

AMERICAN UNIVERSITY OF BEIRUT

PLACE, SMART, AND AFTERIMAGES

by
HOUSSAM NADIM GHANEM

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Arts
to the Department of Philosophy
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
August 2023

AMERICAN UNIVERSITY OF BEIRUT

PLACE, SMART, AND AFTERIMAGES

by
HOUSSAM NADIM GHANEM

Approved by:



Signature

Dr. Hans D. Muller, Associate Professor
Department of Philosophy

Advisor



Signature

Dr. Bana M. Bashour, Associate Professor
Department of Philosophy

Member of Committee



Signature

Dr. Christopher Johns, Associate Professor
Department of Philosophy

Member of Committee

Date of thesis defense: August 28, 2023

AMERICAN UNIVERSITY OF BEIRUT

THESIS RELEASE FORM

Student Name: Ghanem Houssam Nadim
 Last First Middle

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my thesis; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

- As of the date of submission
- One year from the date of submission of my thesis.
- Two years from the date of submission of my thesis.
- Three years from the date of submission of my thesis.



06/09/2023

Signature

Date

ABSTRACT OF THE THESIS OF

Houssam Nadim Ghanem

for

Master of Arts

Major: Philosophy

Title: Place, Smart, and Afterimages

This paper argues that U.T. Place and J.J.C Smart's shared version of identity theory fails on two accounts. The first is their account of what an introspecting subject means when they report a nonveridical experience. This paper argues that the subject cannot mean what the duo claim they do – namely, that they are having an experience similar to a veridical experience – given the physicalist premises they employ. The second is their account of what a scientist can explain about the subject's nonveridical experience through observation of their brain processes. This paper contends that a scientist cannot adequately explain all types of nonveridical experiences. The paper then investigates whether or not Place and Smart's project would benefit from redefining reports of experiences as expressives. Lastly, it posits that Place and Smart are stuck in a sort of paradox.

PREFACE

Introduction

In this thesis, I will argue that in their effort to make a convincing case for identity theory, U.T. Place and J.J.C Smart fail on two accounts, which is to say on two separate fronts. I shall demonstrate this in two chapters. The first is that of their account of what an introspecting subject means when they report a nonveridical experience. Place and Smart assert that all the subject means when they report a nonveridical experience is that they are having an experience similar to a veridical one. I do not believe the metaphysics Place and Smart employ allow the subject's report to mean that. At the very least, if it does, it does not do so convincingly. The second is that of their account of what a scientist can explain about the subject's nonveridical experience through observation of their brain processes. Place claims the scientist ought to be able to explain the subject's experience by simple observation of the appropriate brain process. I believe this only works for a minority of experiences. In the first half of my third chapter, I will consider a possible move Place and Smart could make that might help avoid some of the problems identified in the first chapter. The chapter will then conclude with a presentation and explanation of the paradox I see facing Place and Smart. In presenting it, my goal will not be to imply that their identity theory does not or cannot work, but only to elucidate a fundamental difficulty they will need to deal with if they are to make it work.

Exposition

In his paper "Is Consciousness a Brain Process?", Place investigates whether or not the thesis that mental processes are in fact brain processes can be considered a

reasonable scientific hypothesis. He considers the notion that conscious experience is seemingly saturated with phenomenal properties, that is, qualities of experience such as colors, sounds, smells, and so forth. It is often argued that because such properties are clearly absent in the physical brain and its processes, they must be properties of mental processes existing over and above brain processes. But this would mean that at best mental processes merely correlate with brain processes.

Place claims that this line of reasoning commits what he terms the *phenomenological fallacy*, “the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste or feel to him, he is describing the literal properties of objects and events on a peculiar sort of internal cinema or television screen” The fallacy is based on the mistaken assumption that because our describing of the environment requires that we be conscious of it, our descriptions are primarily descriptions of our experience and only secondarily those of the environment itself. In reality, the opposite is the case: conscious experience is described not in terms of any phenomenal properties but by reference to the actual physical properties which give rise to it. Just because we may know green apples by our experience of them, it does not follow that our experience is itself green.

The phenomenological fallacy, Place explains, is what unsurprisingly leads to the assumption of an introspecting subject reporting a green afterimage, for example, that there literally exists some mental object within them which is green. But no such thing actually exists. The report of the afterimage is merely the report of an experience of the sort the subject normally has when they are observing a green light. Smart, in his paper “Sensations and Brain Processes,” puts this idea more clearly: “When a person says, “I see a yellowish-orange afterimage,” he is saying something like this: “*There is*

something going on which is like what is going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange.”” The italicized words are topic-neutral between physicalism and dualism, Smart explains, because they are abstract enough to be compatible with either metaphysics: a physicalist could use them to refer to brain processes where a dualist could use them to refer to phenomenal properties. Their topic-neutrality explains why physicalists and dualists can talk “happily together” about their aches, pains, and so on.

So long as we do not commit the phenomenological fallacy, it does not seem inconceivable that mental processes should be strictly identical with brain processes and that Smart’s topic-neutral description of sensation reports should in fact be referring to brain processes. To say that a brain process cannot be green is not to say that it cannot be the experience of a green afterimage, where the experience is described in the topic-neutral way by reference to the external environment. All a scientist needs to do in order to explain the subject’s experience of the green afterimage, Place concludes, is demonstrate that the brain process occurring within them at that moment is the same sort of brain process which occurs when they observe a green light (Place 1956; Smart 1959; Smart 1972).

TABLE OF CONTENTS

ABSTRACT	1
PREFACE	2
SMART’S TOPIC-NEUTRAL DESCRIPTION OF SENSATION REPORTS	6
1.1. The First Half.....	7
1.2. The Second Half	16
PLACE’S SCIENTIST	21
2.1. Observing Brain Processes	22
REFLECTIONS	28
3.1. Redefining Reports as Expressives.....	28
3.2. The Paradox	36
CONCLUSION	43
REFERENCES	44

CHAPTER 1

SMART'S TOPIC-NEUTRAL DESCRIPTION OF SENSATION REPORTS

The following chapter investigates Smart's topic-neutral description of sensation reports. When a subject reports their experience of a green afterimage, Place and Smart tell us that what the subject is effectively uttering is "*there is something going on which is like what is going on when I am looking at a green light,*" with the italicized words being topic-neutral.

Now, I am not convinced a subject is able to meaningfully utter such a report without access to phenomenal properties. Without access to green phenomenal properties, on what basis might they compare their experience of a green afterimage to that of a green light? Surely, the reason they compare their experience of the afterimage to that of a green light – and not a blue light, for example – is because of the greenness factoring in both experiences. Well, Place and Smart deny this. I am here committing the phenomenological fallacy, they would say. The reason the subject is able to meaningfully utter the report has to do entirely with brain processes, not anything green floating about in their skull.

To investigate this, then, I would like to divide the topic-neutral description in two. The first half, "*there is something going on,*" concerns the subject having an experience, that is, the subject identifying their experience, in this case their experience of a green afterimage. In section 1.1, I will show that there is little sense to the idea the subject is having an experience at all, let alone one which they can subsequently report. The second half, "*which is like what is going on when,*" concerns the subject describing that experience as like another, in this case as like an experience of observing a green

light. In section 1.2, I will show that the subject lacks all ability to meaningfully do so. Without access to phenomenal properties, I aim to argue, Place and Smart's physical tools do not allow the subject sufficient means to meaningfully utter their report, at least not convincingly.

1.1. The First Half

When a subject reports their experience of an afterimage, Place and Smart tell us that the experience, identical to a brain process, certainly exists but that the afterimage itself does not. In his short paper *A Note on the Fallacy of the 'Phenomenological Fallacy,'* J.R. Smythies raises the question of what exactly is going on when an introspecting subject reports their experience of an afterimage. If the afterimage does not exist, asks Smythies, is the introspecting subject really having an experience? "Surely, to experience nothing is not to have an experience?" (Smythies 1957). To put it another way, if the afterimage does not exist, in what way does it make sense to claim that the introspecting subject is nonetheless having an experience of one such that they should be able to meaningfully utter the phrase "*there is something going on*"? On what basis could they do so? Before investigating this it is important we establish what Place and Smart take a conscious experience to comprise.

At the very least, both Place and Smart must understand conscious experience as constituting the noticing of objects by virtue of their features. As much is clear given Place's definition of the phenomenological fallacy, something which Smart adopts. Recall that Place defines the fallacy as "the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste or feel to him, he is describing the literal properties of objects and events on a peculiar sort of

internal cinema or television screen” Place defines the subject describing their experience as them describing how things look, sound, smell, taste, or feel to them. It follows from that that Place must at minimum understand experience as a noticing of looks, sounds, smells, tastes, or feelings, that is, as a noticing of features. This definition of conscious experience is all we need. There is no need to investigate the matter further. What is important in understanding how an experience of a nonexistent afterimage might be possible is only what Place and Smart must at the very least understand an experience to comprise.

To Place and Smart, therefore, to have a conscious experience of an object is at the very least to notice that object by virtue of its features. Now, to reiterate, the two tell us that when a subject reports their experience of an afterimage, only their experience but not the afterimage itself exists. But if the afterimage does not exist it cannot have any features. Features can only be predicates of something which exists; that is the very definition of a feature, that it is the way *something* is. But if the afterimage has no features, but an experience is defined as the noticing of something by virtue of its features, how is an experience of it possible? What is being noticed by virtue of its features? The definition of experience Place and Smart employ necessitates that an object have features the noticing of which comprises the experience, but the object cannot have said features unless it actually exists.

But the idea here, that afterimages cannot have features unless they exist, is obviously something neither Place nor Smart accept. They routinely refer to afterimages by way of their features, such as by being green, yellowish-orange, or blurry. When a subject reports their experience of an afterimage, they tell us their experience may be of a green or yellowish-orange one. Now, if afterimages can indeed be green or yellowish-

orange, then there certainly must be some sense in which they do exist after all. That they can be green or yellowish-orange necessarily presupposes so. *Something* must be green or yellowish-orange.¹ And Smart seems to imply as much when he refutes the existence of afterimages only *in a sense* (Smart 1959). So there does seem to be a sense in which things like afterimages and their features exist to Place and Smart. The question now is what can be said of this sense of existence? How can an afterimage be green or yellowish-orange in such a sense?

I would like to anticipate a seemingly powerful yet ultimately misplaced argument one could make on behalf of Place and Smart, one which *prima facie* might appear to stop our line of reasoning in its tracks. Suppose it is posited that an afterimage is in fact identical to the experience of it, that the two are in reality one and the same. Would this adjustment not rescue Place and Smart from the problems we have thus far been highlighting? If the afterimage were identical with the experience of it and the experience of it identical with a brain process, then since brain processes certainly exist (and have plenty of properties), it would follow that afterimages likewise exist (and have plenty of properties). The problem of how an experience of something which does not exist is possible would be done away with.

The issue here is that this supposed solution raises an obstacle for identity theorists the very distinction between the afterimage and the experience of it is meant to overcome. If it is claimed that a green afterimage, for instance, is identical to the experience of it, it would follow that the experience of it is also green. The problem for

¹ To be clear, the argument being alluded to here, that a feature is the way *something* is and that therefore green or yellowish-orange afterimages must necessarily exist simply on account of having features, though compelling, is unnecessary. The ontology of things like afterimages is less important to the project at hand than what Place and Smart assert regarding the nature of conscious experience, and how that nature relates to their claims regarding afterimages. It is less important how afterimages might exist and more important how Place and Smart's notion of conscious experience makes sense if they do not.

identity theorists like Place and Smart is that it is certainly not the case that the brain process then identical to that green experience is also green. There is no greenness to be found in the brain. This is the exact reason Place does away with phenomenal properties at least in a specific sense at the very outset of his and Smart's project by way of the phenomenological fallacy. If phenomenal properties exist, they do not exist as the means for an introspecting subject's effectively uttering the topic-neutral description "*there is something going on which is like what is going on when,*" but only topic-neutrally (Smart 1959). What serve as the means for the description's utterance, what is in fact "*going on,*" are like brain processes. But an experience which itself was green on account of being identical with a green afterimage would be means enough for the description's utterance. It would be comparable to the experience of a green light purely on its own merits as a green experience. But this would entail a form of dualism. Physicalist identity theorists like Place and Smart clearly cannot abide by this. This is precisely why they strip conscious experience of its phenomenal oomph, by way of the phenomenological fallacy in the case of ordinary veridical experience, and by way of distinguishing the experience of the object from the object itself in the case of nonveridical experience, relegating phenomenal properties to topic-neutral obscurity. Once we are no longer looking for green brain processes, explains Place, identity theory seems far less insuperable (Place 1956). Identifying afterimages with experiences of them is therefore not an option for Place and Smart. On the contrary, the success of their project hinges on their separating the two.

Returning to the task at hand then, neither Place nor Smart gives us satisfactory answers to the question we left off with: in what sense could it be said that an afterimage exists such that it is able to have the features of being green or yellowish-

orange? Smart attempts to do so by likening the existence of things like afterimages to the existence of the average plumber. Just as the average plumber exists merely as a derivative of plumbers, things like afterimages exist merely as derivatives of the experiences of them (Smart 1963). But this analogy does not address our concern whatsoever. The average plumber exists as a derivative of plumbers in the sense that sentences about the average plumber can be translated into sentences about plumbers. But the issue we are facing has nothing to do neither with sentences about afterimages nor sentences about experiences of them, let alone with their translatability. The issue we are facing has to do with how conscious experience itself, regardless of what can or cannot be said of it, is possible given the nonexistence of whatever it is experience of. The ostensible translatability of sentences about the afterimage into sentences about the experience of it does not provide us with a sense of the afterimage's existence such that the concept of an experience of one then makes sense. Unless Smart believes that the average plumber existing as a derivative of plumbers means that the average plumber is an actual thing which can then be experienced, his analogy does not even apply to the matter at hand.

All we are left with, therefore, is the implication that the only sense in which an afterimage can be said to exist is in the sense that an experience of one can. If the experience is an experience of an afterimage, the afterimage must exist one way or another. To call this argument a poor one is an understatement. A few glaring issues. First, this is to flip the nature of conscious experience on its head, at least when it comes to things like afterimages. Most objects exist as things with certain features and then are had conscious experiences of by virtue of the noticing of said features. But when it comes to things like afterimages, first comes the noticing of said features, the conscious

experience, from which the existence of the afterimage being noticed is then derived. Surely, advocates of Occam's razor such as Smart himself would find this idea suspect at best (Smart 1959). Second, this quite ironically leads to a form of idealism where the conscious experience of a certain type of object constitutes in its entirety that object's existence. This is obviously something physicalists like Place and Smart cannot abide by, but something which it appears they cannot avoid if they cannot elaborate further on the matter. The opposite would be an equally unpalatable form of dualism: they could not concede things like afterimages to be parts of "the furniture of the world" (Smart 1963).

A quick recap. The aim of this section has been to investigate how the first half of Smart's topic-neutral description of sensation reports "*there is something going on*" could be meaningfully uttered given Place and Smart's premises. We began by questioning how a subject could have an experience of an afterimage if the afterimage does not exist. We worked through the issue, employing what Place and Smart must at the very least understand conscious experience to comprise. We ultimately arrived at the conclusion that if there was a sense in which the afterimage could be said to exist after all and therefore a way in which the subject could meaningfully utter the phrase about their experience, it was a very unconvincing one.

So, if we cannot make much sense of the experience of an afterimage in terms of what it is an experience of, in terms of the existence of the afterimage, perhaps we can do so in some other way. Perhaps we can set aside entirely the issue of the afterimage. Let us accept the premise that the experience itself nonetheless exists. Indeed, considering experiences are identical to brain processes and that brain processes certainly exist, we could not even deny it. Well, if the experience exists, then it has

phenomenal properties. This is something identity theorists like Place and Smart must – and do – readily grant. If the experience is identical to some brain process, then it must have properties – phenomenal properties – that are identical to the neurological properties of that corresponding brain process.²

Now, could such properties not be the grounds for the subject’s uttering “*there is something going on*”? Though they may be having an experience of an afterimage which might not exist, the experience itself has phenomenal properties which do. Could these properties not be the basis upon which the experience makes sense and the report is made? The answer is that they cannot. If the experience is identical to a brain process, then its phenomenal properties must be identical to neurological properties, that is, must be *one and the same* with the corresponding neurological properties. But the introspecting subject certainly has no access to the neurological properties of their brain processes. They may not even be aware that they have a brain.

Of course, this cannot be the end of the matter. Place and Smart could simply reassert that since the experience is identical to a brain process and that the subject obviously has access to their experience and its phenomenal properties, then it follows that they must by definition have access to the corresponding brain process and its neurological properties. The problem with this idea is that it is unclear what it exactly means. Suppose the subject in question is lying down on an operating table and that stood behind them is a scientist who has cracked open their skull and is observing their brain. Surely, whatever properties the subject has access to during their experience

² The question of whether or not Smart does in fact admit phenomenal properties – instead of merely topic-neutral ones – could be raised here. It could be argued that since he does not, it is quite difficult to make sense of the experience in this way. What is most important to the remainder of this section however is the idea that whatever the properties of the experience may be, they are identical to neurological properties. The issue of their topic-neutrality will be saved for chapter 3.

cannot be identical to the properties the scientist does. The subject certainly does not see, hear, smell, taste, or feel what the scientist does. But Place and Smart tell us that experiences and brain processes, that phenomenal properties and neurological properties, are one and the same. If that is indeed the case, they will need to explain how it is that the subject having access to their experience means that they have access to the corresponding brain process if they have no access to what the scientist stood behind them does.

But all Place and Smart need reply here is that I am operating on a misconception. They could argue that the subject need not have access to exactly what the scientist does in order for it to be true that they have access to their brain process. They could argue that the subject's phenomenal properties serve as a route of access to the very same neurological properties that the scientist has a different route of access to. The subject would have access to their neurological properties *as* phenomenal properties. The problem with this is that Place and Smart quite simply cannot allow it – the topic-neutral description cannot allow it.

Recall that when a subject reports their experience of a green afterimage, Place and Smart tell us that what the subject is effectively uttering is “*there is something going on which is like what is going on when I am looking at a green light,*” with the italicized words being topic-neutral. “*Something going on*” could either be referencing brain processes in the case of a physicalist metaphysics or phenomenal properties in the case of a dualist one. To the duo, what is in fact “*going on*” is of course a brain process. But herein lies the problem: *all* that is “*going on*” is a brain process. *All* that is “*going on*” such that the subject is able to utter the description regarding their experience of the afterimage is a brain process. They have no access to any green phenomenal property by

virtue of which they might be able to liken their experience to that of observing a green light. Such a property has no place in a physicalist metaphysics. What the utterance is based on is solely and entirely the fact that what is taking place, what is “*going on*,” is a brain process similar to that which takes place when the subject is observing a green light.

All this means is that there can be nothing that the introspecting subject has access to which the scientist stood behind them does not that serves as the reason for the subject’s uttering “*there is something going on*.” For Place and Smart to invoke the idea that the subject has access to their neurological properties by way of having access to their phenomenal properties is to imply that said phenomenal properties play any part when it comes to the reason the subject utters the topic-neutral description in the first place. But they clearly do not. There can be no such properties floating about to which only the subject has access. All that exists, all that the utterance is based on, is a brain process, one which the scientist but not the subject has access to. This is my response to Place and Smart, then. Either the subject has access to their neurological properties *as* phenomenal properties or else all that is referenced by the topic-neutral description, all that is “*going on*,” is a brain process. They cannot have both.

Still, Place and Smart might respond that how they understand phenomenal properties is not the same as how a dualist does. They might argue that the two possibilities I have laid out, of neurological properties as only the scientist has access to on the one hand and phenomenal properties as things floating about on the other, represent a false dichotomy. The problem with this is that it is far from clear what these phenomenal properties that are supposedly neither of those two things then are or how they exist. They are neither what the scientist sees looking into the subject’s skull nor

are they things floating about. They exist quite confusingly in a sort of metaphysical limbo.

All that said, it remains conceivable to me that the introspecting subject does have direct access to their neurological properties. It rings true intuitively, and I am certainly not yet convinced of its impossibility. The issue is Place and Smart have not convincingly made the case as to how that might work.

1.2. The Second Half

In his paper *Mental Events and the Brain*, Jerome Shaffer puts into words a certain issue he has with identity theory:

Let us take the case where a person reports the having of some mental event, the having of an afterimage, a thought, or a sensation of pain. Now such a person has surely noticed that *something* has occurred, and he has surely noticed that this something has *some* features (or how could he report it was an afterimage rather than a sensation of pain?). Now it seems to me obvious that, in many cases at least, the person does not notice any *physical* features – he does not notice that his brain is in some particular state, nor does he notice any external physical stimulus, nor any physical event between the stimulus and the neurological response. Yet he does notice *some* feature. Hence he must notice something other than a physical feature. The noticing of some nonphysical feature is the only way to explain how anything is noticed at all. (Shaffer 1963)

Now, when a subject reports their experience of a green afterimage, Place and Smart tell us that what the subject is effectively uttering is “*there is something going on which is like what is going on when I am looking at a green light,*” with the italicized words being topic-neutral. Well, would it not be on the basis of such nonphysical features that the introspecting subject is able to describe their experience of a green afterimage as like that of a green light? Surely, the greenness of the afterimage is what prompts the describing of the likeness of its experience to that of a green light. Does it

not take the existence of such features for the subject's utterance of the second half of the topic-neutral description "*which is like what is going on when*" to make sense? Not so, according to Smart.

Anticipating this objection, Smart argues that it is possible for a subject to notice a likeness between two of their experiences without noticing *how* the experiences are alike, provided we think of the issue cybernetically. "The strength of my reply," he explains, "depends on the possibility of our being able to report that one thing is like another without being able to state the respect in which it is like. I do not see why this should not be so. If we think cybernetically about the nervous system we can envisage it as able to respond to certain likenesses of its internal processes without being able to do more" (Smart 1959). To illustrate the point, he provides two analogies for the nervous system concerning machines.

His first analogy is of a machine built to analyze objects. It would be easier, he claims, to build a machine which could report whether two objects were similar than it would be to build one which could also report wherein those similarities consisted (Smart 1959). The idea being represented here is that it is certainly more likely that the nervous system could simply react to the occurrence of similar brain processes than it is that it could also react to *how* the brain processes were similar.

His second analogy is of a machine built to produce sequences of numbers. It would be easier, he claims, to build a machine which produced the sequences 1, 2, 3, 4, 5 ... and then 1001, 1002, 1003, 1004, 1005 ... than it would be to build one which produced the sequences 1, 2, 3, 4, 5 ... and then 1002, 1004, 1006 ... or 1, 2, 3, 4, 5 ... and then 1001, 1003, 1005 ..., because the first machine would be producing sequences more alike than the latter would be, even if both machines could be described as

producing like sequences (Smart 1963). The idea being represented here is that it is certainly more likely that similar brain processes could occur within the nervous system than not.

Smart's solution, then, is quite straightforward. It is to identify the introspecting subject with their nervous system – as any identity theorist might – and argue that if the nervous system, like a machine, is able to react to its like internal processes without awareness of what constitutes the likenesses, then by definition the subject ought also to be able to.

Simple as it may be, I do not think this answer does the trick, at least not convincingly. In response to the question of how an introspecting subject might be able to notice that two of their experiences are alike in the absence of any nonphysical features, Smart's reply effectively amounts to: "Well, if the introspecting subject is identical with their nervous system and if the nervous system functions like certain machines, then in the same way those machines are able to be aware that such and such internal processes are alike without being aware *how* they are, the nervous system is able to about brain processes as well, and that means the introspecting subject must be able to about any experiences identical with those brain processes, too." But such an answer falls short for two reasons.

First, this is just to punt the issue. It does not tackle the matter at hand. Recall that the purpose of this section has been how to make sense of the second half of Smart's topic-neutral description "*which is like what is going on when.*" That concerns an introspecting subject reflecting on their experience of something – like a green afterimage – and describing it as like their experience of something else – like a green light. It concerns something the subject has access to. But for Smart to respond that the

subject is nothing other than their nervous system and that that is how it might be possible for them to notice that their experiences are alike without being able to notice *how* is to undermine the meaning behind those words, words he and Place claim the subject effectively utters upon reporting their experience of the afterimage. Indeed, the assumed meaningfulness of the topic-neutral description is the premise for the objection the introspecting subject would not be able to notice *how* their experiences are alike without access to Shaffer's nonphysical features. In the absence of such features, it was implied that there does not appear to be any means available to the subject, nothing that they have access to, in order for them to meaningfully describe how their experiences are alike. But Smart's response was merely to say that there must be, that the subject must be able to notice that their experiences are alike provided we thought of them as their nervous system and their nervous system like a certain kind of machine. If the subject is identical with their nervous system, then they must have access to their like internal processes.

Second, this line of thinking also runs us into precisely the same difficulties we encountered at the end of section 1.1. It is not so clear how to make sense of how the introspecting subject might have access to the internal processes of their nervous system, their brain processes. We discovered that Place and Smart cannot – and need not – argue that the subject has access to said processes in the same way that a scientist stood behind them looking into their skull does. The subject also has no access to any phenomenal properties – no nonphysical features – being that Place and Smart are physicalists. But that only meant that the conclusion we had to settle on was that the subject has access to physical properties *as* phenomenal properties, which are neither nonphysical on the one hand nor physical in the sense in which they are for the scientist

stood behind them on the other. We described these properties as existing in a sort of “metaphysical limbo” where we could not sufficiently make sense of them as either nonphysical nor physical.

All that said, it certainly remains possible for an introspecting subject to notice similarities between their experiences without being able to say much more. Subjects utter phrases like “this looks familiar” all the time without being able to fully qualify how it might be. But that is quite clearly not what Place and Smart claim the subject means when they report their experience of an afterimage. They claim that what the subject is effectively uttering when they report their experience of a green afterimage is *“there is something going on which is like what is going on when I see a green light.”* What their subject utters goes beyond “this looks familiar” in that it establishes exactly how it is familiar, yet neither Place nor Smart are able to adequately substantiate how such a qualification might make sense in the absence of Shaffer’s nonphysical properties. At most, what we get is the argument that because the introspecting subject is identical with the nervous system and because the nervous system is able to react to likenesses of its internal processes without doing much more, the introspecting subject must be able to as well.

CHAPTER 2

PLACE'S SCIENTIST

This chapter investigates what Place claims a scientist is able to explain about an introspecting subject's nonveridical experiences. I will be concerned here not with what Smart's topic-neutral description means to its utterer but what it refers to. When a subject reports their experience of a green afterimage, Place tells us that all a scientist has to do in order to explain the experience is demonstrate that the brain process occurring at that moment is the same sort of brain process which occurs when the subject is observing a green light.

Now, it certainly seems conceivable how this might work. Suppose science identified a particular brain process which typically fired upon the observation of the color green. Certainly, whenever a scientist observed the firing of that brain process within a subject's brain in the absence of any external physical stimuli, it would stand to reason the subject was experiencing some green quality, perhaps a green afterimage. Indeed, it could even be said the scientist would be able to explain the subject's experience entirely by observation of that brain process.

But suppose things were a little different. Suppose the subject was having a more complex nonveridical experience, perhaps one composed of multiple properties. Would Place's scientist be able to explain the subject's experience by observation of their brain processes, at least to the same degree they can ostensibly do so about experiences of green afterimages? I do not believe so. In this chapter, I will show that there at least some experiences a scientist can never adequately explain.

2.1. Observing Brain Processes

Toward the end of his seminal paper, Place claims that when an introspecting subject reports their experience of a green afterimage, all a scientist has to do in order to explain the experience is demonstrate that the brain process occurring at that moment is the same sort of brain process which occurs when the subject is observing a green light. Now, before diving into the matter, some clarification is necessary. If during the subject's observation of the green light the greenness of the light is entirely a property of the light itself, as the phenomenological fallacy would suggest that it is, it is not clear what it is about the brain process occurring at that moment that then makes it so that its recurrence at a later date in time results in the subject experiencing a green afterimage. Place is not explicit on the issue. The simplest and shortest answer might be to describe the brain process taking place during the subject's observation as a specific sort of neurological reaction to the greenness of the light, such that whenever the reaction reoccurs for reasons science will presumably eventually be able to explain, the subject has the experience of a green afterimage. This seems like a reasonable response, one which Place would almost certainly give himself. Well, though all of this may appear a satisfactory account of simple experiences like those of green afterimages, things begin to fall apart when we consider experiences more complex.

What of hallucinations? Suppose an introspecting subject reports hallucinating a pink elephant with yellow spots. Obviously, Place's scientist would not be able to explain the subject's experience by reference to the brain process which occurs upon the observation of a pink elephant with yellow spots. There are no pink elephants with yellow spots around to observe, and it is certainly possible the subject might have never

even seen a representation of one.³ So how might the scientist explain such an experience? Well, Place could respond that though pink elephants with yellow spots cannot be observed, the various properties which make up the hallucination of one can be. For the sake of argument, I will assume said properties number only four: one of the elephant or “elephantness,” one of the spots or spottiness, and two of the colors pink and yellow. It is certainly likely that the subject has at one point or another observed each of *these* properties (or at least representations of them). So, upon observation each property would have induced in the subject a specific brain process resulting in four different brain processes. The scientist could certainly explain the subject’s hallucination by reference to the reoccurrence of these brain processes, Place might conclude. Unfortunately, this breakdown does not leave the scientist much better equipped to explain the experience.

Let us suppose that when a subject reports their hallucination of a pink elephant with yellow spots, all a scientist has to do in order to explain the experience is demonstrate that the brain processes occurring at that moment are of the sort which occur when the subject observes elephants, when they observe spots, and when they observe pink and yellow lights. We immediately run into a minor issue here: Place and Smart tell us in the case of the experience of a green afterimage that only the experience but not the afterimage itself exists, and that the experience is strictly identical with the corresponding brain process. Well, here there are four relevant brain processes. Which of them serves as the experience? It cannot be any single one of them if the scientist needs all four to explain the experience. But there is no need to press Place further on

³ Of course, Place could here respond that were pink elephants with yellow spots to exist the observation of one *would* result in a specific sort of neurological reaction. But this is neither here nor there. Subjects nonetheless hallucinate all sorts of things which do not exist.

the issue. It is certainly conceivable – if not likely – that multiple brain processes should combine to serve as an experience of a hallucination – what remains is for Place to spell out how.

But all this only leads to a bigger, more serious problem. If it is possible multiple brain processes combine to serve as an experience of a hallucination, that is, if all the scientist has to do in order to explain the subject's experience is point to the four brain processes, it does not appear as if they can ever really tell us much. The hallucination of a pink elephant with yellow spots is not being described by the subject as similar to the experience of observing an elephant, *and* similar to the experience of observing spots, *and* similar to the experience of observing a pink light, *and* similar to the experience of observing a yellow light. The subject is describing their experience of a single entity – a pink elephant with yellow spots, not “elephantness,” and then spottiness, and then pinkness, and then yellowness. But that appears to be as much as the scientist observing their brain processes could ever say. They could point to the four brain processes serving as the subject's experience of the hallucination, but that would not be explaining exactly what it is that the subject is having an experience of.

Smart's topic-neutral description only elucidates the problem. The subject's report certainly does not mean “*there is something going on which is like what is going on when I see an elephant, and when I see spots, and when I see a pink light, and when I see a yellow light.*” They are reporting a hallucination of a pink elephant with yellow spots. That is what they mean. They certainly do not mean they are having a hallucination of an elephant, and then spots, and then the color pink, and then the color yellow.

Now, Place and Smart could respond here that all this is just an issue of reference, not referent. Just because all the scientist may be able to tell is that the subject is experiencing a hallucination involving elephants, spots, and the colors pink and yellow, does not mean that the subject is not in fact hallucinating a pink elephant with yellow spots. Both the scientist's description and the subject's report would still be referring to the same thing. But the issue here is not whether the two are referring to the same thing but what the scientist can explain of what that referent is. Place tells us of a subject's experience of a green afterimage that a scientist's observation of the subject's brain would explain the experience. But things do not seem as straightforward in this case. To put this more clearly, suppose the subject is instructed not to report their experience and that the scientist stood behind them is tasked with explaining it. The scientist observing brain processes to do with elephants, spots, and the colors pink and yellow, could *guess* that the subject is hallucinating a pink elephant with yellow spots, but they could also guess a myriad of other combinations. They would have no way of explaining exactly what the subject is having an experience of. They would be lacking essential information regarding the subject's experience, which in the case of the subject experiencing a green afterimage, they would not.

Place and Smart could return yet again and maintain that this supposed problem is not a problem at all. They could argue that the issue we are facing has to do with our conflating the difference in meaning between statements about sensations and statements about brain processes, which they claim are not translatable into one another. The subject's report of a hallucination of a pink elephant with yellow spots means one thing and the scientist's observation of however many brain processes are taking place at that moment in time means another. It would not make sense they might contend to

try to “find” the hallucination in the brain so to speak, just like it would not make sense to try to “find” the brain processes in the hallucination. They might even argue that all sorts of combinations of brain processes may reflect in introspecting subjects as experiences of various hallucinations, and that it matters not which hallucination is specifically reported so long as it is accounted for by the presence of brain processes. The introspective report does not ultimately matter. What matters is what is physically going on inside the subject’s skull.

But such a response misses the point. We are not here concerned with identifying in the subject’s brain what their report means and vice versa. It is clear enough that statements about sensations mean things entirely different from statements about brain processes. What is in question here is only what the scientist observing the subject’s brain is able to explain. Though they may well observe all sorts of brain processes occurring, they have no way of determining, simply by that observation, what the subject is exactly having an experience of. But it is certainly relevant to the subject which of those hallucinations they are experiencing: they report experiencing the hallucination of a pink elephant with yellow spots, not anything else. This is not a matter of mere meaning. We began this section with Place’s claim that a scientist is able to explain a subject’s experience of a green afterimage by observing the subject’s brain. Well, it quite simply does not appear the scientist can tell us much if we consider something more complex. If the experience of the hallucination of a pink elephant with yellow spots is nothing over and above a combination of brain processes, the scientist observing the brain ought to be able to explain it. But they cannot. All they see are brain processes. Place’s scientist must be able to account for why the subject reports the one

experience and not anything else if Place hopes for a physicalist account of consciousness.

CHAPTER 3

REFLECTIONS

This chapter serves as a response to the previous two. Divided in two parts, the first will consider whether or not redefining sensation reports as expressives rather than reports helps avoid the problems encountered in chapter 1. If it is difficult to account for an introspecting subject's report of their experience of a green afterimage using the physical tools Place and Smart provide us, if we are asking too much of physicalism, then perhaps the physicalist can simply ask less of the report, in this case by turning it into an expressive. I will consider both the advantages and disadvantages of such a move, but argue that the latter make it one too costly to make from the perspective of an identity theorist, specifically from that of Place and Smart.

The second part will engage with the implications of the central problem identified in chapter 1, that of not being able to convincingly make sense of an introspecting subject's utterance of the topic-neutral description given Place and Smart's physicalist premises. Here is where I will propose what I see as a paradox facing Place and Smart. The goal will not be to imply that their identity theory does not or cannot work, but only to elucidate a fundamental difficulty they will need to contend with in making it work.

3.1. Redefining Reports as Expressives

In his paper "Sensations and brain processes: A reply to Professor Smart," W.D. Joske makes a distinction between reports and expressives. Subjects, he explains, make reports on the basis of something going on but they make expressives simply because

something is going on, but not on the basis of that something. A report is made on the basis of what its utterer knows or believes to be the case while an expressive merely indicates that something is going on within its utterer (Joske 1960).

Perhaps redefining reports of things like afterimages as expressives may help us avoid some of the more serious problems encountered in chapter 1. Recall how in sections 1.1 and 1.2 we tried to make sense of Smart's topic-neutral description of sensation reports by dividing it in two and working out how either half could mean what Place and Smart tell us it means given the physicalist premises they employ. We ran into serious trouble on that front, concluding that if the description did make sense – that if the introspecting subject's utterance meant the subject was (a) having an experience they (b) could compare to another – it did so on the basis of murky metaphysics. Redefining reports of things like afterimages as expressives could alleviate much of this semantic burden reports carry.⁴

Now, it is not entirely clear what an expressive exactly is. Though Joske contrasts it from a report, we could do with an example. Fortunately, at the very beginning of his paper, Smart himself explores the whole issue, providing one. Perhaps, he wonders, there is nothing more to a report of something like an afterimage than its behavioral expression, and that we ought to understand it as nothing more than a sort of sophisticated pain-wince (Smart 1959). The problem with this is deducing how much work “sophisticated” is doing here, and what it means. Smart does not expand on the matter. Though it would be accurate to understand an expressive as something at least to some degree involuntary – much like a wince as a result of pain is – it would also be

⁴ None of this helps us with the central problem of chapter 2, of course, that of how seemingly little a scientist observing a subject's brain can explain about the subject's experience. Defining the subject's utterance as either a report or an expressive has no bearing on the nature of their experience except to say that the former is based on knowledge or belief about it while the latter is not.

more sophisticated. I think the best course of action given this lack of clarification might be to define an expressive as what it is not rather than what it is. What suffices for our purposes in this section is how an expressive falls short of being a full-fledged report. An expressive of an afterimage, then, would be something (a) which was not made on the basis of what its utterer knew or believed to be the case (this, we take from Joske) and (b) which was not entirely voluntary (this, we take from Smart).⁵

Two important points, however. First, our definition is clearly not true of all expressives. Suppose one winced not as a result of pain but because they tasted some food they thought was disgusting. That disgust-wince, that expressive, *would* be based on what they knew or believed to be the case, namely, that the food was disgusting. There would also be nothing involuntary about it. Indeed, the only difference between it and a full-fledged report would be that though both were based on what the subject knew or believed to be the case, the former does not report that knowledge or belief. Instead of uttering “this food tastes bad,” the subject would only wince or exclaim “yuck!” But such an expressive does not help us. Remember, Smart’s topic-neutral description of sensation reports is only a translation. When a subject reports their experience of a green afterimage, Place and Smart tell us that what the subject is *effectively* uttering is “*there is something going on which is like what is going on when I am looking at a green light,*” with the italicized words being topic-neutral. The issues we wrestled with in chapter 1 did not have to do with the very words “there is something going on which is like what is going on when,” but with the knowledge or belief they imply the subject holds, namely, that (a) they are having an experience

⁵ Of course, this definition is also true of many other things, but I do not think its breadth a problem. Again, what is important about expressives is less so what they themselves may be and more so how they are not full-fledged reports.

which (b) they can compare to another. Redefining reports of things like afterimages as these sorts of expressives, as disgust-winces, then, would not alleviate that semantic burden reports carry.

But this only leads to my second point. I am not so certain even pain-winces would. Though they do seem at least to some degree involuntary, it is not so clear they do not involve knowledge or belief. Surely, one ought to know or believe themselves to be in pain in order to wince? They ought to believe their experience to be painful. If they did not, they would not wince. But if it is also true that pain-winces are indeed to some degree involuntary, then it follows that, at least to that degree, one winces regardless of what they know or believe. It follows that, at least to that degree, it is not up to subject whether or not they wince. I must express some uncertainty here, then. It is not entirely clear to me that pain-winces serve as examples of the type of expressive Place and Smart could make use of. Furthermore, even if we accept that their being partly involuntary gives us what we need, that it abstracts knowledge and belief from the process, pain-winces are still only *partly* involuntary. The question we would then face is whether or not that degree of involuntariness abstracts *enough* knowledge and belief from the process, whether or not that degree of involuntariness sufficiently denudes pain-winces of enough knowledge and belief such that they could serve as the type of expressive we seek.

In truth, then, I am undecided. Pain-winces could go either way. Either they involve too much knowledge or belief – in which case they are not too different from disgust-winces – for Place and Smart to be able to make use of them, or their being partly involuntary sufficiently denudes them of enough knowledge and belief that they can serve as the type of expressive we need. For the purpose of the remainder of this

section, I will assume the latter. Indeed, I have no choice but to, since not doing so would mean Place and Smart could not play the expressives game, something which might rescue them from the problems of chapter 1. Not doing so would mean reports of things like afterimages would remain reports, and as we saw in chapter 1, Place and Smart's physicalist premises make it difficult to substantiate what their utterer means.

To proceed, then, how might redefining reports of things like afterimages as expressives, as sophisticated pain-winces, make Place and Smart's case more convincing – or less unconvincing? For better or for worse, the answer is quite straightforward. Recall how we dealt with the first half of Smart's topic-neutral description in section 1.1. We attempted to make sense of the utterance "*there is something going on*" by exploring how it is an introspecting subject could have an experience of something – an afterimage – Place and Smart told us did not itself exist. Working through the problem, we ultimately arrived at the conclusion that if the utterance did mean something, it did so on shaky grounds. Rendering the report an expressive would simply sidestep the issue at the very outset. There would be no need for an investigation into the nature of an experience in the first place since the utterance "*there is something going on*" would not be being made on the basis of any knowledge or belief of one. The question of how the subject could meaningfully utter what they do, how they could possibly know or believe they were having an experience, given Place and Smart's physicalist premises, would be moot.

Likewise with section 2.2. There, we investigated the second half of Smart's topic-neutral description "*which is like what is going on when.*" We questioned whether or not the introspecting subject's ability to describe two of their experiences as alike necessitated their noticing any nonphysical features. We investigated Smart's solution

to the problem and concluded that its viability ultimately rested on the same shaky grounds arrived at in 1.1. But redefining such sensation reports as expressives would similarly render that entire investigation unnecessary. The subject's utterance "*which is like what is going on when*" would not be being made on the basis of what they knew or believed to be the case about their experiences and thus would not seem to necessitate their noticing any nonphysical features (or any features whatsoever, for that matter).

Now, straightforward as this all may be, it quite clearly cannot work. It introduces more problems than it solves. First and foremost, it simply strikes of a cheap trick. Instead of presenting solutions to the problems identified in sections 1.1 and 1.2, it merely circumvents them by abstracting meaning from the topic-neutral description. Though the description remains, though it still serves as a translation of an introspecting subject's report of their experience, we are expected not to make sense of it. But this only begs the question of the purpose of the description in the first place. Its purpose has been to provide a method through which a physicalist might make sense of a subject's report. But if we ought not to make sense of it any longer, then we might as well not have translated the report at all. We could just as well have described the subject's report as an expressive and been done with it.

Second, there is something quite eliminativist about all this. If redefining a report of an afterimage as an expressive denudes enough meaning from the utterance such that it is no longer based on what its utterer knows or believes to be the case about their experience, it follows that the utterer is in error about the things they say when they describe their experience as a matter of knowledge or belief. When a subject reports their experience of an afterimage as being green, yellowish-orange, or blurry, for example, they must be in error not because they are not having any such experience but

because what they utter is not based on knowledge or belief about it. But this is quite the pill to swallow. It seems ludicrous *prima facie* that one should be in error about the very experience they are themselves having. Though a subject may certainly be in error about what their experience is about, they cannot be, at least not to any major degree, in error about the experience itself, that is, about what it at the very least appears to be about. That *is* the experience. But redefining the report as an expressive implies a major error. It implies the subject is wrong when they so much as describe their experience of the afterimage as being of a green one, or a yellowish-orange one, and so on.

Finally, it undermines the very topic-neutrality Smart assigns his description. Recall that Smart defines it as topic-neutral because he claims both physicalists and dualists are able to explain sensation reports using it. The words “*there is something going on which is like what is going on when*” are compatible with either metaphysics. The problem is they now belong to two different descriptions. Since dualists do not face the issues encountered in sections 1.1 and 1.2, to them, the description remains a report. Replete with phenomenal properties, a dualist metaphysics would provide the means necessary for the description’s utterance to make sense. But, to physicalists, the description is now an expressive. But Smart’s assertion that the description’s topic-neutrality is what allows physicalists and dualists to talk “happily together” about their aches, pains, and so on, can only be true if they share the same description, which is to say mean the same thing.

I would like to emphasize this point. If Place and Smart tell us that sensation reports can be translated into the topic-neutral description, then what they are committing to is a congruence in sense between the physicalist and the dualist. Of course, the point of the description’s topic-neutrality is that it is compatible with either

metaphysics. The referent of the description, what is “*going on*,” is a brain process to the physicalist and phenomenal properties to the dualist. But they both nonetheless understand the translation in the same way, which is to say share the same sense, as the introspecting subject identifying their experience and likening it to some other experience of theirs. This commitment precludes Place and Smart from reducing reports to expressives. The description cannot remain topic-neutral when the physicalist understands the subject’s utterance one way and the dualist another.

Smart’s hunch, therefore, that it would not do to reduce sensation reports to expressives, was a good one (Smart 1959). Still, I think what we ought to take from all this is that a move of this sort is always open to the physicalist. If it is difficult for the physicalist to account for an introspecting subject’s report, if we are asking too much of physicalism, then the physicalist may always respond by simply asking less of the report, in our case by turning the report into an expressive. Of course, however much success the physicalist might derive from this depends on their commitments. A physicalist whose understanding of the report is shared, at least in part, with a dualist will naturally have a harder time redefining the report to better fit physicalism while also leaving it something the dualist could still make use of. But a physicalist who makes no such commitments to the dualist, who perhaps argues that the dualist gets the subject’s report entirely wrong, has more room to toy with it as they see fit. I think I have sufficiently demonstrated that Place and Smart are of the former kind, and as such would have an especially hard time doing so.

3.2. The Paradox

I think it important to begin this section with a quick recap of those parts of Place and Smart's project that have been most relevant to us. Recall that Place introduces what he terms the phenomenological fallacy, "the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste or feel to him, he is describing the literal properties of objects and events on a peculiar sort of internal cinema or television screen" In reality, he explains, the opposite is the case: conscious experience is described not in terms of any phenomenal properties, but by reference to the actual physical properties which give rise to it. The phenomenological fallacy is what leads to the assumption of an introspecting subject reporting a green afterimage, for example, that there literally exists some mental object within them which is green. But the report of the afterimage is merely the report of an experience of the sort the subject normally has when they are observing a green light.

Smart puts the idea more clearly with his topic-neutral description of sensation reports, explaining: "When a person says, "I see a yellowish-orange afterimage," he is saying something like this: "*There is something going on which is like what is going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange.*"" Being topic-neutral, the italicized words could refer either to brain processes or phenomenal properties, though they in fact refer to brain processes. All a scientist needs to do in order to explain the subject's experience of the green afterimage, Place concludes, is demonstrate that the brain process occurring within them at that moment is the same sort of brain process which occurs when they observe a green light.

Now, the manner in which Place and Smart make their case for identity theory ought to be pretty clear. Their strategy is to challenge dualism and allow physicalism to fill in the gaps. They do so through two avenues. First, they do so through the phenomenological fallacy. It is a mistake to assume the existence of phenomenal properties when an introspecting subject reports their veridical experience. Just because we may know green apples by our experience of them, it does not follow that our experience is itself green. Second, they do so through the topic-neutral description of sensation reports. When a subject reports their nonveridical experience, their report points to merely topic-neutral properties, not phenomenal ones. The report is only compatible with phenomenal properties; what is “*going on*” *could be* referring to phenomenal properties. Once we learn to avoid these misconceptions, it does not seem inconceivable that experiences should be strictly identical with brain processes. What remains is to work out how physicalism might fill in those gaps the absence of any phenomenal properties leaves, gaps that concern how to make sense of a subject’s experiences.

The obvious problem is this strategy has not been very convincing. The purpose of chapter 1 has been to demonstrate that. In it, we investigated the issue of how an introspecting subject’s effective utterance of the topic-neutral description might make sense given Place and Smart’s physicalist premises. We found that we could not make much sense of it. Without the aid of phenomenal properties, we found that the utterance only made sense on the basis of murky metaphysics.⁶

⁶ Being only tangentially related, the central problem of chapter 2 does not feature here. To recall, that was the problem of how seemingly little a scientist observing a subject’s brain can explain about the subject’s experience. But the existence (or lack thereof) of phenomenal properties has little bearing on that issue.

Well, for the purpose of this section, I would like to suppose that none of this was the case. Let us suppose that in making their case for identity theory Place and Smart had not employed the strategy identified above and chipped away at phenomenal properties, first through the phenomenological fallacy and second through the topic-neutral description of sensation reports. Let us suppose that they had instead admitted them. How might things have turned out differently? I would like to make the case that this would have put Place and Smart in an even worse position, and that, as such, they are left stuck in a sort of paradox.

Let us consider the example we have all the while been using, that of an introspecting subject reporting their experience of a green afterimage. Suppose that Place and Smart agreed that the subject's experience was now saturated with phenomenal properties. It would now make complete sense as to how their report might be translated into the topic-neutral description. When the subject effectively uttered "*there is something going on which is like what is going on when I am looking at a green light,*" it would make sense as to how they might. They would have access to the two sets of phenomenal properties necessary to meaningfully make that assertion, the first being the greenness of their experience of the afterimage and the second being the greenness of their experience of a green light. They would be able to liken their experiences to one another on account of both of them being green.

Prima facie, this would appear to solve Place and Smart's problems, at least those we identified in chapter 1. The introspecting subject would have all the tools necessary to meaningfully utter the topic-neutral description, and everything would still be taking place within a physicalist universe. Those green phenomenal properties would still be nothing over and above neurological properties of brain processes. The problem

with all this is that it would strip the physical side of the identity relation of any meaningful efficacy in accounting for the subject's experience. The brain processes would have very little to do.

Recall that in chapter 1 we investigated how an introspecting subject's effective utterance of the topic-neutral description might make sense given Place and Smart's physicalist premises. That was an investigation into how brain processes and brain processes alone could manifest in the meaningful utterance of the description. In section 1.1, we had trouble accounting for the first half of the description, which concerned the subject having an experience they could identify. Without the existence of the afterimage, we were forced to settle on some murky metaphysics. We did not fare much better in section 1.2. That was about the second half of the description, which concerned the subject likening their experience to another. There, Smart provided us a physicalist story as to how they might, one which concerned thinking of the subject as nothing over and above a nervous system and the nervous system as a sort of machine. We found this unconvincing, not least because it committed us to the same murky metaphysics section 1.1 did.

Now, regardless of whether or not we thought it convincing, the case being made by Place and Smart in chapter 1 was nonetheless clear. They were positing a physicalist account using physical tools of nonveridical experiences. But introducing phenomenal properties into the mix would render much of that story unnecessary. There would be no need to explain how it might be possible that an introspecting subject could recognize their having an experience in the absence of an actual afterimage, or how they could liken that experience to another in the absence of any nonphysical features. There would be no need to explain how brain processes could manifest in such and such a way

if only we think of them as this and that. All the phenomenal properties necessary to make sense of the topic-neutral description would be available, and they would be nothing over and above neurological properties of said brain processes.

But exactly here lies the problem with this move. Introducing phenomenal properties into the mix would eclipse all the explanatory power brain processes – the physical side of the identity relation – formerly had to offer. They would have nothing to contribute to Place and Smart’s physicalist story other than to say that their neurological properties were what the phenomenal properties were nothing over and above. But all this only leaves Place and Smart worse off, not better. To strip the physical side of the identity relation of all explanatory power and give it over to the mental side leaves their identity theory weaker. It amounts to arguing that dualists are in fact correct about the nature of reality and experience, but that it just so happens that all the mental properties they assert are nothing over and above physical properties.

I think it important to emphasize what I am not arguing. I am not arguing that were Place and Smart to admit phenomenal properties, brain processes would no longer be the ultimate reason the introspecting subject was able to meaningfully utter the topic-neutral description. I am not arguing that the universe would not still be a physical one. What I am arguing is that the existence of phenomenal properties would render those brain processes – the physical side of the identity relation – explanatorily fruitless. There would be nothing to say about them other than their neurological properties serve as what the phenomenal properties are nothing over and above. But though this would still amount to a physicalist account of reality, more specifically an identity theory, it would amount to a very poor one. It would be as compelling a case for physicalism as is the subject’s experience that the afterimage is green would be for dualism. Obviously,

just because the afterimage may seem green to the subject does not mean dualism must be true. That is precisely what Place and Smart reject. Likewise, it cannot be the case that physicalism must be true simply on account of it being conceivable that the greenness of the subject's experience is in fact nothing over and above a neurological property of a brain process of theirs. The physicalist ought to say more. The story cannot end there. They ought to make a compelling case for how a physical universe gives rise to experiences of things like green afterimages, one which does not effectively amount to accepting dualism but simply arguing that all the mental properties dualists employ are in fact nothing over and above physical properties.

I would like to make the claim, then, that Place and Smart are trapped in a sort of paradox. On the one hand, they may chip away at the mental side of the identity relation. They may explain away phenomenal properties. This gives the physical side, brain processes, much to do in explaining an introspecting subject's report of their nonveridical experience. The problem here is that the physical side does not do a very good job at it. Chapter 1 demonstrates as much. On the other hand, they may afford the mental side its full phenomenal breadth. Phenomenal properties do a much better job at explaining the subject's report. The problem here is that this leaves the physical side with very little to do. Its only relation to the whole matter is that it serves as what is in fact identical to the mental side, as what the mental side is nothing over and above. There is no longer any need to make sense of how brain processes might manifest in this and that. The phenomenal properties do all the explanatory work while the brain processes merely take the credit.

None of this is to imply that Place and Smart's identity theory cannot work, let alone that identity theory more generally cannot. I only elucidate what I see to be a

serious difficulty Place and Smart – or any identity theorist employing a similar strategy – will have to overcome. They will need to stick to their physicalist convictions and adequately explain how the physical universe is able to give rise to things like nonveridical experiences. This is something Place and Smart try to do. But in doing so they will also need to avoid affording the mental side of the identity relation too much explanatory power. Their case for physicalism cannot hinge solely on the mental side's being in fact identical with the physical side. That leaves the physical side devoid of the explanatory utility we expect of it in a physical universe.

CONCLUSION

In this thesis, I have argued that Place and Smart's identity theory fails on two accounts. I demonstrated my position in two chapters. The first was that of their account of what an introspecting subject means when they report a nonveridical experience. The second was that of their account of what a scientist can explain about the subject's experience through observation of their brain processes. In my third chapter, I considered a possible move Place and Smart could make that might help avoid some of the problems identified in the first. The chapter concluded with a presentation and explanation of the paradox I see facing Place and Smart. In presenting it, my goal was not to imply that their identity theory does not or cannot work, but only to elucidate a fundamental difficulty they will need to deal with if they are to make it work.

REFERENCES

Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47, no. 1 (1956): 44-50.

Smart, J. J. C. "Sensations and Brain Processes." *The Philosophical Review* 68, no. 2 (1959): 141-56.

Smart, J. J. C. "Materialism." *The Journal of Philosophy* 60, no. 22 (1963): 651-62.

Smart, J. J. C. "Further Thoughts on the Identity Theory." *The Monist* 56, no. 2, (1972): 149-162.

Smythies, J. R. "A Note on the Fallacy of the 'Phenomenological Fallacy'." *British Journal of Psychology* 48 (1957): 141-144.

George Pitcher & W. D. Joske. "Sensations and brain processes: A reply to Professor Smart." *Australasian Journal of Philosophy* 38, no. 2 (1960): 150-160.

Shaffer, Jerome. "Mental Events and the Brain." *The Journal of Philosophy* 60, no. 6 (1963): 160-66.