, June 1st, 2009; the Universidad Autónoma de Madrid, November 12th, 2009. I am grateful to all these audiences for their comments and suggestions and especially to Peter Simmons, Amy Lara, Agustín Rayo, Maite Ezcurdia, Jesús Vega, and my students at the graduate seminar: "The problem of consciousness and phenomenal concepts" at the University of Buenos Aires. I am also grateful to Eamonn McDonagh who helped me with the English version, and to the anonymous referees from **teorema** and Ángel García Rodríguez, for helpful suggestions. Research for this work was partially funded by the University of Buenos Aires (UBACyT F-130), CONICET (PIP-2531), and the Spanish Ministry of Science and Innovation (C-Consolider HUM 2006-08236 and Consolider Ingenio CSD2009-0056).

<sup>2</sup> I use the expression "folk (psychological) conceptualization" in order to point to the usual way in which normal adult human beings understand their own and others minds. This folk conceptualization involves the deployment of psychological concepts such as belief, desire, fear, thought, being in love, experience, etc. Philosophers and scientists usually create new concepts in order to formulate coherent theories about different phenomena. But I will only be concerned with the ways in which ordinary people think about their own and others minds.

<sup>3</sup> By "phenomenal concepts" I refer to those concepts used in order to conceptualize our subjective experience. Their very nature will be the main issue in this paper.

<sup>4</sup> I assume that each concept has (essentially) its own possession conditions and that if a concept has a complex structure, the possession conditions of its simpler parts are involved in the possession conditions of the complex concepts (or their categorical bases if there are any).

<sup>5</sup> The exceptions are Churchland (1986) and Dennett (1991), (2007).

<sup>6</sup> In case we hold that phenomenal concepts have modes of presentation. Some accounts of phenomenal concepts do not associate modes of presentation with phenomenal concepts (see below, for example, Tye's (2003) proposal).

<sup>7</sup> As usual, names of concepts are written in small capitals.

<sup>8</sup> I will return to her proposal in Section IV. In particular, see footnote 28.

<sup>9</sup> As is well known, Chalmers defends a naturalistic dualism, holding that qualia are real and irreducible to the physical. But for our purposes, this does not matter. In his 2002 he tries to specify the nature of phenomenal concepts, that is, of the concepts needed in order to account for the new knowledge Mary gains when released. While trying to understand the nature of phenomenal concepts it does not matter if these concepts refer to the physical or, as Chalmers holds, to irreducible qualitative features. It will only be relevant when we try to use these concepts to undermine arguments for dualism. If we replace, in Chalmers' definitions, the expression "the phenomenal quality..." by "the state..." we will get a neutral definition for the phenomenal concepts mentioned. If this replacement can be made without assuming qualia realism, the result could be useful for physicalist purposes.

<sup>10</sup> The same problem exists in other languages, like French or Spanish, and it seems to be a general fact about the expression of a phenomenal concept in natural languages, as far as I know.

<sup>11</sup> Excluding MacDonald and anyone else who assumes that RED is itself a phenomenal concept. But I do not think that RED is a phenomenal concept. Maybe PHENOMENAL RED is, but this concept is complex too.

<sup>12</sup> For developing my own view I was inspired by Carruthers's (2005) second order theory of phenomenal consciousness which appeals to the theory of mind in order to give an account of phenomenal consciousness. But my view is compatible with a first order theory, because I consider some concepts belonging to the theory of mind as constituents of our phenomenal concepts (and not of phenomenal consciousness itself). For a more detailed exposition of these differences see Pérez (2008).

<sup>13</sup> My own account, as we will see, can be at first sight included in the second horn of the dilemma. But see footnote 27.

<sup>14</sup> We can now ask: How many phenomenal concepts for the same phenomenal state can Mary learn when released? Why should we suppose that there could be *more than one* new phenomenal concept directly associated with the new phenomenal state Mary is entertaining? How can Mary individuate the four different phenomenal concepts Chalmers mentions if they all refer directly to the same phenomenal state that Mary entertains? In my opinion, those who wish to defend an account of phenomenal concepts along these lines should claim that there is only one new phenomenal concept associated with each new experience (-type) that Mary enjoys. The relevant phenomenal concepts that Mary acquires when released, the ones that could help solve Mary's puzzle within physicalism, seems to be pure phenomenal concepts, those concepts without any structure, which are simple labels that Mary attaches to her new experience while enjoying it: the R that Chalmers postulates.

<sup>15</sup> Or at least, the demonstrative account needs to offer an explanation of how to reach this fineness of grain. Levin (2007), for example, adds a functional role to the phenomenal concept. The difference with my account is that, once the functional role – or theoretical information as I prefer to say – is added, there is no need to appeal to



some special phenomenal demonstrative concepts, as we will see in the last section of this paper.

<sup>16</sup> Recently, Tye (2009) offered a different solution to this puzzle, without phenomenal concepts, adopting the distinction between knowledge by description and knowledge by acquaintance. In Mary's case, she does not lack any factual knowledge, but she lacks the appropriate acquaintance with red, a kind of "object" knowledge. In any case, Mary can entertain new thoughts after being released, and those thoughts are the ones I want to explain in this paper. Furthermore, if the ability hypothesis (Lewis 1988) turns out to be plausible, what Mary lacks is some kind of know-how, not knowledge-that. I will not discuss these proposals here. I assume that she is lacking a kind of factual knowledge.

<sup>17</sup> In Perry's words, we should accept that Mary's new knowledge is attached (and not detached) knowledge [Perry (2001), p. 137].

<sup>18</sup> It might be objected that it is possible to formulate a version of the knowledge argument according to which Mary knows all physical facts about pain but due to some physiological condition, she has never experienced pain herself. In this case - the objection follows-, it seems extremely intuitive to say that if one day she experiences pain for the first time, she would learn something new. My answer to this objection is the following (quite debatable, of course): the concept PAIN - as the concept RED - has among its possession conditions, a "recognitional" condition, because in normal conditions, people learn the concept PAIN applying it to themselves. It follows from my account of psychological concepts that those who have not been in certain states will not possess our concept of PAIN (I mean: the concept that normal adult human beings who have had pain states in their lives have). In other words, our everyday conception of pain is not fully theoretical. This is the reason why the knowledge argument can also be formulated for the case of PAIN, even if in this case there is no appearance-reality distinction to consider.

<sup>19</sup> The case of people is more similar to the color case than the others, though. Because in both cases the proper condition for classifying someone as "the same person that I saw before" or "the same color that I saw before" involves sub-personal mechanisms, i.e. nonconceptual recognitional capacities, for example, the mechanisms of face recognition, voice recognition or color recognition.

<sup>20</sup> In fact, these are not thought experiments at all. It is what happens to all of us all of the time. We are having new experiences all the time in our lives, and only some of them have the structure Mary's has.

<sup>21</sup> This alteration of Mary's case is similar to Nida-Rümelin (1998), but it has a different purpose.

<sup>22</sup> After developing these ideas, I read Knut Nordby (2007). Nordby was a person with achromatopsia, and like Mary, he became interested in the perception of colors. He holds that people with achromatopsia usually learn some facts about colors, memorizing them, such as "fire engines are red, violets are blue, grass is green and lemons are yellow" because he says: "This enables them to "know" color without experiencing the color hues" [Nordby (2007), p. 79], it was he who placed the quotation marks around the verb!), by discriminating similarities in brightness, a thing that Mary (and all normal people) will not be able to do, according to Nordby. I.e. achromatopsic people deploy alternative nonconceptual abilities to those used by people with normal sight in order to apply color terms.

<sup>23</sup> The reasons why the blind man and Mary-inside-the-room lack these capacities are obviously different. But they both lack the capacity to discriminate between the different colors, because they are both deprived of the relevant experiences.

<sup>24</sup> It could be objected that there can be different ways to perform this nonconceptual ability, in fact Nordby (2007) says that he can deceive people, because he can discriminate the colors of different objects not by the peculiar subjective experience they produce in us, but by differences in brightness and other physical characteristics that normal perceivers do not perceive. I assume that there is a normal condition in play here, and that “our” color concepts are those where the “normal” nonconceptual ways of discrimination are operating.

<sup>25</sup> As Nordby holds. It can be objected that it is implausible to hold (as I do) that someone who had only seen splashes of red paint and did not know what objects are normally red, would fail to know what it is like to see red. Of course, that person will know what it is like to see objects *like these*, or to see *that color*, or to see *things called “red”* (if the way in which color terms are learned is by showing red splashes and nothing else, i.e. no paradigmatic red objects). But we usually associate a wide body of knowledge and a set of paradigmatic emotional responses with each color concept, and all this knowledge is involved in the folk thoughts we can entertain when we think about what it is like to see red.

<sup>26</sup> In my opinion the four concepts that Chalmers distinguishes are not in fact four concepts but four parts of a single concept, our folk concept of SENSATION OF RED. And something similar occurs with color concepts. I think that this intuition is what Wittgenstein intends to say, when he enigmatically affirms: “If humans were not in general agreed about the colors of things, if undetermined cases were not exceptional, then our concept of color could not exist. No: – our concept *would* not exist” [Zettel § 351].

<sup>27</sup> On this account, the possession conditions of RED include a nonconceptual ability, so RED is not reducible to physical/theoretical concepts but to physical/theoretical concepts plus a certain ability (the ability to discriminate red things from other things by sight). And if a phenomenal concept is a complex concept constituted by a psychological attitude concept LOOKS or APPEARS TO ME and a color concept RED, then this phenomenal concept cannot be reduced to physical/theoretical concepts. Hence, everything is physical but not everything can be taught using the language of physics. (This is, as I read it, what *Philosophical Investigations* § 381 says). I do not think that this turns my position into a non-physicalistic one. The trouble is that the physicalist thesis that Jackson (1982) assumes is too narrow. My conception of physicalism is wider because I assume that there are some nonconceptual abilities which could be acceptable to a physicalist. In other words, phenomenal concepts as I understand them, are irreducible to physical concepts (i.e. to the concepts used in theoretical physics), but they are not irreducibly non-physical, i.e. they can refer to physical (including nonconceptual) phenomena such as our subjective experience (understood within a physicalist framework).

<sup>28</sup> MacDonald (2004) accepts not only that RED is a recognitional concept, which everybody accepts, but also that it is a phenomenal concept, and I am not able to understand why she does so. Nobody identifies phenomenal with recognitional concepts. It seems clear to me that there may be recognitional concepts that are not phenomenal. For a concept to be phenomenal, it should fulfill the six requirements I proposed. The first one was that phenomenal concepts are those used in reflecting on

our own minds. And I am not sure that RED is a concept used in our reflections about our mental lives. At least, it is not obvious that we need a concept like that. A further argument is necessary to include RED as a phenomenal concept.

<sup>29</sup> I assume that RED SENSATION is more similar to RED SQUARE, than to PET FISH, that is, that the abilities involved in the possession conditions of the components are inherited by the complex concept.

<sup>30</sup> I think my solution is better than MacDonald's (2004). She argues that Mary learns a new *conception* of an old *concept*, a concept she already had when she was inside the black-and-white room. So, in her opinion, Mary does not learn a new concept, but extends the application to new cases. I think that the intuition behind the thought experiment Jackson proposed is that there is something that Mary did not know inside the room, something that is inevitably learned through experience. My proposal makes room for this intuition more easily.

<sup>31</sup> I did not argue for this thesis here. But it is supported among others by Strawson's *Individuals*, and, in my opinion, Wittgenstein's *Philosophical Investigations*. In any case this is a general thesis about all our psychological concepts, not only about phenomenal ones. I leave the development of this idea to another occasion.

#### REFERENCES

- ABATH, A.J. (2005), "Nonconceptual Content, Fineness of Grain and Recognitional Capacities" *Abstracta*, 1, 2, pp.196-203.
- ALTER, T. and WALTER, S. (2007), *Phenomenal Concepts and Phenomenal Knowledge*, Oxford, Oxford University Press.
- AYDEDE, M. and GÜZELDERE, G. (2005), "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness", *Noûs*, 39:2, pp. 197-255.
- BALOG, K. (2002) "The 'quotational' Account of Phenomenal Concepts; unpublished.
- CARRUTHERS, P. (2005), *Consciousness. Essays from Higher Order Perspective*, Oxford, Oxford University Press.
- CHALMERS, D. (2002), "The Content and Epistemology of Phenomenal Beliefs" in Smith, Q and Jovic, A. (2002) *Consciousness. New Philosophical Perspectives*, Oxford, Oxford University Press.
- CHURCHLAND, P. S. (1986), *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge (Mass.), MIT Press.
- CRANE, T. (2005) "Papineau on Phenomenal Concepts", *Philosophy and Phenomenological Research*, vol. LXXI, no. 1 pp.155-162.
- DENNETT, D. (1991), *Consciousness Explained*, London, Little & Brown.
- (2007), "What Robo Mary Knows", in T. Alter and S. Walter. *Phenomenal Concepts and Phenomenal Knowledge*. Oxford, Oxford University Press, pp. 15-31.
- JACKSON, F. (1982), "Epiphenomenal Qualia", *Philosophical Quarterly*, XXXII 127, pp. 127-136.
- (1986), "What Mary Didn't Know", *Journal of Philosophy*, 83:5, pp. 291-5.
- LEVIN, J. (2007), "What is a Phenomenal Concept" in Alter, T and Walter, S (2007) *Phenomenal Concepts and Phenomenal Knowledge*, Oxford, Oxford University Press.

- LEWIS, D. (1988), "What Experience Teaches", reprinted in Block, B, Flanagan, O and Guzeldere, G (2002) *The nature of consciousness*, Cambridge (Mass.), MIT Press.
- LOAR, B. (1990), "Phenomenal States" reprinted in Block, N., Flanagan, O, and Guzeldere, G. *The Nature of Consciousness*, Cambridge (Mass.), MIT Press.
- LYCAN, W. (2002), "Perspectival Representation and the Knowledge Argument", in Smith, Q and Jovic, A. (2002) *Consciousness. New Philosophical Perspectives*, Oxford, Oxford University Press.
- MACDONALD, C (2004), "Mary Meets Molyneux: The Explanatory Gap and the Individuation of Phenomenal Concepts", *Noûs*, 38:3, pp. 503-524.
- NIDA-RÜMELIN, M (1998), "On Belief About Experience. An Epistemological Distinction Applied to the Knowledge Argument Against Physicalism", *Philosophy and Phenomenological Research*, Vol. LVIII, no. 1, pp. 51-73
- NORDBY, K (2007), "What Is This Thing You Call Color: Can a Totally Color-Blind Person Know About Color?", in Alter, T and Walter, S (2007) *Phenomenal Concepts and Phenomenal Knowledge*, Oxford, Oxford University Press.
- PAPINEAU, D. (2002), *Thinking about Consciousness*, Oxford, Oxford University Press.
- (2007) "Phenomenal and Perceptual Concepts", in Alter y Walter (2007), pp. 111-144.
- PÉREZ, D. (2005), "Mental Concepts As Natural Kind Concepts", *New Essays in Philosophy of Language and Mind, Canadian Journal of Philosophy* (supplemental volume 30), Stainton, R., Viger, C. and Ezcurdia, M. (eds.), University of Calgary Press, pp.197- 221.
- (2008), "Why Should Our Mind-Reading Abilities Be Involved in the Explanation of Phenomenal Consciousness? *Análisis Filosófico*, vol. XXVIII no 1, pp. 35-48.
- (2010), "Cómo pensamos acerca de nuestra experiencia consciente. Acerca de la idea misma de 'concepto fenoménico'", en Pérez, D., Español, S., Skidelsky, L. y Minervino, R. (comps.) *Conceptos. Debates contemporáneos en filosofía y psicología*, Buenos Aires, editorial Catálogos.
- PERRY, J. (2001), *Knowledge, Possibility and Consciousness*, Cambridge (Mass.), MIT Press.
- PRINZ, J. and CLARK, A. (2004), "Putting Concepts to Work: Some Thoughts for the Twenty First Century", *Mind and Language*, Vol. 19 No. 1, pp. 57-6.
- STOLJAR, D. (2005), "Physicalism and Phenomenal Concepts", *Mind and Language*, Vol. 20, no. 5, pp. 469-494
- TYE, M. (2003), "A Theory of Phenomenal Concepts", en O'Hear, Anthony (ed.) *Minds and Persons, Royal Institute of Philosophy Supplement 53*, Cambridge, Cambridge University Press, pp. 91-105.
- (2009), *Consciousness Revisited. Materialism without phenomenal concepts*, Cambridge MIT Press.
- WITTGENSTEIN, L. (1953), *Philosophical Investigations*, Oxford, Basil Blackwell.
- (1967), *Zettel*, Oxford, Basil Blackwell.



OXFORD

PENELOPE MADDY

# Defending the Axioms

ON THE PHILOSOPHICAL  
FOUNDATIONS OF SET THEORY

# CRITICAL PRAGMATICS

An Inquiry into Reference and Communication

Kepa Korta and John Perry



CAMBRIDGE