Croatian Journal of Philosophy
Vol. XVII, No. 52, 2018

Is the Antipathetic Fallacy Responsible for the Intuition that Consciousness is Distinct from the Physical?

FRANÇOIS KAMMERER*
Ecole Normale Supérieure, PSL Research University,
Institut Jean Nicod (ENS/EHESS/CNRS), Paris, France

Numerous philosophers have recently tried to defend physicalism regarding phenomenal consciousness against dualist intuitions, by explaining the existence of dualist intuitions within a purely physicalist framework. David Papineau, for example, suggested that certain peculiar features of some of our concepts of phenomenal experiences (the so-called "phenomenal concepts") led us to commit what he called the "Antipathetic Fallacy": they gave us the erroneous impression that phenomenal experiences must be distinct from purely physical states (the "intuition of distinctness"), even though they are not. Papineau's hypothesis has been accepted, though under other names and in different forms, by many physicalist philosophers. Pär Sundström has tried to argue against Papineau's account of the intuition of distinctness by showing that it was subjected to counterexamples. However, Papineau managed to show that Sundström's counterexamples were not compelling, and that they could be answered within his framework. In this paper, I want to draw inspiration from Sundström, and to put forth some refined counterexamples to Papineau's account, which cannot be answered in the same way as Sundström's. My conclusion is that we cannot explain the intuition of distinctness as the result of a kind of "Antipathetic Fallacy".

Keywords: Consciousness, dualism, physicalism, introspection, concepts, intuition.

Introduction

Many philosophers recognize that phenomenal consciousness seems to pose a metaphysical problem. On the one hand, we have various

* I would like to thank Samuel Webb and Joseph Levine for their comments and their help. I acknowledge the financial support of the two following grants: ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*

reasons to suppose that physicalism is true. Physicalism is the thesis that phenomenal states are fully identical with physical states (broadly construed, as to include physically realized functional states), such as brain states. On the other hand, the identity of phenomenal states and physical states appears very counter-intuitive. This is rendered manifest when we focus introspectively on one of our current experiences: how can *this* (say, this sensation of pain) be *the same thing* as some electrochemical activity that takes place in my brain?

Some dualist philosophers have argued that this intuition, once elaborated and transformed into arguments, simply shows that phenomenal states are *really* distinct from physical states. Others have tried to defend physicalism against this intuition, by giving an explanation of this intuition in a purely physicalist framework. For example, they have tried to show that this intuition is a by-product of certain (purely physical) features of some of our *concepts* of phenomenal states –concepts of phenomenal states which are notably applied through introspection, and which are called "phenomenal concepts".

One particular line of thought has emerged as especially popular: some philosophers, such as David Papineau (Papineau 1993, 2002, 2007), have tried to explain this dualist intuition (which Papineau labelled the "intuition of distinctness") as being the result of a peculiar feature of phenomenal concepts. These concepts, according to Papineau, display a "use/mention feature": whenever a subject uses them, she tends to *activate* the very experience thought about *via* this concept, or at least a "faint copy" of this experience. For this reason, when we think about phenomenal experiences *qua* experiences, *i.e.* with phenomenal concepts, our thought has a distinctive feeling, which it has not when we think about the same states *qua* brain states, using purely physical concepts. We then succumb to a fallacy that Papineau calls the "Antipathetic Fallacy", when we infer that this phenomenological difference between the two *thoughts* indicates that the *thing* thought about with a phenomenal concept *must be itself different* from the thing thought about with a purely physical concept.

Pär Sundström (Sundström 2008) addressed an objection to this physicalist account of the intuition of distinctness. He tried to show, on the basis of an imaginary counterexample, that this account cannot be correct, as it makes *false predictions*. It predicts that an intuition of distinctness should arise in a case in which it obviously doesn't. David Papineau answered this objection, by showing that it was possible to reinterpret Sundström's counterexample in order to make it harmless for his own account. In this paper, I want to draw inspiration from Sundström. I will formulate new counterexamples, inspired by Sundström's, which I think cannot be answered in the same way. I will then use these counterexamples to make a case for the idea that we cannot explain the intuition of distinctness as resulting from the "Antipathetic Fallacy" described by Papineau.

In a first section, I will explain how, in Papineau's theory, the antipathetic fallacy is supposed to account for the existence of the intuition of distinctness in a physicalist framework. In a second section, I will expose Sundström's criticism, as well as Papineau's answer to this criticism. In a third section, I will present two thought experiments, one of which is a clear counterexample to Papineau's account but cannot be answered in the same way as Sundström's objection. In a fourth section, I will present more thought experiments—with the aim of showing that the intuition of distinctness really has little to do with a hypothetical "use/mention" feature of phenomenal concepts. In a fifth section, I will consider one possible response to my argument, and I will try to counter it. The sixth section will be devoted to concluding remarks

1. *The antipathetic fallacy and the intuition of distinctness*

Phenomenal states are states such that *there is something it is like* to be in these states. A headache, a visual sensation of red, an olfactory sensation of honeysuckle, are typical examples of phenomenal states. These states are said to be endowed with *phenomenal properties*, which are properties in virtue of which these states are such that there is something it is like to be in them, and which are properties that determine *what it is like* to be in these states. *Being a visual sensation of red*, for example, that we can also label "phenomenal redness", is a typical example of a phenomenal property.

We have numerous reasons to think that these properties must be wholly identical with physical properties—broadly construed, as to include physically realized functional properties.[1] However, we are often deeply puzzled when we focus on our phenomenal states, and when we then try to think that they are fully identical with physical states, merely endowed with physical properties. How can *this* (thought by focusing, say, on a current visual sensation of red) be *the same thing* as a certain electrochemical activity in my visual cortex? Many of us, even convinced physicalists, admit that it seems to be a mystery. It has been said that in this kind of situation we face an *explanatory gap* (Levine 1983, 2001). David Papineau described this situation by saying that, in these cases, we encounter a strong *intuition of distinctness* (Papineau 2002): the intuition that our phenomenal states *are not* identical with physical states,[2] but truly are distinct from them. This explanatory gap

[1] These reasons have generally mostly to do with causal considerations (Levine 2001: Chapter 1; Papineau 2002, Chapter 1). I won't expound them here, as my goal is not to argue in favor of physicalism.

[2] I take these two descriptions of the issue to be roughly equivalent. This is confirmed by Levine's own words: "Whether we think of [the explanatory gap] as an explanatory gap or a distinctness gap, the problem is really the same" (Levine 2007: 148). Also see (Papineau 2011) for the idea that the explanatory gap has to be interpreted as constituted by the intuition of distinctness.

(or intuition of distinctness) fuels, in one way or in another, many anti-physicalists arguments regarding phenomenal consciousness (Chalmers 1996; Jackson 1982; Kripke 1980).

Some physicalists have suggested that it is possible to defend physicalism against this intuition, and against the arguments that it supports, by providing an *explanation* of this intuition within a physicalist framework. This explanation is supposed to rely on certain special features of some of the *concepts* we use to think about our phenomenal experiences. These concepts are called "phenomenal concepts", and they are the concepts we notably (but not only) use when we focus introspectively on our phenomenal experiences. In this view, phenomenal states indeed *seem* distinct from physical states. However, this happens merely in virtue of some features of the way in which we *think* about phenomenal states—features which are themselves purely physical. And, from a metaphysical point of view, phenomenal states really are identical with physical states. This kind of defense of physicalism has been labelled the "Phenomenal Concept Strategy" (Stoljar 2005). Numerous versions of this Strategy have been developed in the recent years (Aydede and Güzeldere 2005; Balog 2012; Hill 1997; Levin 2007; Loar 1997; Papineau 2002; Sturgeon 1994; Tye 1999).

Many of the theories belonging to the Phenomenal Concept Strategy have a common way to tackle the intuition of distinctness. They interpret it as a result of what David Papineau called the "Antipathetic Fallacy" (Papineau 1993, 2002, 2011). According to Papineau indeed, the intuition of distinctness arises because of a special feature of phenomenal concepts. Phenomenal concepts present a "use/mention feature": each occurrence of a given phenomenal concept involves the *instantiation* of the phenomenal property (identical to a physical property) this concept refers to, or at least of a property resembling it. This means that every time I think about a type of phenomenal experience using phenomenal concepts, I crucially activate a version of this experience, or at least what Papineau calls a "faint copy" of this experience. Phenomenal concepts, in this view, are peculiar because they *make use* of the property they *mention*.[3]

Why does this feature give rise to an intuition of distinctness? The explanation, according to Papineau, goes as follows. When we try to consider that a given phenomenal state (say, a visual experience of red) and a given physical state are *identical*, we make use of two different concepts. The first of them, being a phenomenal concept, brings the instantiation of phenomenal redness whenever we use it, while the other does not. Therefore, the phenomenal way of thinking about this property has itself a distinctive "feeling": *it is like* something to think about phenomenal redness with a phenomenal concept. On the other hand, there is no distinctive feeling when I think about one of my brain

[3] The details of this theory have changed over the years in Papineau's work (Papineau 1993, 2002, 2007), but the general idea has remained the same.

states using physical concepts. So, according to Papineau, "there is an intuitive sense in which exercises of material concepts 'leave out' the experience at issue. They 'leave out' […] the technicolour phenomenology, in the sense that they don't activate or involve these experiences" (Papineau 2002: 170). And this is where we commit what Papineau calls the "Antipathetic Fallacy": we can't help thinking that the fact that our physical conception "leaves out" something when compared with our phenomenal conception shows that these two conceptions simply *are not* about the same thing. This is why it seems to us that phenomenal states and physical states are distinct; this is how we get the intuition of distinctness.

Although David Papineau has been an early and a forceful defender of this kind of explanation, he is not the only one who has proposed something in the vicinity. An explanation of this type can indeed be found in the work of numerous philosophers proponents of the Phenomenal Concept Strategy. Brian Loar, one of the other main defenders of this Strategy, writes:

> A phenomenal concept exercised in the absence of the phenomenal quality it stands for often involves, not merely a recognitional disposition, but also an image. And so, as a psychological state in its own right, a phenomenal concept—given its intimate connection with imaging—bears a phenomenological affinity to a phenomenal state that neither state bears to the entertaining of a physical-theoretical concept. When we then bring phenomenal and physical-theoretical concepts together in our philosophical ruminations, those cognitive states are phenomenologically so different that the illusion may be created that their references must be different. (Loar 1997: 605)

Even if this feature is not the only feature that is supposed to account for the explanatory gap in Loar's account, it still plays an important role. Besides, the Antipathetic Fallacy, though not by this name, also plays a role in Michael Tye's and Katalin Balog's theories of phenomenal concepts (Balog 2012: 30–31; Tye 1999: 712–713). For reasons of simplicity, I suggest to call the hypothesis according to which the Antipathetic Fallacy (or something roughly equivalent) explains the birth of the intuition of distinctness the "Antipathetic Fallacy Hypothesis". I think that it is safe to say that this hypothesis constitutes one of the major lines of thought developed by proponents of the Phenomenal Concept Strategy in order to account for the explanatory gap in a physicalist framework.[4]

[4] The other aspect that proponents of the Phenomenal Concept Strategy usually insist upon is the conceptual independence of phenomenal concepts and physical concepts, which cause an absence of conceptual derivation from physical truths to phenomenal truths. Papineau does not insist upon this trait in his theory, however, as he does not think that the explanatory gap is primarily a matter of lack of conceptual derivation (Papineau 2011). I tend to agree with him as well as with Joseph Levine (Levine 2001, 2007: 200) on this point, even though I won't talk about it here.

2. *Sundström's counterexample and Papineau's response*

The Antipathetic Fallacy Hypothesis has been subjected to many criticisms. One of them, that I find quite compelling because it does not bear on many theoretical assumptions, relies on counterexamples. It has been developed by Pär Sundström (Sundström 2008).

The general idea of Sundström's criticism, as I understand it, can be exposed as follows. Let's accept that, as the Antipathetic Fallacy Hypothesis says, the use/mention feature of phenomenal concepts causes the intuition of distinctness.[5] If this is the case, then we should expect that, whenever we consider an identity statement of a certain kind (which I will describe in detail), an intuition of distinctness arises.

The relevant identity statements are statements which relate two conceptions of the same phenomenal property (identical with a physical property, given that the hypothesis is physicalist), with only one conception being systematically accompanied by the instantiation of this very phenomenal property. Let's call statements of this kind "phenomenologically contrasted identity statements". So, if the Antipathetic Fallacy Hypothesis is true, whenever we consider phenomenologically contrasted identity statements, we should have an intuition of distinctness concerning the two things identified in the statement.

Sundström then shows that there are cases that can intuitively count as counterexamples to this prediction: cases in which we *do* consider phenomenologically contrasted identity statements and yet do *not* have an intuition of distinctness. Sundström puts forth two examples of this kind. The first one essentially relies on some particular details of Papineau's account of phenomenal concepts, and notably Papineau's hypothesis that there are "derived" phenomenal concepts.[6] The second counterexample seems to me to be more compelling, as it does not bear on any specifics of the targeted theory, and therefore could apply to *any* theory that tries to explain the intuition of distinctness in a similar

[5] Papineau seemed at first to imply that the Antipathetic Fallacy was the only cause of the intuition of distinctness, but he later explicitly stated that it was likely to be just *one* cause of this intuition amongst others (Papineau 2011: 17–19).

[6] Roughly: Papineau says that there are, aside from "full-blown" phenomenal concepts, *derived phenomenal concepts* (Papineau 2007: 127–128). They are mental representations that are informationally deeply connected to "genuine" phenomenal concepts, so that they can refer to the same property, but whose instantiations do not necessitate the instantiation of the phenomenal property referred to (these concepts are required in order for me to be able to think thoughts such that: "I am not having an experience of this kind right now"). Sundström then builds a counterexample to Papineau's theory, crucially using these concepts. In a nutshell, his counterexample goes like this (Sundström 2008: 141): he notes that any identity statements relating a derived phenomenal concept and a genuine phenomenal concept will constitute what I called a phenomenologically contrasted identity statement, and then should cause an intuition of distinctness. However, according to him, this is obviously not the case.

manner as Papineau's—even one which is not committed to the existence of "derived" phenomenal concepts. For this reason, I will focus on this particular counterexample.[7]

The counterexample goes like this (Sundström 2008: 141–142). Consider an identity statement such as "My brother's most salient current experience = an experience of white". Let's say that the second half of the identity statement is thought while focusing on my current experience of the whiteness of the background of my Word document. On the other hand, if we consider the first half of the identity statement, it seems that it can be thought without any instantiation of phenomenal whiteness. After all, I can think about my brother's most salient current experience without having in mind a particular experience—even without knowing what kind of experience it is. Therefore, this identity statement is a *phenomenologically contrasted identity statement*. If the Antipathetic Fallacy Hypothesis is true, an intuition of distinctness should arise. But, according to Sundström, it is obviously not the case. I have no trouble entertaining the hypothesis that my brother's most salient current experience is an experience of white. I am in no way puzzled by this statement—while I am puzzled when I think that an experience of white *is* a certain neural activation in my sensory cortex. So, this counterexample seems to show that the Antipathetic Fallacy Hypothesis makes false predictions, and should therefore be abandoned.

Papineau later responded to this counterexample (Papineau 2011: 16–17). Acknowledging that Sundström's point is "well-taken", he seemed to agree with most of the premises of Sundström's objection. He notably seemed to accept that the Antipathetic Fallacy Hypothesis predicts that, when we face a phenomenologically contrasted identity statement, an intuition of distinctness should arise. He also recognized that we do not face such an intuition when we consider the identity statement: "My brother's most salient current experience = an experience of white".

His defense strategy against Sundström's objection amounted to arguing that this identity statement is *not* a phenomenologically contrasted identity statement after all. Maybe, he says, we "tend surreptitiously to activate the experience" of white when we think the first half of the identity statement. Or maybe we *don't* activate the experience of white when we think the second half of the identity statement—for example, because we are making use of a derived phenomenal concept instead of a "genuine" phenomenal concept in order to think about the experience of white. We just have to stipulate that one of these two possibilities is the case in order for the Antipathetic Fallacy Hypothesis to be protected against Sundström's objection.

[7] I also find it more interesting to focus on this counterexample as it is the only one (to the best of my knowledge) which has received an explicit response from Papineau.

Is this defense successful? I think that it can partially succeed, as one of the two possibilities described by Papineau could indeed be the case. It *could be* that, when I think about "My brother's most salient current experience" (and then try to equate it to an experience of white), I "tend surreptitiously to activate" an experience of white. Nothing, in Sundström's description of this situation, can guarantee that this is not the case. As for the other possibility, I don't think (*pace* Papineau) that it constitutes a way out for the defender of the Antipathetic Fallacy Hypothesis. Indeed, Sundström explicitly supposed that, when I thought the second half of the identity statement, I focused on my *current* experience of the whiteness of the background of my Word document. It couldn't be the case then that my thinking is not accompanied by an experience of white.

However, one possibility is enough to protect the Antipathetic Fallacy Hypothesis against Sundström's counterexample. Therefore, I think that it allows Papineau to block Sundström's objection.

My opinion is that Sundström's point is mostly right, and that the Antipathetic Fallacy Hypothesis does not constitute a satisfying explanation of the intuition of distinctness. My goal is to draw on Sundström's proposal and to propose some refined counterexamples, which do not allow for the same kind of defense move as the one suggested by Papineau. I will devote the rest of this paper to the description of these refined counterexamples.

3. *A refined counterexample to the Antipathetic Fallacy Hypothesis*

Suppose that I am sitting on a couch with my sister Elise, facing a large TV screen. Both of us have our eyes open, and we are visually paying attention to the screen. A computer feeds the screen with images—let's say, for reasons of simplicity, that they are only images of colored geometrical shapes: a red triangle, a blue square, a green rectangle, etc. In my hand, I hold a remote control. Every time I press a button on the remote control, the image on the screen changes. The software that runs in the computer makes it so that the succession of images is "random", in the sense that I have absolutely no way to predict what the next image will be.

I will now expose a few thought experiments which all use this device. Let's start with a rather innocuous one. Suppose that I am trying to consider a kind of naïve version of representationalist physicalism concerning consciousness, which states that when I have a conscious experience of an object, this conscious experience is identical with the physical state of my brain when it detects this object. Say that I am trying to decide if such a position is plausible while I am facing the screen—which, at the time, displays an image of a blue square. So: while visually focusing on the square, I consider the identity "My experience of this

blue square = The state of my brain when it detects this blue square". Let's stipulate that I think the first half of the identity statement by introspectively focusing on my experience of the blue square, and that, when I think the second half, I think the "blue square" component on the basis of my visual perception of it, by focusing on it. That means that the identity statement I am considering is *not* a phenomenologically contrasted identity statement. Indeed, my thinking of *both* halves of the identity statement crucially relies on me instantiating the phenomenal property associated with an experience of a blue square. But I take it that, in this case, I will still have a clear intuition of distinctness: I will be puzzled, as in any other case, by the fact that my *experience* could be identical with *a certain state of my brain*. I think that it shows that phenomenologically contrasted identity statements are not *necessary* for intuitions of distinctness to arise[8], which in turn shows that the Antipathetic Fallacy, if it can be *a* cause that gives rise to such intuitions, clearly cannot be their *only cause*. This point, as noted previously, has been clearly recognized by Papineau himself in recent papers.

Let's turn to a second thought experiment. Let's say that, still facing the screen, I close my eyes, and then press the button of the remote control. I now know that a new image is being displayed on the screen, and I know that my sister is visually experiencing it. However, I have no idea what the image is. Now let's say that, with my eyes still closed, I start thinking about "My sister Elise's most salient current experience". I assume that this experience is the experience that she is currently having as she watches the screen. However, given that I have my eyes closed, I have no idea what this experience is. This guarantees me that, when I think about Elise's most salient current experience with my eyes closed, I am *not* (surreptitiously or not) activating an experience of the same kind.[9] Let's say that I now prepare myself to entertain the hypothesis that Elise's most salient current experience, about which I am thinking with my eyes closed, is type-identical with the experience that I would myself get if I opened my eyes. I then consider that Elise's most salient current experience is identical with… (and here I open my eyes) *an experience of a purple hexagon* (where this part is thought with a phenomenal concept, and while focusing on the very experience I got as soon as I opened my eyes).

In this situation, do I have an intuition of distinctness? I take it to be obvious that I don't. Of course Elise's experience *can be* an experience of a purple hexagon. Of course she can be in a state that *feels like*

[8] Sundström also put forth a thought experiment aiming at showing that, though it relied on the details of Papineau's account regarding *derived phenomenal concepts* (Sundström 2008: 141).

[9] Except if (1) I am imagining what this experience *could be* and, by chance, I just *got it right*—but this would happen only very rarely; or (2) We suppose that I am able to activate together and at the same time *hundreds* of different visual experiences (of a blue square, of a red diamond, of a yellow star, all in different sizes and hues, etc.). I take this to be completely implausible.

this. But here is the point: the identity statement I was considering was a phenomenologically contrasted identity statement, and it seems very difficult to deny it. Indeed, I forced myself to think about Elise's current experience *without any possibility of knowing what this experience was*, so that I was obviously *not* activating a copy of this experience (even "surreptitiously"). That guarantees that the first half of the identity statement was thought about *without* activating an experience of a purple hexagon. And then I forced myself to think about the experience of a purple hexagon by introspectively focusing on the very experience I got when I opened my eyes—which guarantees that, this time, I instantiated the corresponding phenomenal property when I thought about the experience.

So, in the situation described in the thought experiment, I consider a statement, which is quite certainly a phenomenologically contrasted identity statement, and I nonetheless get no intuition of distinctness. This case therefore constitutes a counterexample to the Antipathetic Fallacy Hypothesis, which draws inspiration from Sundström's case but cannot be answered in the same way.

4. *Pulling apart the intuition of distinctness and the phenomenologically contrasted identity statements*

I have shown previously, in my first thought experiment, that phenomenologically contrasted identity statements are not *necessary* for an intuition of distinctness to arise. This is something Papineau himself recognized. I then presented a thought experiment that gives a reason to think that they are not *sufficient* for an intuition of distinctness. I tend to think that this shows that we should pull apart the issue of intuitions of distinctness, and the issue of phenomenologically contrasted statements. I now want to quickly present a few more thought experiments that could bring our intuition in the same direction.

Suppose that I am still facing the same screen, with my sister Elise still by my side. I then close my eyes and press the button of the remote control. At this point I know that my sister is looking at an image, but I don't know which image it is. Let's say that, my eyes closed, I start thinking about "the current state of Elise's visual cortex". I then decide to consider the fact that it is identical with… (and then I open my eyes) the state of Elise's visual cortex when she looks at *this*—where *this* is, say, a red oval, about which I think *on the basis of my visual perception of it*. Do I then have an intuition of distinctness? I take it to be obvious that I don't: the identity I am considering seems perfectly reasonable. *Of course* the current state of her cortex can be identical with a certain state of her cortex! But the identity statement I was trying to consider at the time was nonetheless a phenomenologically contrasted identity statement. This *also* shows that the phenomenological contrast between two conceptions is not sufficient to create an intuition of distinctness concerning the two objects referred to by the conceptions.

Now, let's consider one more thought experiments in which an intuition of distinctness *does* arise while I consider a phenomenologically contrasted identity statement, but in such a way that it does not fit well with the Antipathetic Fallacy Hypothesis. Let's say that, while I am in the same kind of situation as described previously, I think, with my eyes closed, about Elise's most salient current experience. Again, I have no idea what it is at the time. I then consider that this experience is identical with… (and then I open my eyes) the state of Elise's visual cortex when she looks at *this*—where *this* is, say, a blue spiral, about which I think on the basis of my visual perception of it.

I take it that, in that case, an intuition of distinctness would arise: how could Elise's experience be identical with *a state of her cortex*? This identity statement would seem as strange as any other physico-phenomenal identity statement. I may *believe* it, but I will find it puzzling nonetheless. In that case, the identity statement I consider happens to be a phenomenologically contrasted identity statement. However, we can see here that the phenomenologically loaded conception, which is in the *second half* of the identity statement, is not at all the conception that seems to refer to an irreducibly phenomenal entity. In fact, that is exactly the contrary. The intuition of distinctness arises, but what strikes me as being irreducibly phenomenal is the thing thought about in a non-phenomenologically loaded way: eyes closed, and while *not knowing* what kind of experience is thought about.

Our previous examples had shown that phenomenologically contrasted identity statements were not necessary, nor sufficient, for intuitions of distinctness to arise. The further examples I just put forth are cases in which the two relevant factors (the intuitions of distinctness in the one hand, the phenomenologically contrasted identity statements on the other hand) can vary quite independently from each other. I hope they will incite the reader to completely pull apart these two things. An intuition of distinctness can arise, whether or not we are considering a phenomenologically contrasted identity statement, and I don't think that we have solid reasons to believe that one causes the other.

I don't intend to assert here that the specific way in which we grasp our phenomenal experiences through introspection is *not* crucial when it comes to explaining the arising of the intuition of distinctness. I actually think that our introspective grasp of consciousness has special features, which explain this intuition. But the Antipathetic Fallacy hypothesis, understood as one particular way to interpret in what way our introspective grasp of experiences contributes to the presence of this intuition, is mistaken. It is not true that we are reluctant to equate a phenomenal experience (thought about introspectively) with a purely physical state (thought about with purely physical concepts) because the first thought activates the concerned experience (or a copy of this experience), while the other doesn't.

5. *Objection: can we really separate two steps in our thinking of identity statements?*

The counterexamples I just presented may be subject to objections. I would like to consider one of them, and then try to answer it.

In order for my counterexamples to be immune to the kind of answer that Papineau gave to Sundström's objection, I had to describe them in such a way that it is guaranteed that the identity statements considered are indeed phenomenologically contrasted identity statements. This crucial task was fulfilled thanks to some specific features of the situation described in the counterexamples.

It is especially crucial to the counterexamples that the first half of the identity statements considered is always thought *with the eyes closed* (while ignoring the image displayed on the screen), and that the other half is thought *with the eyes open*, and on the basis of the visual perception I then get. But one could object the following: when we consider an identity statement, our thought cannot be *temporally divided* in such a clear and cut way. Thought is not like speech, in this respect. In fact, whenever we consider an identity statement, we have to think the two conceptions of the two things that are being identified *at the same time*, and thus *bring together* these two conceptions in our mind, so to speak. For this reason, when we look at the counterexamples I just described, there is no sense in saying that we only think the first half of the identity "the eyes closed", and without activating the relevant experience, because we also have to think it with the eyes open. Indeed, it is only when we have the eyes open that we can properly think the second half of the identity statement, and then "assemble" in our minds the first half and the second half of the identity statement. Therefore, when we think the relevant identity statement, we *must* think *even* the first half of the identity statement with our eyes open, visually attending at the screen, which means that we must think it while having the relevant experience. This means that the identity statements we are considering in the counterexamples are not phenomenologically contrasted identity statements after all. Therefore, they are not counterexamples to the Antipathetic Fallacy Hypothesis.

This objection has a certain appeal. However, I don't think that a defender of the Antipathetic Fallacy Hypothesis could make use of it in order to defend her theory against the counterexamples I put forth. Indeed, this objection crucially relies on the thesis that, when we think an identity statement, we have to activate *at the same time* the two conceptions (of the two things we try to identify) in order to "bring them together" in our mind. But this would destroy the very possibility of phenomenologically contrasted identity statements. For this would mean that, anytime I think a given identity statement, I have to entertain the *two conceptions* thought about at the same time, which means that, at least at that moment, both conceptions would be accompanied

by whatever phenomenology accompanies the other. That means that I could *never* think an identity statement in such a way that my thinking of one part of the statement would be accompanied by a given phenomenology, while my thinking of the other part would *not*. But, if there are no such things as phenomenologically contrasted identity statements, then the *explanans* posited by the Antipathetic Fallacy Hypothesis does not exist. Therefore, this hypothesis is false.

For this reason, even if this objection can seem appealing, I don't think a defender of the Antipathetic Fallacy Hypothesis can make use of it in order to repeal my counterexamples.

6. *Concluding remarks*

In this paper, I have devised thought experiments in order to show (conclusively, I hope) that phenomenologically contrasted identity statements are neither necessary, nor sufficient, for intuitions of distinctness to arise. The counterexamples I designed to show that they are not sufficient were inspired by Pär Sundström's counterexample to Papineau's theory. I tried to construct them in such a way that it was not possible to answer them in the same way Papineau answered Sundström's. I also tried to put forth a variety of cases, in which these two features of the situations (whether or not the identity statements considered are phenomenologically contrasted, and whether or not they create an intuition of distinctness) vary independently. My goal was to incite the reader to pull apart these two features of the situation: the phenomenological contrast that can exist between two halves of an identity statement, and the birth of an intuition of distinctness vis-à-vis the two objects identified in the statement.

In this paper, I argued against the Antipathetic Fallacy Hypothesis, which is a hypothesis that aims at defending physicalism against the dualist intuition (the "intuition of distinctness"), by giving an explanation of this intuition within a physicalist framework. However, I did not plan to argue against physicalism. I did not even plan to argue against the Phenomenal Concept Strategy, if we understand it as the general attempt to defend physicalism against dualist intuitions (such as the intuition of distinctness) by appealing to some physically explainable features of our way of thinking about conscious experience in order to explain the birth of this intuition. I therefore think that numerous versions of the Phenomenal Concept Strategy (Aydede and Güzeldere 2005; Hill 1997; Levin 2007; Sturgeon 1994) are left untouched by my argument. My own point of view is that physicalism is true, and that we can account for the intuition of distinctness within a purely physicalist framework, by showing how this intuition arises as a consequence of some of the features of our introspective grasp of consciousness. However, as I tried to show, the Antipathetic Fallacy Hypothesis does not give us a satisfying explanation of this intuition. My own view,

for which I did not argue here, is that the only satisfying physicalist theory of our introspective grasp of consciousness is an *illusionist* one, according to which we introspectively represent ourselves as conscious even though consciousness does not really exist (Frankish 2016). Illusionist views of consciousness escape the objection made here, as well as objections usually made against the Phenomenal Concept Strategy. Of course, they encounter problems of their own, whose solution may not be trivial (Kammerer 2018, 2016).

References

Aydede, M. and Güzeldere, G. 2005. "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness." *Noûs* 39 (2): 197–255.

Balog, K. 2012. "Acquaintance and the Mind-Body problem." In C. Hill & S. Gozzano (ed.). *New Perspectives on Type Identity: The Mental and the Physical*. Cambridge: Cambridge University Press.

Chalmers, D. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Frankish, K. 2016. "Illusionism as a Theory of Consciousness." *Journal of Consciousness Studies* 23 (11–12): 11–39.

Hill, C. 1997. "Imaginability, Conceivability, Possibility and the Mind-Body Problem." *Philosophical Studies* 87: 61–85.

Jackson, F. 1982. "Epiphenomenal qualia." *Philosophical Quarterly* 32: 127–136.

Kammerer, F. 2016. "The hardest aspect of the illusion problem—and how to solve it." *Journal of Consciousness Studies* 23 (11–12): 123–139.

Kammerer, F. 2018. "Can you believe it? Illusionism and the illusion metaproblem." *Philosophical Psychology* 31 (1): 44–67.

Kripke, S. 1980. *Naming and Necessity*. Cambridge: Harvard University Press.

Levine, J. 1983. "Materialism and qualia: the explanatory gap." *Pacific Philosophical Quarterly* 64: 354–61.

Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.

Levin, J. 2007. "What is a Phenomenal Concept?" In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

Levine, J. 2007. "Phenomenal Concepts and the Materialist Constraint." In T. Alter and S. Walter (ed.). *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

Loar, B. 1997. "Phenomenal States (Revised Version)." In N. Block, O. Flanagan and G. Güzeldere (ed.). *The Nature of Consciousness*. Cambridge: MIT Press: 597–616.

Papineau, D. 1993. "Physicalism, Consciousness, and the Antipathetic Fallacy." *Australasian Journal of Philosophy* 71: 169–183.

Papineau, D. 2002. *Thinking about Consciousness*. Oxford: Oxford University Press.

Papineau, D. 2007. "Phenomenal and Perceptual Concepts." In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.

Papineau, D. 2011. "What Exactly is the Explanatory Gap?" *Philosophia* 39 (1): 5–19.

Stoljar, D. 2005. "Physicalism and phenomenal concepts." *Mind and Language* 20 (2): 296–302.

Sturgeon, S. 1994. "The Epistemic View of Subjectivity." *The Journal of Philosophy* 91 (5): 221–235.

Sundström, P. 2008. "Is the mystery an illusion? Papineau on the problem of consciousness." *Synthese* 163 (2): 133–143.

Tye, M. 1999. "Phenomenal consciousness: the explanatory gap as a cognitive illusion." *Mind* 108 (432): 705–725.