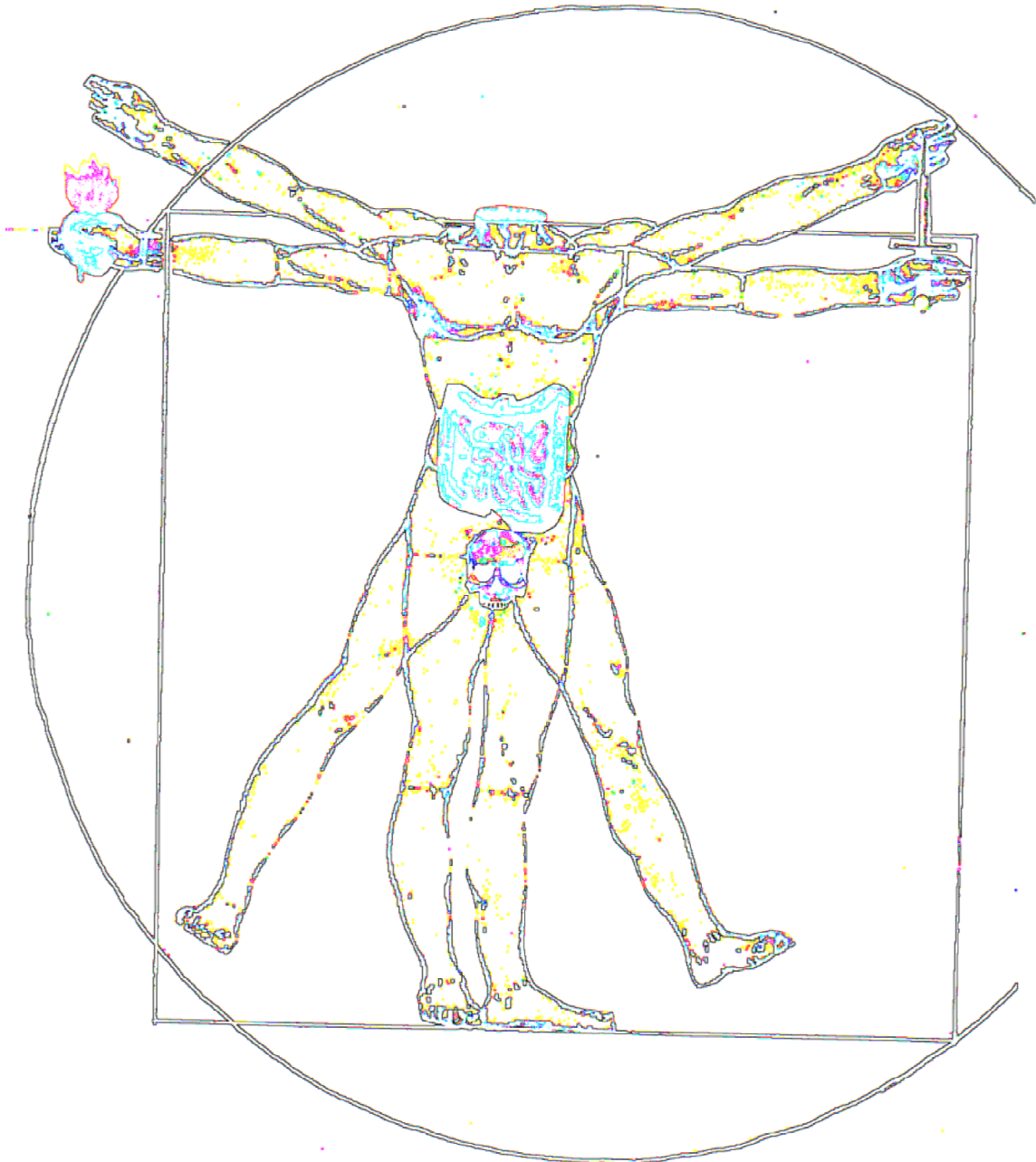


COGITO

j o u r n a l o f p h i l o s o p h y



Vol. IV (New Series) No. 3, 2008

A joint publication of the Socratic Society of the University of New South Wales, the Russellian Society of the University of Sydney, and the Platonic Society of Macquarie University.

COGITO

Journal of Philosophy

COGITO is the journal of the Socratic Society of the University of New South Wales, the Russellian Society of the University of Sydney and the Platonic Society of Macquarie University.

Vol. IV (New Series) No. 3, published in Sydney, 2009

© 2008, remains with the authors of the articles

ISSN 1833-2005

Contributors

Swantje Lorrimer
Steph Rennick
Melanie Rosen
Stef Savanah
Tim Smartt

Editors

Marco Crespo
Andy Crosbie
Rishi Kristnan

Layout and design

Marco Crespo

Cover art

After Masson after da Vinci 2009

Cover artist

Marco Crespo

Reviewers

Lise-Marie Anderson
Rick Benitez

Karen Bland
Damion Buterin
John Cusbert
Anson Fehross
Chris Maden
Monte Pemberton
Elizabeth Schier

Acknowledgements

Wilson Cooper
Ben Schulz
Mikey Slezak
Mary Walker

Contacts

Platonic Society

<http://au.geocities.com/platonicsociety/>

Russellian Society

<http://www.russoc.com/>

Socratic Society

<http://geocities.com/socraticssociety/>

COGITO is produced by students and is one of the main activities of the year for the three Societies. The first journal was published in 1966 and can be found in the UNSW Library under S105/19. It is generally comprised of articles, interviews, essays, and book reviews submitted by students of all levels. Sometimes, an academic may contribute an article to the journal.

Contact the Societies for past issues of COGITO.

COGITO is a peer-reviewed journal. The editors are forever grateful to our referees for kindly volunteering to donate some of their precious time toward this project.

COGITO is generously supported by funding from the Department of Philosophy, Macquarie University.

COGITO

J o u r n a l o f p h i l o s o p h y

Vol. IV (New Series) No. 3 2008

Essays

- Colour Perception: It's All in the Mind 3
Stef Savanah
- In Defence of Mysterianism 12
Melanie Rosen
- What Can Kaplanians Learn from Perspectivist Accounts of Creativity? 24
Swantje Lorrimer
- Harry Potter and the Time Travel Paradox: an analysis of predestination and ontological paradoxes as encountered in J. K. Rowling's *Prisoner of Azkaban*. 38
Steph Rennick

Book Review

- A New Platonic Theism? 53
Tim Smartt

Editorial

Welcome to the 2008 edition of **C O G I T O** . In this issue we have contributions from students at various levels of study writing on a diverse range of topics in philosophy. We would first of all like to extend our gratitude to all of our contributors for making submissions to **C O G I T O** . Without students submitting their work there would be no student journal of philosophy in Sydney. Second, we would like to sincerely thank all the referees for sacrificing some of their valuable time to review the submissions. They presented us with sound advice so that we were able to select and publish papers of the highest standard.

Steph Savannah's paper explores the ontological status of colour. The problem is whether colour is actually in the world, the objectivist position, or is a mental phenomenon, the subjectivist position. Savannah adopts the subjectivist view. He argues that the subjectivist view is vindicated if we take into consideration the apparent constancy of an object's colour under low luminosity, an argument employed by objectivists in defence of their position, in conjunction with the results of the "coloured shadow effect" experiment.

On the very possibility of being able to know the unity of brain and mind, Melanie Rosen considers the view of mysterianism. Mysterianism asserts that we are cognitively closed to accessing the link between introspection and brain activity. First, Rosen evaluates whether we can be cognitively closed. Second, in a subversive twist she argues that mysterianism can give support to physicalism, since knowledge about the brain will only increase, but not to the extent that science could give us the very experience of introspection, which is, ontologically speaking, tethered to brain activity.

Swantje Lorrimer discusses the involvement of creativity in the evolution of natural language. Drawing into her analysis Fregean mediated theory of reference, as well as modern direct reference theorists such as Kapitan and Hintikka, she argues that the polysemia of indexicals and demonstratives presuppose creative speaking agents.

We are sure that many readers of **C O G I T O** are devoted fans of J. K. Rowling's *Harry Potter* series. Steph Rennick explores the paradoxes that arise when we think about the possibility of time travel. Rennick mediates her discussions about the philosophical problems involved with time travel by recourse to Rowling's 3rd book in the Harry Potter series, *Prisoner of Azkaban*. A number of theories, including the 4-D view of time and the mutable universe hypothesis, are brought to bear on two contradictions of time travel, the predestination paradox and the ontological paradox, in which Harry is caught up.

Finally, Tim Smartt reviews Daniel A. Dombrowski's *A Platonic Philosophy of Religion*, published by SUNY Press, Albany in 2005. Contrary to the traditional interpretation of Plato's contribution to theology, one in which it is alleged that Plato posits eternal, static being in opposition to fleeting, mutable becoming, Smartt draws our attention to Dombrowski's innovative reading of the divine in the works of Plato. Dombrowski claims that what we actually find in Plato is a pantheism that is dipolar in structure. The world is "in" God. Being and becoming are the correlated principles of the world such that the divine is actualised in the unfolding process of life itself.

Essays

Colour Perception: It's All in the Mind

Stef Savanah
MPhil Candidate
Macquarie University

Abstract

Is colour in the world or in the mind? Intuitively, colour would seem to be a property of objects, but in this paper I will argue that colour is a construct of the mind. The latter view may be termed *subjectivist*, due to its dependency on the existence of subject. In contrast we may call the former view *objectivist*. I concentrate on the two strongest arguments on each side: *metamerism* (the ability of different spectral profiles to cause the same colour experience in an observer), which argues for subjectivism, and *colour constancy* (the persistence of an object's apparent colour under poor lighting conditions), which argues for objectivism. I show that colour constancy, when combined with the results of an experiment known as the "coloured shadow effect," actually supports the subjectivist position.

I: Introduction

The first question that must be asked is "why is colour important to philosophy?" Some philosophers contend that it is a natural starting point for investigations into appearance and reality.¹ To my mind it goes to the very essence of perceptive consciousness; if such an apparently common sense physical reality as colour turns out to be a "figment of our imagination" we must give pause to much of our other such preconceptions. The central question, then, is whether colour is "in the world" (is a property of objects) or only exists as an experience in our minds. The latter position is known as subjectivism as it holds that colour only exists as a subjective experience. The opposing view, in which colour is an inherent property of objects in the world, irrespective of the existence of subjects, is known as objectivism. The question is itself open to interpretation as we shall see when we examine the various viewpoints, so for the purposes of this paper I will first establish a very strict criterion for the objectivism/subjectivism determination. I will then briefly review the main competing philosophical theories of colour and their strongest arguments/objections. A quick look at the science of colour will throw some light on the discussion. With this background I will then lay out my argument to show that colour is in fact all in the mind.

¹ Byrne, Alex and Hilbert, David R (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, p. xi

II: Objectivism vs. Subjectivism

Let us break the chain of colour perception into three parts: the *object*, the *signal* that transmits colour information about the object (i.e., light) and the *receiver* of the signal. (I deliberately choose "receiver" rather than "percipient" to allow that the receiver could be an instrument.) By object I include anything that can emit a light signal by reflection or fluorescence or in any other way, and so include flames, sky, etc. *The criterion for establishing that colour is "in the world" (is a property of the object) is that the object can emit a signal identifying its colour independent of there being a subject to perceive it.* The object does not need to be actually transmitting a signal (for example, if it happens to be in an unlit room at any time) but it must have the ability to do so. If the above criterion cannot be met, then we will have established that colour is a subjective effect. At first sight this criterion would seem to favour the objectivist viewpoint as, after all, objects do emit light and the wavelengths of light are related to colour. However, as we shall see, the difficulty is in the phrase "identifying its colour."

Philosophical theories of colour do not neatly fall into two opposing camps; indeed, Barry Maund identifies no less than nine distinct competing theories in the spectrum.² I consolidate these into three broader categories that I name the *pure objectivist*, the *hybrid* and the *pure subjectivist* theories. In pure objectivist theories colour is an intrinsic property of the object and is perceiver-independent. In pure subjectivist theories colour is not a property of the object and is perceiver-dependent. In hybrid theories colour is both a property of the object *and* perceiver-dependent. In line with my criterion, as hybrid theories involve a perceiver they will be classed as subjectivist. Let's now examine the *pure objectivist*, *hybrid* and *pure subjectivist* views in a little more detail.

In pure objectivist theories of colour there are several ways in which colour may be deemed to be a property of the object. Hacker argues that colour is not a 'property of seeing but of what is seen.'³ In this extreme view the property of colour is inherent in the object and is not reducible to microstructures or in any other way. A less extreme view holds that colour is a property of objects that is reducible to 'complexes of such properties of the object's surface and immediate surroundings.'⁴ In other words, colour is a direct consequence of the microstructures of an object. In yet another view, colours are dispositional properties of objects: powers to induce an experience of colour in a perceiver. A variant of this view is that the disposition is to emit light of certain characteristics that can be justifiably labelled with a colour predicate. Isaac Newton held that colour is neither a property of the object nor is it perceiver-dependent, but rather it is a property of the light ray (or *signal* in my parlance.)⁵

² Maund, Barry. "Color," *The Stanford Encyclopedia of Philosophy* (Fall 2002 edition), Edward N Zalta (Ed.) <http://plato.stanford.edu/archives/fall2002/entries/color/>

³ Hacker, P. M. S. *Appearance and Reality: a philosophical investigation into perception and perceptual qualities* Basil Blackwell Ltd., Oxford, 1987, p. 116

⁴ Jackson, F and Pargetter, R. "An Objectivist's Guide to Subjectivism about Colour," in Alex Byrne and David R. Hilbert (eds.) *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 67-80

⁵ Thompson, Evan. Chapter 1 "The Received View," in *Colour Vision* Routledge, London and New York, 1995, p. 12

Colour Perception: It's All in the Mind

As the properties of this light ray are a direct consequence of the object from which it is emitted, I group this view with the objectivists. In all these cases, of course, the common ground is that colour is perceiver-independent, which is what defines them as objectivist by my criterion.

Early objections to the objectivist view relate mainly to the disparity between an object's microstructure and its macrostructure. At the micro level the surface structures of an object do not resemble the properties that colours have. There does not appear to be a way to predict an object's colour based on the composition of its surface. Furthermore, many different physical microstructures can lead to objects appearing the same colour, an objection known as the Problem of Multiple Realisations. To my mind these types of objections lack strength. I see no reason to require an obvious isomorphism between an object's microstructures and its colour appearance. It is feasible that identical or nearly identical light signals might be emitted from differently structured objects. For example, the reflection from a pane of glass can look identical to that from the surface of a pool of water, whose microstructures are clearly different. An objectivist need only point out that the colour of an object is determined by the frequency (or frequencies) of light that emanates from it, and this property of the emanating light signals will partially be determined by the action of an object's microstructure on reflected or transmitted light. In fact, in more recent times the frequency of emanating light may be predictable given certain knowledge of microstructures, such as the electron absorption levels of the constituent elements of the object. A discussion of this topic, however, is beyond the scope of this paper.

The only truly devastating argument against objectivism is *metamerism*. That is, the fact that objects emitting completely different wavelength patterns (i.e. different colour signals) can appear the *same* colour to human observers. This seems to imply that the colour experience depends on a construction in the mind rather than the particular signal transmitted by the seen object. Otherwise, we would expect that different patterns of light frequencies should induce different colour experiences in the same observer. We'll come back to metamerism later, but first we'll take a look at the subjectivist position.

Pure subjectivist theories (sometimes referred to as projectivist) are far more common amongst scientists than philosophers.⁶ In his Stanford Encyclopaedia article, Maund cites only one pure subjectivist case in his list of theories, which he attributes to J. van Brackel. In these theories colour is a projection by a subject onto the world. More common among philosophers are hybrid theories, such as those propounded by Merleau-Ponty.⁷ In such cases colour is in the world but is also phenomenological. The light signal emanating from an object depends on the object, but the way individual observers experience the resultant colour is unique to the individual. Indeed, many of us as children have wondered at the

⁶ Boghossian, P. A and Velleman, J. D. "Colour as a Secondary Quality," in Alex Byrne and David R. Hilbert (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 94-95

⁷ For example, see Maurice Merleau-Ponty, Chapter 5 "Eye and Mind," in James M. Edie (ed.), *The Primacy of Perception: : and other essays on phenomenological psychology, the philosophy of art, history and politics* Northwestern University Press, Evanston, Ill, 1964

privacy of our own subjective experience, and that for all we know the colour we all label as “red” may be experienced in completely different ways by different people.

Many arguments against subjectivism are along the lines that the consistency of our colour experiences makes the objectivist view more plausible. For example, subjectivists must explain *why* we usually take colours to be objective, if they are truly a construct of the mind. It seems problematic to hold that our colour experience is systematically mistaken. Michael Tye, although he is an objectivist, actually provides a neat response to these objections using the example of optical illusions, such as those that make straight lines appear curved.⁸ Anyone who has experienced such illusions will accept the ease with which our minds may be deceived into believing objects have properties that they in fact do not. These types of arguments, based on the apparent implausibility of consistent error in our colour experiences, do not seem to me to hold much water. Implausibility always seems to me to be less of an argument and more of an appeal. Rather, for subjectivists, a truly problematic objection concerns the phenomenon of *colour constancy*. This is the property of objects to retain their colour appearance under poor lighting conditions or even under differently coloured illumination. In a darkened room, a visible object will still appear to have the same colour as seen during normal daylight. Even when the only light available is not white but coloured differently to viewed objects, they are experienced as the same colour as when viewed in normal white light. Different lighting conditions may affect of vision generally, yet viewed objects retain their disposition to emanate light that determines their colour. This seems to support the view that the seen colour is an intrinsic property of the object: the object’s colour persists under different lighting conditions because it “belongs” to the object.

I propose to concentrate henceforth on the strongest arguments on both sides: metamerism and colour constancy. The aim of the objectivist is to explain how colour can be said to be independent of colour percipients even though an indefinite number of unrelated and disjunctive light signals can produce the same colour experience in percipients (the problem of metamerism). The aim for subjectivists is to explain colour constancy, which strongly suggests that colours are illumination-independent properties of objects (Tye’s “common sense” view.)⁹ Before explaining my own responses to each of these arguments, it will be helpful to briefly cover some of the relevant background science.

III: Colour Science

In Newton’s classic colour experiment a beam of white sunlight is split by passing it through a glass prism, producing a line of seven “rainbow” colours.¹⁰ These colours represent *specific wavelengths of light*, covering the visible range (360-780 nanometres.)¹¹ However, our experience of colours is not a linear but a circular continuum. Colour charts are usually

⁸ Christie, R. M. *Consciousness, Colour and Content* MIT Press, Cambridge, Massachusetts, 2000.

⁹ Christie, R. M. *Consciousness, Colour and Content* MIT Press, Cambridge, Massachusetts, 2000, p. 148

¹⁰ See, for example, Evan Thompson, Chapter 1 “The Received View,” in *Colour Vision* Routledge, London and New York, 1995

¹¹ Christie, R. M, Mather, R. R and Wardman, R. H. *The Chemistry of Colour Application* Blackwell Science, Great Britain, 2000.

displayed as a "colour wheel": yellow, green, cyan, blue, magenta, red and then back to yellow again. Colours on opposite edges of the wheel are called "opponent" colours. This is because the colour of (say) a dye is caused by the dye *absorbing* light from the region of the spectrum representing the opponent colour. For example, a yellow dye *absorbs* blue light and reflects all the rest. This effect was first noticed by Arthur Shopenhauer in the late 19th century and formalised by Kek Hering in 1878. Well before then, Thomas Young (1801) and separately Hemholtz had formalised rules for producing colour by *adding* hues: a wide range of colours can be produced by mixing different proportions of the three primary colours: red, green and blue (for example, red + blue = magenta).¹² The primary colours are so deemed because they cannot be produced by mixing together any other two colours. The principle is used, for example, in colour television. The upshot of all this is that our colour experience cannot be equated to a particular wavelength of light. Christie, Mather and Wardman show spectral profiles for dyes of several colours that are not spikes at a single wavelength but any number of various curves, including double broad spikes, dips, and plateaus (for example, their magenta resembles a double dipping roller coaster). This is of course the science behind metamerism: the same colour can be produced in multiple ways, by various combinations of light at different wavelengths. For convenience, when hereafter referring to colours caused by a "single" wavelength such as when white light is split by a prism, I will use a prefix "lambda-" (the symbol scientists usually use to designate wavelength). Thus, "lambda-green" refers specifically to light of a very narrow wavelength range centred at 545 nanometres.

Metamerism is a little less mysterious when we understand how the human eye responds to colour. Of the two types of light-sensitive elements in the human retina, only one type ("cones") is colour-sensitive. There are just three variants of cones, those that are sensitive to lambda-blue, lambda-green and lambda-red respectively. There are some 6-7 million cones, which, according to Christie, enable normal observers to distinguish about ten million separate colours. Christie goes on to say that "...colour sensation perceived by the eye is governed by the response of the three types of cone cells to the particular wavelength profile with which they are interacting."¹³ Given that the eye is sensitive to only three colours it is tempting to immediately assert that the mind must be constructing the ten million possible experiences of colour, and as such that colour is subjective, but I will not rest my case yet. Let us instead allow that it should be theoretically possible to construct an instrument to mimic the response of the human eye and register colours with the same sensitivity as humans. If this is possible it would support the case for objectivism (by proving percipient-independence). We will have recourse to this idea of a "colour sensing machine" below.

IV: Metamerism Examined

J. J. C. Smart and D. Armstrong argue the "so what" response to metamerism. Their argument is along the lines that the colour property of an object is that which causes it to emit a particular spectral profile, and as such it belongs to the object irrespective of what

¹² Nassau, Kurt. *The Physics of Chemistry and Colour* 2nd edition, Wiley Interscience, New York, 2001

¹³ Christie, R. M. *Colour Chemistry* The Royal Society of Chemistry, Cambridge, UK, 2001, p. 17

colour experience it causes in a percipient.¹⁴ Now, say we built a colour-sensor machine with a receiver of colour signals based on the human eye, and the output display is a colour indicator - perhaps a pointer over a colour spectrum. Somehow this machine would need to be calibrated. We can easily calibrate it for the "lambda-" colours by simply tuning it to their specific wavelengths. For example we could tune this instrument to recognise an input of lambda-green by transmitting to it a signal at a wavelength exactly 545 nanometres. But, as previously explained, an experience of green might be elicited by different spectral profiles. How is our machine to recognise the greens caused by other spectral profiles? The machine has no objective standard against which to match a particular spectral pattern in order to identify the associated colour. Of course, we could calibrate the machine using known spectral profiles, perhaps hundreds or thousands of them. And we could even devise some clever algorithms for the machine to interpret the colours of spectral patterns that have not been specifically built in, by interpolating between the patterns that have. But these are not *objective* standards, they have been provided by human observers. Maund, referring to Hardin, makes a similar point, pointing out that any such standard would be quite arbitrary as humans do not all have the exact same colour perception.¹⁵ But in any case, arbitrary or not, my point is that the standards would depend on the perception of one or more *human observers*, i.e., they would have a direct dependence on colour-percipients. The spectral profile of the emitted signal may indeed "belong" to the object as Smart and Armstrong assert, but it *cannot* identify a colour. It always has to be or have been a human observer that identified (or "matched") the colour. A spectral pattern causes a colour experience *in the mind*; it cannot be said to "have colour." We can now see that this is true even for lambda-green: the spectral profile of lambda-green just happens to be a conveniently sharp spike at 545 nanometres, but it is still the mind that experiences this as green just as the mind has colour experiences for more complex spectral patterns.

V: Colour Constancy Explained

As an avowed subjectivist, then, how am I to explain colour constancy? I turn to the work of John Mollon who describes an experiment inspired by a lecture given to the Académie Royale in 1789 by Gaspard Monge.¹⁶ The phenomenon is known as the "coloured shadow effect." First, illuminate a white screen with a projector shining pink light. In front of the screen place a green object. Despite that fact that the green object is being illuminated essentially by its *opponent* colour, it still appears green. So far this is just a demonstration of colour constancy, and appears to support the claim that the greenness is in the object. But now consider the following experiment: illuminate a white screen with *two* projectors - one shining pink light as before and the other shining white light. Then place an opaque object

¹⁴ Smart, J. J. C. "On Some Criticisms of a Physicalist Theory of Colours," in Alex Byrne and David R. Hilbert (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 1-10; See also in the same volume D. Armstrong, "Smart and the Secondary Qualities," pp. 33-46

¹⁵ Maund, Barry. "Color," *The Stanford Encyclopedia of Philosophy* (Fall 2002 edition), section titled "Objectivism: Problems & Solutions."

¹⁶ Mollon, John. "Colour in Nature," in Trevor Lamb and Janine Bourriau (eds.) *Colour : art & science* Cambridge University Press, Cambridge, 1995, pp. 127-150. Johann Wolfgang von Goethe describes a similar experiment to Monge's in his *Theory of Colours* published in 1840. See the translation by Charles Locke Eastlake, MIT Press, USA, 1970, pp. 30-33

Colour Perception: It's All in the Mind

in front of the pink source so as to cast a shadow on the illuminated screen. The *shadow* of the opaque object appears green! To emphasise the point: there is no green source in this experiment, just pink and white, and yet the *shadow* cast by the object blocking the pink light appears green to observers. According to Mollon, the spectral profiles of these two cases are equivalent and the startling effect in each case is due to the brain *compensating* for the pink background. In the first case, the green object cannot reflect the pink light and so should appear black or be invisible, but the brain constructs a colour for it – the *opponent colour* of the background pink. In the second case the brain constructs a green colour for the shadow in the same way, according to the surrounding opponent colour, pink. Mollon appears to be the first person to link the *coloured shadow effect* to the phenomenon of *colour constancy*. The underlying cause is the same in both cases: the mind adjusts the colour experience of an object according to the background colour. This experiment demonstrates that colour experience is a construct of the mind. In this context, colour constancy actually argues the case for subjectivism rather than objectivism.

In fact, C. L. Hardin lists a host of other conditions apart from background colour that affect the sensation of colour, even when the emitted light from the object remains constant. Some examples include whether the eye and/or brain transmission channels are working correctly; if the illuminating light is flickering and if so at what frequency; the size of the object; and even what colours were viewed immediately before the object.¹⁷ So even in cases where the *same* light signals (that is, identical spectral profiles) emanate from an object, this cannot guarantee a repetition of the same colour experience in a *single* individual human observer, let alone all. Colour experience is dynamically constructed in the mind of the perceiver.

VI: An Analogy with Music

Despite the strength of the subjectivist argument and its scientific basis, it is still difficult to grasp the concept of colour being in the mind rather than in the object. I will present an analogy to make this easier. First though, a general comment on the usage of analogies by proponents on both sides. It seems to me that they are often rather poorly chosen, for example, the aforementioned analogy by Tye regarding the “curved line illusion.” Tye’s intention was to demonstrate that just because colour is relative to its background that does not make it “in the mind.” His rationale being that the property of being straight still belongs to a line even though it appears curved to an observer, and in the same way we can say that colour belongs to the object no matter how it is perceived by an observer. But this analogy is not cogent, because in the case of the straightness of a line we have an objective standard we can apply to verify that property - i.e. a ruler. As argued previously, we have no such objective standard in the case of colour. As Kurt Nassau emphasises, it is the perception of the human observer that is the final arbiter as to whether or to what degree two colours match.¹⁸

¹⁷ Hardin, C. L. “The Philosophy of Colour,” in K. Nassau (ed.), *Colour for Science, Art and Technology* Elsevier Science, Amsterdam, 1998, pp. 209-220

¹⁸ Nassau, Kurt. *Physics and Chemistry of Color: the fifteen causes of color* 2nd ed., Wiley Interscience, New York, 2001

Stef Savanah

Now consider the analogy of music to colour. Not just sound, but music, for music is to sound as colour is to light: both music and colour carry that extra emotional content. In both cases we have an object that emits a signal. In the case of colour the transmission medium is light waves, for music the transmission medium is the perfectly analogous sound waves. In the case of colour the signal is received by the perceiver via the optic nerve, for music via the auditory nerves. The analogy is very close, and yet when a violinist produces a musical note by drawing a bow across a string no one would assert that the music is "in the string." We have no difficulty accepting that the music is not in the instrument. We say the musician "makes" music, but the music is not in the act itself; the musical experience is constructed in the mind from the perceived sound signals. The sound waves emanate from the object but the music we hear does not belong to the object. In exactly the same way we should accept that light waves emanate from an object but the colour we perceive does not "belong to the object."

VII: Conclusion

Is colour in the world? Only in the weak sense that objects have *dispositional power* to emit light signals that normal percipients will *usually* experience as a particular colour. But we can never get around the requirement for an observer: the emitted signal cannot identify the colour; colour is an experience of the mind. As the coloured shadow effect demonstrates, the mind constructs colour experience according to any number of factors. As such, by my criterion for the objectivism/subjectivism determination, as colour is not percipient-independent we must conclude that colour is indeed subjective.

Bibliography

Armstrong, D. "Smart and the Secondary Qualities," in Alex Byrne and David R. Hilbert (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 33-46

Byrne, Alex and Hilbert, David R (eds.) *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997

Boghossian, P.A and Velleman, J.D. "Colour as a Secondary Quality," in Alex Byrne and David R. Hilbert (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 81-103

Christie, R. M. *Colour Chemistry* The Royal Society of Chemistry, Cambridge, UK, 2001

Christie, R. M, Mather, R. R and Wardman, R. H. *The Chemistry of Colour Application* Blackwell Science, Great Britain, 2000

Goethe, Johann Wolfgang von *Theory of Colours* Charles Locke Eastlake (trans.), MIT Press, USA, 1970

Colour Perception: It's All in the Mind

Hardin, C. L. "The Philosophy of Colour," in K. Nassau (ed.), *Colour for Science, Art and Technology* Elsevier Science, Amsterdam, 1998, pp. 209-220

Hacker, P. M. S. *Appearance and Reality: a philosophical investigation into perception and perceptual qualities* Basil Blackwell Ltd., Oxford, 1987

Jackson, F and Pargetter, R "An Objectivist's Guide to Subjectivism about Colour," in Alex Byrne and David R. Hilbert (eds.) *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 67-80

Lamb, Trevor and Bourriau, Janine (eds.) *Colour : art & science* Cambridge University Press, Cambridge, 1995

Maud, Barry. "Color," The Stanford Encyclopedia of Philosophy (Fall 2002 edition), Edward N Zalta (ed.) <http://plato.stanford.edu/archives/fall2002/entries/color/>

Merleau-Ponty, Maurice. *The Primacy of Perception : and other essays on phenomenological psychology, the philosophy of art, history and politics* James M Edie (ed.), Northwestern University Press, Evanston, Ill, 1964

Mollon, John. "Colour in Nature," in Trevor Lamb and Janine Bourriau (eds.) *Colour : art & science* Cambridge University Press, Cambridge, 1995, pp. 127-150

Nassau, Kurt (ed.) *Color for Science, Art and Technology* Elsevier Science, Amsterdam, 1998

Nassau, Kurt. *The Physics and Chemistry of Color: the fifteen causes of color* 2nd ed., Wiley Interscience, New York, 2001

Smart, J. J. C. "On Some Criticisms of a Physicalist Theory of Colours," in Alex Byrne and David R. Hilbert (eds.), *Readings on Color Vol 1: The Philosophy of Color* MIT Press, Cambridge, Massachusetts, 1997, pp. 1-10

Thompson, Evan. *Color Vision : a study in cognitive science and philosophy of perception* Routledge, London and New York, 1995

Tye, Michael. *Consciousness, Color and Content* MIT Press, Cambridge, Massachusetts, 2000

In Defence of Mysterianism

Melanie Rosen
PhD Candidate
Macquarie University

I: Introduction

Mysterianism is a position adopted¹ by Colin McGinn as an alternative to physicalism or dualism. This theory has provided a challenge to the physicalists who attempt to solve the mind body problem, and to the dualist explanation of the mind. According to McGinn, we cannot solve the mind body problem because we are cognitively closed to understanding the answer to this problem. Our ability to know this answer is limited by our methods of attaining knowledge, which consists of introspection and epistemic experiences. We can experience consciousness via introspection and we can experience the brain through science, but we lack the ability to understand how one consciousness arises from neurons. Of course, this has been the topic of much contention and my aim is to provide possible counter arguments to some of the criticisms of McGinn's theory. I will look at two criticisms of mysterianism, firstly, whether or not we can be cognitively closed in the way McGinn describes; I will argue that such cognitive closure is not incoherent as Uriah Kriegel argues. Secondly, I will discuss the issue of whether mysterianism can give support to physicalism, despite the fact that we are cognitively closed to understanding *how* physicalism could be true. This is a difficult claim to support, and is well criticised by Barbara Montero, but I will show how mysterianism may lend more support to physicalism than to dualism. In the final section, I will try to show that mysterianism provides support to physicalist naturalism.

II: What is Mysterianism?

The reason the problem of consciousness is so difficult to solve, according to McGinn, is that we are cognitively closed to understanding the relationship between the mind and the brain. We have two methods of inquiry which we can use to investigate mental states and brain states. Firstly, we have direct access to at least some aspects of our phenomenal consciousness through introspection.² If I ask myself, "How am I feeling," I can very easily realise that I have a slight headache, or that I am excited or thirsty. I know exactly what it feels like to be in whatever state that I'm in simply by thinking about it. Knowledge we have about the brain, on the other hand, is discovered in a quite different way. We cannot introspect into our brains the way we do our minds. We must use scientific research and technology to discover how a brain functions; neural imaging shows us images of the brain and we can even see the functional workings and neuron firing using modern technology. But without knowledge about correlations between behaviour and neuronal firings in the brain, looking at a pattern of neurons firing *alone* will not tell us whether *this person is in*

¹ McGinn, C. (1999) and (2004)

² Although there is debate that much of even our phenomenal consciousness is not accessible through introspection.

In Defence of Mysterianism

pain. We require the testimony of the individual to discover how they feel, and then we can use this information to discover which neurons fire or which functions occur when a person is in pain. For example, when I feel pain, my T cells may fire and when I look at disturbing pictures or beautiful pictures, different areas of the brain are used.³ Unfortunately, this type of information cannot tell us *which* brain functions are associated with which emotions, feelings and thoughts. The fact that certain brain functions do not entail certain types of qualia is problematic for the physicalist because without such entailment we leave open the possibility that there could be “zombies” who are functionally identical to us but do not experience qualia. An even more difficult issue is that it seems that no amount of research will tell us *why* qualia or experience occurs when the brain is stimulated. Why do humans experience such rich sensation, rather than just being complex automatons? We could talk about the functions of the brain at a neuronal level, or the functions of the mind at a phenomenological level, but as argued by David Chalmers, no amount of functional information can solve the problem of how phenomenal consciousness arises from neuronal activity, because functions do not have such explanatory power.⁴

Chalmers argues that no matter how extensive our research into the behaviour of neurons and functions of the brain, such functions can only causally explain other functions, and they seem unable to explain why we have conscious experience in the first place. Chalmers argues that because introspection and study of the brain are irreducible, we must conclude that dualism is correct. Chalmers does not believe that the truth of dualism entails that consciousness is beyond the grasp of science, just that science needs to focus on new types of answers. McGinn on the other hand doesn't think that dualism is correct, but rather that there is something about the mind that is beyond our ability to explain, and will be forever unexplainable to humans. The fact that introspection and brain science are irreducible does not show dualism to be true, but instead we lack an essential method of understanding what can link introspection with empirical study.

Introspection, according to McGinn, tells us what the contents of consciousness are. In this way, we are directly acquainted with our own consciousness, which McGinn refers to as an *acquaintance based* knowledge.⁵ This can give us no knowledge about neurons or brain processes. We are directly acquainted with our consciousness, where as we tend to only have knowledge by description of the brain. This is why it is hard for us to find a satisfactory reductionist account of consciousness, which, McGinn states ‘is possible only in the presence of a priori entailments from one set of facts to the other’⁶ as aforementioned. In other words, we can only reduce consciousness to the brain if the existence of the brain entails the existence of consciousness. Pretheoretically, this doesn't seem to be the case, as

³ Layard, (2005) p. 18

⁴ Chalmers, (1995)

⁵ This draws on the distinction between knowledge by acquaintance and knowledge by description discussed by Russell. (Fumerton, (2004)). We have knowledge by acquaintance of only those things we are directly acquainted with, e.g. I am acquainted with my parents, but I have only knowledge *by description* of McGinn, whom I have never met. I am directly acquainted with my experiences, but not with my own brain. I could become acquainted with my brain by using a saw and mirror, but I could not gain acquaintance to the extent that I have with my experiences, which I know far more intimately.

⁶ McGinn, (2004) p. 13

we can conclude from the zombie argument. This argument is about how it seems possible for a world to exist which is physically identical to ours but lacks consciousness; thus those who inhabit the world are *zombies*; creatures that are physically identical to us yet have no phenomenal experiences. Of course, most physicalists would deny this possibility. But the real problem, according to McGinn is that if the mind *is* reducible to the brain (and he believes this is true), we have no method of discovering how. But why should we believe physicalism is true, nonetheless?

Physicalism itself, according to McGinn, is an ambiguous philosophical position. He sets out two formulations of physicalism which show how broad the concept of physicalism can be. Narrow physicalism is the thesis that consciousness is reducible to the properties of neurophysiology and physics we have already discovered. This seems quite unlikely, as neuroscience is still making new discoveries about the way the brain works, and it seems like we are a long way off having a complete understanding of the functions of the brain. Wide physicalism, alternatively, states that consciousness is reducible to the properties explained in the perfect theory of the universe. Of course it is not clear that such a perfect scientific theory of everything necessarily would not include a non-physical aspect of the universe. So we make no headway in the debate between physicalist and dualist accounts. McGinn's naturalist position is on the scale between wide and narrow physicalism. His naturalist position is a physicalist one as he rejects dualism altogether. Any adequate theory of the mind, according to McGinn, 'must exclude dualism as an option: once we see how the brain generates consciousness we will have no inclination to think that mind and brain are separate entities.'⁷ This is because in discoveries such as that water is H₂O, once we fully understand the connection, reduction becomes obvious, and any magical type theories become implausible. Although McGinn believes such reduction is *ontologically* true, we will never be able to understand how or why this is the case because of our *epistemic* gap. But how do we know there is no ontological gap? We need to show that it is at least more likely that there is a real epistemic or explanatory gap, but no ontological gap in order to agree with McGinn's claim that physicalism is true despite our cognitive closure. If there was an ontological gap between the mind and the brain, our inability to understand would not be from our own cognitive closure, but because of the way the world is. Of course we can never know which is true, but scientific theory and Occam's razor seem to support the epistemic gap. It is simpler to postulate a lack in our own ability than a fundamental gap in the ontology of the universe. This is also supported inductively; for example, the success science has had in explaining the mysteries of the universe and often disproving accounts of nature which suggest that nature is an ontological mystery or that it requires some non-physical higher power as an explanation. We tend to rule out ontological mysteries concerning unexplained phenomena as we learn more about the world.

McGinn argues that there are three responses people have to new and unexplained phenomena. They either try to domesticate the phenomena by reducing it into another theory, declare it irreducible, propose some magical explanation or try to eradicate it from their ontology.⁸ I will note four types of responses a one can take over the mind/body

⁷ McGinn, (2004) p. 15

⁸ *Ibid.*, p. 183

In Defence of Mysterianism

problem. One could try to reduce consciousness into theories about the brain, claim that there is a non-natural spiritual element of which consciousness is composed, claim there is no such thing as consciousness, or claim that the mind is not reducible to the brain, but physicalism is still true. McGinn instead opts for a method in which mysterianism explains away the mysterious aspects of naturalism; the mind only appears to be mysterious to us because of cognitive closure, not because of an ontological gap between mind and brain. Mysterianism does not attempt to deny the existence of consciousness as some theories do, but instead explains why naturalism is hard for us to understand whilst supporting the acceptance of naturalism. I will now discuss some counter arguments to McGinn's claims.

III: Can we be Cognitively Closed in the way McGinn Proposes?

Many of the unsolved mysteries of nature involve a lack of empirical evidence to support one theory over the other, which prevents us from understanding a phenomenon. One such example of an unsolved mystery is that we don't know how the dinosaurs died due our inability to gain epistemic knowledge of events that occurred that far in the past. There are a variety of hypotheses, such as a meteor, climate change or atmospheric change, but it is unclear which of these options, or which variety is correct given our lack of evidence. This of course, is a different type of epistemic disconnection than the one McGinn discusses. Our lack of knowledge of the dinosaurs comes from the difficulty of knowing what occurred in the distant past. McGinn argues that the mind has a more fundamental difficulty; we do not even *know* of any explanations that are adequate to solve the problem. McGinn says: 'we do not have a plethora of theoretically adequate options from among which we cannot empirically choose; rather, nothing that we can think of has a chance of explaining what needs to be explained.'⁹ Uriah Kriegel on the other hand, argues that such cognitive closure is incoherent. In this section, I will provide some counter arguments to Kriegel's thesis.

Kriegel characterises McGinn's thesis as *the new mysterianism*. He argues that McGinn's mysterianism is incoherent because we cannot truly *understand* a problem without understanding the types of possible solutions to that problem. Either we must be able to comprehend which *types* of answers would solve the problem of consciousness, or we do not really understand the problem. Kriegel believes that McGinn's mistake arises from failing to distinguish between cognitive *closure* which is 'principled and permanent' and cognitive *closedness* which is 'unprincipled and provisional in character.'¹⁰ Cognitive closedness is a temporary or contingent state of ignorance. *We do not* know the answer, but it does not mean we *could not* know. McGinn believes that we suffer from cognitive closure, which is *not contingent* because the way our brains function prevents us from ever knowing an answer. According to McGinn, no human with cognitive capacity like ours could ever solve the mind body problem. This is because of the separate ways we have knowledge of our own minds and brains. Knowledge of my own mind comes from introspection, whereas knowledge of the brain comes from aforementioned scientific methods. The only explanation for this type of cognitive closure to be possible, Kriegel maintains, could be that we do not really understand the question. When someone asks a question such as *how much*

⁹ McGinn, (2004) p. 62

¹⁰ Kriegel, (2003) p. 5

does a man weigh, for us to understand the question we must also understand what would suffice as an answer. One does not understand what it means to have a weight if one doesn't understand some form of system of weight measurement, for example, John weighs 80 kg. In the same way, one does not understand the question *are you happy* without understanding that the possible answers are *yes I am Happy* and *no I am not happy*. Or in other words, '...there is a conceptual connection between understanding a problem and understanding its possible solution(s)...what makes a question the question it is, with the specific meaning it has, is that all and only such-and-such assertions would constitute an answer to it.'¹¹ To not understand the possible answers is to not understand the question. Kriegel thinks we should reject the theory of cognitive closure, instead opting for a weaker form of ignorance termed *cognitive closedness*; although we may never know the solution, we are not ignorant in the radical way McGinn describes. Our ignorance is only contingent. For example, before the discovery that water is made up of H₂O molecules, we lacked an empirical discovery which would have provided the solution. So mysteriousness may be caused by the lack of an empirical discovery, rather than some radical incompleteness of our view of reality.

I disagree that the concept of cognitive closure is incoherent. Kriegel's examples are not wide ranging enough to be convincing. Firstly, the example given by Kriegel about asking how much a man weighs may be understood in different ways. What about someone who uses a simpler method of weighing things? For example, I may not understand any measurement system, but answer the question by saying that John is too heavy to carry or light enough to carry; these might be the only answers I am capable of understanding. If I was then asked what John weighs in kilograms I would not understand the question or how to answer it, but this does not mean I did not understand the initial question. Although this type of ignorance is not McGinn's cognitive closure, this shows that understanding questions is not quite as simple as Kriegel assumes. Secondly I think Kriegel's analysis does not take into account all types of questions. In the weight example, there is only a limited range of weights John can possibly be. I would argue that this is only true for some simple types of questions. It is true that to understand the question "does Mary love John?" one would need to understand what it is to both love someone and not love someone, not all questions can be answered with a yes or no. For the *how* and *why* questions, answers can be far more complex and indeterminate. For example, perhaps I discover one day that my best friend has secretly been torturing stray cats in her basement. I could ask her "*why* would you do that?" without knowing all the possible answers to the question. I may not understand what schizophrenia is, for example. I may not know that some people lack a vital function in the brain that allows for empathy. This would mean that my friend is incapable of feeling sorry for a cat, but is just interested in the sounds it makes whilst being tortured. Even if I am told this information, I may still not understand how someone could possibly be this way. Alternatively, questions like "*why* do I always dream of electric sheep?" could be explained by some slight brain malfunction which, even if I knew about it, I may not understand the connection between having certain brain functions and dreams. I think *did* and *would* type questions which tend to have a yes or no answer fit into Kriegel's examination, whereas *how* and *why* questions could have multiple possibilities, the answer

¹¹ *Ibid.*, p. 9

In Defence of Mysterianism

to which may be sometimes beyond our comprehension. "What started the big bang?" is a question that I believe I understand, but do not know all the possible answers to. There may be some type of cause I am unaware of, or that transcends by ability to understand. To take an example from the realm of human experience, consider *what it is like to see the colour ultra violet*. I'm sure you understand the question, but could you list all *or any* possible answers? This is a more accurate portrayal of the question "*how* does consciousness emerge from the brain?" Even if the answer is within our ability to comprehend it, there may be such a wide range of possible answers that we may *never* be able to imagine all of them.

Kriegel concludes that for us to understand a question, we must be able to comprehend what would constitute an answer to that question. McGinn argues that we *can* comprehend, through intuition, that there *is* a problem in understanding how we can reduce experience to brain functions. We have two methods of experiencing the mind, through introspection and through study of the brain, and we know that they are connected somehow, but we do not know how they could be connected. If it is true that we understand the problem then according to Kriegel we should be able to comprehend what would serve as an answer to the problem of consciousness, but this does not seem to be the case.

IV: Can Mysterianism be an Argument for Physicalism?

Now I shall discuss the possibility of mysterianism being an argument for physicalism. Megill argues that Locke's mysterianism is more plausible than McGinn's because Locke remains agnostic on the physicalism or dualism question. I shall argue why this view is unsatisfactory. I shall then look at Montero's rejection of mysterianism as an argument for physicalism, and provide counter arguments to her claims.

1: Locke vs. McGinn

Jason Megill argues that although Locke and McGinn both share a mysterianist attitude towards the mind body problem, Locke's mysterianism is more plausible because it remains agnostic on the physicalism vs. dualism debate. Megill distinguishes between two types of mysterianism: strong and weak. Strong mysterianism, unlike McGinn, claims that both the answer to the mind body problem and whether or not physicalism is true are *both* mysterious, and hence we should not embrace physicalism. Weak mysterianism, as advocated by McGinn, embraces physicalism although we can never prove it to be true or understand it. Megill notes that 'new mysterianism assumes the physicalist frame work ... [that] consciousness is a physical entity,¹² but he finds strong mysterianism more plausible. Mysterianism is insufficient to provide support for physicalism, according to Megill, because if we can never explain *how* physicalism can be true, we will never know whether or not it is true. I agree that we cannot show that physicalism is true using the mysterianism argument alone, but unlike Megill, I think mysterianism could at least explain *why if physicalism is true* we can't comprehend exactly how the mind supervenes on the brain, and why we have experience at all, which is the *hard* problem of consciousness.

¹² Megill, (2006) p. 121 ff.

Megill claims that Locke is a strong mysterianist. In Locke's writings, he argues that we do not know the solution to the mind body problem, and may never know. I would argue that although Megill's claim that a mysterianism that remains agnostic is stronger than one that supports physicalism seems to be true, but only vacuously so. This is because it is *always* more likely that *a or b* is true rather than *a* is true, irrelevant of what we know about either *a* or *b*. Adding a disjunct to a simple affirmation will *never* decrease the probability of that statement; it will in fact increase its probability, unless the probability of the disjunct is zero. I agree that saying that "either physicalism is true or dualism is true" is more likely to be true than "physicalism is true" but that *does not mean* that physicalism and dualism are *equally probable*. I am unsatisfied with leaving the issue open to the extent that Locke does because if physicalism is more probable than dualism, we should *accept* physicalism. My position is that there are good reasons to think that mysterianism can give support a physicalistic theory of the brain, which I will discuss in reference to Barbara Montero's arguments. Therefore, although I believe Locke's mysterianism to be a strong position, I will attempt to show that McGinn's more tendentious position nonetheless has certain convincing aspects, which I will discuss.

2: Montero and Pessimistic Physicalism

Barbara Montero argues that mysterianism fails to support physicalism due to a logical gap between the statement *we are cognitively closed to the answer of the question of how the brain causes the mind* and the conclusion *we should accept physicalism*. She defines McGinn as a *pessimistic physicalist* who believes that there is no ontological gap between the mind and the brain, but there is an explanatory gap that prevents us from knowing how the two are related. Montero argues mysterianism provides no support for physicalism in the way McGinn suggest, but I will provide a counter argument to this claim.

According to ontological physicalism, mentality is not a fundamental feature of the world. It is reducible to physical features such as the brain. One may be an ontological physicalist whilst being an epistemic dualist, or in other words one can believe that the concepts of mind and brain are irreducible even though ontologically the mind *is* reducible to the brain. Some people believe that the physical science we currently have does not provide an explanation for how the brain produces consciousness *now* but this does not entail that we will never solve the problem. McGinn believes that we require a third method of experience, the ability to look at our minds and brains outside of our current limited abilities to be able to understand how the brain produces consciousness. Of course we will never be able to do this because no one can experience from outside themselves. Consciousness is just a physical process and no explanation of it would be intelligible to us. Our brains would need to be different than they are to understand consciousness fully.

The hardest problem to solve according to Montero is the *generation* problem; how does consciousness and experience come from the firing of neurons? This generation problem is the main issue that divides people. Different people are convinced by different arguments because 'if you find it puzzling how a physical brain could generate consciousness, zombies will seem possible, as will inverted spectra, as will Mary's lamentable condition.... if you are an optimistic physicalist and do not find the generation of consciousness from insentient

In Defence of Mysterianism

matter puzzling, the thought experiments that are supposed to elicit anti-physicalistic intuitions will not have their desired effect.¹³

So our decisions over which is the more plausible theory may be based on intuitions that are not well founded. If opinions among intellectuals vary broadly as to which are convincing arguments, we need to show why there is so little agreement, and attempt to find arguments that are more widely convincing.

According to the pessimistic physicalist, the generation problem lies on the “epistemic side of the epistemic/ontic divide,”¹⁴ but Montero states that it is unclear that this is the case. An ontological gap between brain and mind would lead to the same epistemic gap that McGinn describes, so we cannot know whether the gap is ontological or epistemic. Nonetheless, even if there is only an epistemic gap, our problem is that we are in an epistemic situation in relation to mental and physical properties that makes the problem unsolvable. Saying the generation problem is epistemic is not the same as saying physicalism is true, or that there is no metaphysical problem. This only leads us to think that whatever the case may be, we will not be able to know the answer.

Montero doesn't think we can leap from the acceptance of mysterianism to the acceptance of physicalism, because we should never make such assumptions without further reasons. This next example seems, *prima facie*, to be a case where one should accept something without further information. If a person has a dream about talking to their dead father, in most cases upon awakening they will believe the dream was in fact a dream and not reality *without* being given any further information; it seems more likely that such things, such as talking with the dead, only occur in dreams. But Montero claims that for this to be a reasonable assumption, he *would* need a good reason to believe that it wasn't real, such as prior knowledge about certain facts about the world. Our world is such that we commonly believe that ghosts are not real. The man, given his cultural background and assumptions, has good reason to accept that it was a dream. If we are told that we cannot solve the mind body problem, we would need a good, separate reason to accept physicalism, and physicalists *don't have this reason* according to Montero. But to accept Montero's argument we must assume that there is no reason to accept physicalism. There are many good arguments for both physicalism *and* dualism, which makes the problem so difficult to solve, but I think that mysterianism reduces the intuitive force of the generation problem, which tips the scales in favour of physicalism. After explaining the generation problem, mysterianism allows us to appeal to physicalism as the best explanation. I will discuss how mysterianism explains away the generation problem in more detail later. First, I will look at some arguments for physicalism.

Some may argue that the causal closure of the physical, which is widely accepted by scientists, gives strength to the claims of physicalism. But firstly, some theorists, for example, David Chalmers, reject causal closure of the physical, and secondly, causal closure of the physical leaves open many possible alternatives to physicalism. For example, if

¹³ Montero, (2003) p. 407 ff.

¹⁴ *Ibid.*

epiphenomenalism, the theory that the mental only effects the mental and not the physical, were true, the world would still abide by causal closure of the physical. If the mental has no effect on the physical, then it would be undetectable and the world would appear to be entirely physical. Alternatively, given evidence that not all physical events have physical causes, we could reject causal closure of the physical, and accept interactionism; that the physical and non physical interact. One would require very strong evidence to reject causal closure because if it were the case that causal closure of the physical were false, the perfect scientific system would be impossible unless a scientific system could be made to include that which is outside the scope of modern science; the non-physical. But McGinn accepts that we will never have a perfect system anyway due to our epistemic gap, so why not remain agnostic on whether there is an ontological gap as well? Montero claims that we can only argue for causal closure of the physical if we take the view that everything should be accountable for by science, and dualists would reject this view anyway. I don't agree that *all* dualists would reject this view. One could be a dualist and think that in the future science could take into account the non physical. Secondly, the only reason dualists have for rejecting the causal closure thesis is the mind problem. It seems that the mind is the only aspect in the world that would have a non-physical aspect if dualism were true. Montero is suggesting we can use the absence of causal closure to support dualism, but if the mind is the only reason to reject causal closure, this seems to be to be circular reasoning. Now I shall show that Montero misinterprets McGinn's views.

Montero seems to interpret McGinn as having arrived at physicalism by way of mysterianism, and since mysterianism is insufficient in itself to prove physicalism, we should reject McGinn's claims. But this is an incorrect interpretation. Firstly, McGinn states that his mysterianism has developed from thinking naturalistically of the mind, not the other way around.

'So if I am a mysterian about the mind-brain link, it is because I am a naturalist about the human mind; indeed, since consciousness appears essential to scientific and philosophical understanding, I am a mysterian about consciousness precisely because I am a naturalist about it.'¹⁵

Montero claims that if pessimists were convinced of the arguments for physicalism, they would be physicalists, but I do not think this is the case. McGinn is convinced by some arguments for physicalism but there is one obstacle that is unsolvable: the *hard problem*. This is the argument against physicalism which Barbara Montero refers to as the *generation problem* mentioned earlier; how do firing neurons cause experience? No arguments about causal closure, evolutionary theory or the success of modern science succeed in overcoming this problem. But mysterianism goes at least some way towards describing *why* we find this such a problem, and why this problem should not lead us to dualism. We can't solve the generation problem because of our *incapacity to understand it*. This does not prove the existence of physicalism, but appeases to some extent the main reason to reject physicalism. With the support of other arguments, such as the causal closure argument and

¹⁵ McGinn, (2004) p. 70 ff.

In Defence of Mysterianism

the problem our conception of evolution poses for dualism,¹⁶ we have reason to accept physicalism over dualism as the more plausible theory. Although some interactionists have a causal account of the mind/body relation, there are many problems with interactionism. Interactionism isn't supported by modern science, as we have no way of determining *how* the physical can interact with the non-physical. We can only know about *physical* interactions.

McGinn argues that the human brain has been part of the evolutionary process along with all other living creatures, and hence is prone to cognitive inabilities, such as the inability to remember more than eight digits at a time. In the same way a cat cannot understand ethics. Mysterianism is intended to demystify what Montero refers to as the generation problem. She argues that we have no reason, given mysterianism, to accept either way, because although there are many convincing reasons not to believe dualism, there are equally some *prima facie* reasons to reject physicalism. The generation problem is the hardest problem to solve. But McGinn demystifies this problem by explaining that we are cognitively closed to understanding the answer. If physicalism were true, because of our cognitive makeup, it would still not be understandable, thus we find not a solution to the generation problem, but an explanation as to why we find it problematic without entailing dualism. I think in this way one of the main arguments against physicalism is weakened. Hence, although mysterianism in itself is not a definitive argument for physicalism, it weakens the intuitive force of the generation problem, hence we have one less reason to reject physicalism. In this way mysterianism supports McGinn's naturalistic view of the world.

V: Conclusion

Although I agree that we cannot prove that there is no ontological gap in nature, I have shown that the possibility that we are cognitively closed to the answer to the problem of consciousness is not incoherent in the way Kriegel suggests. I believe mysterianism goes some way towards explaining why there may be an epistemic gap that prevents us from understanding how physicalism could be true. This provides support for physicalism by weakening the intuitive strength of the argument that *our inability to reduce the mind to the brain means that they are irreducible* by giving us an alternative: namely there is an epistemic irreducibility but not an ontological one. Although I agree with McGinn that it is possible that we are causally closed in this way, I myself remain agnostic as to whether this is a fundamental or temporal ignorance. I shall still remain hopeful that there is further work to be done by philosophers, and that such work may give us insight into what consciousness is.

¹⁶ If the only non-material substance is the mind, this must somehow have evolved along with other traits, but it seems highly counterintuitive that the non-physical came into being at some time along the evolutionary track. It seems equally problematic that the non-physical mind appears at some point in the development of a foetus or baby. Aspects such as awareness and understanding can conceivably slowly develop over time, not requiring that a being be at any time either fully aware or entirely unaware, but a non-physical attribute it must either be present or absent, not somewhere in between. It would be hard to say *when* and *how* a being gains this attribute.

Melanie Rosen

Bibliography

Robinson, H. (2007) "Dualism," in the Stanford Encyclopedia of Philosophy
<http://plato.stanford.edu/entries/dualism/#Int>

Fumerton, R. (2004) "Knowledge by Acquaintance vs. Description," in the Stanford Encyclopedia of Philosophy
<http://plato.stanford.edu/entries/knowledge-acquaintdescrip/>

McGinn, C. (1999) *The Mysterious Flame* New York: Basic Books.

McGinn, C. (2004) *Consciousness and its Objects* Oxford: Oxford University Press.

Chalmers, D. (1995) "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies* 2 (3) pp. 200-219

Braddon-Mitchell, D and Jackson, F. (1996) *Philosophy of Mind and Cognition* Oxford: Blackwell Publishers Ltd.

Megill, J. (2006) "Locke's Mysterianism: On the Unsolvability of the Mind-Body Problem," *Locke Studies* 2006 vol 6. pp. 119-147

Kriegel, U. (2003) "The New Mysterianism and the Thesis of Cognitive Closure," *Acta Analytica* 18 (2003) pp. 177-191

Montero, B. (2003) "The Epistemic/Ontic Divide," *Philosophy and Phenomenological Research* 66 (2) pp. 404-418.

Layard, R. (2005) *Happiness: Lessons from a New Science* New York: Penguin Books.

CALL FOR PAPERS

COGITO is inviting submissions for the 2009 edition to come out in early 2010. Anyone interested should deliver their submissions to the Socratic Society, the Russellian Society or the Platonic Society by email:

thesocraticsociety@gmail.com

cogito@russoc.com

platonicsociety@yahoo.com.au

...by post:

The Socratic Society
C/O School of History and
Philosophy
Level 3 - Morven Brown Building
Faculty of Arts & Social Sciences
University of New South Wales, NSW
2052

The Russellian Society
C/O SOPHI,
Administration
A14, The University of
Sydney, NSW 2006

The Platonic Society
C/O Department of Philosophy
Macquarie University, NSW
2109

...or by hand to the administration offices of the respective philosophy departments, directions to which can be found on their websites.

SUBMISSIONS DUE: JULY 31ST, 2009

Submissions to the Journal should meet these basic conditions:

1. Be an original work by a student at any level of university study
2. On a philosophical topic or of philosophical interest
3. Between 2000-6500 words in length
4. Submitted typed on paper or preferably as an email file attachment
5. Having good spelling and punctuation
6. And adhere to the style guidelines which can be found at

http://au.geocities.com/platonicsociety/Style_Guide_For_Submissions.pdf

The editors will also reply to a suggested topic or a summary of an article not yet written.

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

Swantje Lorrimer
BA Candidate
Macquarie University

I: Introduction

Creativity is a broad concept, but in the context of this paper, it refers to mental processes which bring forth original ideas and concepts. These concepts and ideas are such that they can potentially inspire other people, “strike a chord” so to speak, and are fit to be used as paradigms and reference points for cultural developments. Creativity is an antecedent stage of innovation - the application of an idea to a specific use and purpose. In this paper, I am going to discuss the “innovation” of natural language as an ongoing development and a process that involves creativity on behalf of all its users: speakers as well as listeners. This creative and therefore dynamic element of communication has been acknowledged in the literature in regard to the context-sensitivity of indexicals and demonstratives and it is there that I want to bring this notion of creativity to bear.

In this paper I will demonstrate that the exploration of indexicals has illustrated that an account of language needs to be able to accommodate a notion of creativity, in the sense of a practical, innovative application of the standard linguistic conventions; not only in an account of the origins of language, but also in order to explain indexical reflexivity in everyday communication. My aim in this paper is to show that Frege’s mediated theory of reference, as well as modern direct reference theories like the perspectival accounts of Kaplan and Hintikka, can incorporate the necessary creative component that expand on the notion of semantic meaning, provided they endorse an Oxfordian interpretation of Frege’s notion of sense and not a purely descriptive one.¹ I will argue that Kaplan’s traditional theory of direct reference can determine reference, but that the semantic meaning yielded is in fact a very narrow and restrictive concept. I will conclude that in order for an account of language to accommodate the notion of creativity, it needs to contain a certain something, which I believe the essence of Fregean sense captures, which Kaplan neglects altogether, but which modern perspectivalists retain, though they choose to rename it.

II: Creativity and Language

The interpretation as well as the use of language can be ambiguous. This is a well-documented problem for indexicals because their meaning is context-sensitive. David Kaplan (1933-) argues that pure indexicals refer automatically, but only relative to the context they are used in. In contrast to pure indexicals, true demonstratives require

¹ Throughout this paper I will refer to the Oxfordian interpretation of Fregean sense as the one employed by Evans and Dummett, which does not contain a definite description.

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

supplementation of sorts in order to determine a reference successfully: they require either accompanying ostensions or intentions.² I will illustrate my claim that creativity actually and necessarily features in our everyday use of language with three examples. The first will look at the ambiguity in the use of a proper name from the auditor's point of view; the second will look at the shift of meaning of the verb "to grab" from the producer's point of view; the third example illustrates the creative use of indexicals.³ I intend to illustrate with these examples that the Causal Theory of Names held by the direct referentialists can in most cases explain how reference is established by an expression alone, be it a proper name or an indexical, but that it fails in some cases because extra-linguistic elements and semantic features in a wider sense can introduce elements that render the determination of reference ambiguous. I claim that a creative approach is needed in order to reliably establish the reference of an expression in context and, most importantly, that the meaning of the reference in a fuller sense can only be asserted if a creative approach is incorporated into the determination of reference. This creative approach, I believe, defies definition in terms of rule-bound terms, but leaves open a gap that can only be closed with the approach of an open and uniquely human creative attitude.

The first example, based on Evans and Kripke examines the use of creativity in regard to proper names. This is an important misconception to highlight because the meanings of proper names are commonly believed to be fixed and therefore their meaning is thought to be unambiguous. I will consider Evans' example of a TV host asking to name a capital city (Evans 1973: 194) and expand on it. The contestant replies: "Kingston is the capital of Jamaica!" This is a true statement, even though the contestant might have picked up this correct information by way of hearing "Kingston" in an incorrect use; i.e. from a person of racist persuasion travelling on the bus through Kingston-upon-Thames (Great Britain), pointing to a group of Jamaican youths and muttering to his companion: "Kingston is the capital of Jamaica." The two points that Evans makes are, firstly, that the causal origin of the proper name in this case has no bearing on its use in the contestant's correct utterance: the origin of the contestant's knowledge of "Kingston" was not referring to the geographically correct location of that city. This means, no causal connection between one person's use of the name and another's is needed in order to use the name to say something (Evans 1973: 195). A second point that Evans makes, not only in regard to the above example but in relation to Kripke's theory of names generally, is that in the Causal Chain theory, the causal relation has been misplaced. The relation does not lie between an 'item's being dubbed with a name and the speaker's contemporary use of it' (Evans 1973: 197), but in the designated entities' 'states and doings and the speaker's body of information' (*Ibid.*)

The second example examines the change in the use and meaning of the verb "to grab." Twenty five years ago, "to grab" was used in terms of "to seize eagerly and snatch," "to take illegal possession of" or "seizure or acquisition by violent or unscrupulous means" (Macquarie Dictionary, 1982). In our contemporary use it is used in the sense of "to pick up"

² David Kaplan claims in his early work that ostentation is needed and in his later work he claims that intentions are sufficient for demonstratives to successfully refer.

³ I am very grateful to Mrs Ann Lorrimer and her extensive knowledge of the English language. The example of "to grab" resulted from one of our discussions about the mysterious ways of the indexical.

or “to take,” as in “Let me grab your number,” “Please, grab a seat,” or “I’ll grab my keys.” The change in meaning provides the evidence that, at some point, someone used “to grab” against convention. This could have occurred for instance when taking the phone number of someone he or she fancied. The use of “to grab” in this context added a level of meaning beyond taking someone’s phone number; it also expressed desire creatively, in a very personal way. This creative twist in the expression was probably seen very fitting by someone else and re-used until it became an institutionalized convention.

Whereas the first example illustrates the interpretative identification of an expression carried out by a listener, the second introduces the notion of executive identification in which the speaker has room for a creative manoeuvre in choosing his words, rather than relying on a previously uttered expression for the interpretative identification that a listener must utilize (Kapitan 2001: 296). The third example emphasizes the notion of executive interpretation even more clearly, referring to the creativity involved in expressing the cognitive perspective of the speaker in the use of an indexical.⁴ Kapitan mentions the example of self-criticism sometimes being expressed in the second person: “You are a fool!,” rather than “I am a fool!” (*Ibid.*) Although Kapitan asserts that it is not necessarily evident in interpretation which indexical type a speaker employs, I believe that in certain cases, like in the above self-criticism, it is. Both expressions reveal a cognitive perspective that speakers take in relation to their deeds: the former appears to express speakers distancing themselves from their deeds or stupidity, the latter seems to express a more accepting attitude, in terms of owning up to their mistakes. In sum, executive interpretation makes room for the creative use of convention. What is conveyed in the above example goes beyond what is literally said in terms of linguistic meaning, which Davidson terms *first meaning* or the level of meaning initially yielded from an expression, similar to the meaning found by consulting a dictionary (Davidson 1986: 435).⁵

In this first section I have illustrated two things. Firstly, there is a level of meaning beyond reference, in terms of the bare semantic or literal notion of what was said. This meaning relies on semantic, as well as on extra-semantic features, like for instance context, on the historic meaning of a word and, to a certain extent, on the shared knowledge of speaker and interpreter in order for this level of meaning to be conveyed. Secondly, I have shown that thoughts can be expressed in different ways, introducing the notion of the speaker’s cognitive perspective. In the next section, I will explore Frege’s account of semantic meaning in his theory of mediated reference. I believe that Frege’s account can explain the integral feature of language that I have identified as creativity, given an adequate interpretation of his notion of sense.

⁴ Needless to say that executive interpretation can occur with proper names, as in many of the Fregean puzzles. Hesperus is Phosphorus is probably the most well known of them which intend to prove the differing cognitive value of co-referring names. But indexicals can be examined in regard to interpretative identification. I am putting executive interpretation and indexicals together in one example to cover more ground in shorter space in order to bring out the wide ranging scope of creativity.

⁵ The three conditions of *first meaning* are defined by Davidson as (1) first meaning as systematic, (2) first meaning as shared and (3) first meanings being governed by learned conventions or regularities (Davidson 1986: 436)

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

III: Mediated Reference Theory

Contra Russell (1872- 1970), who claimed that singular propositions refer directly to an individual or object in virtue of it being a constituent of that proposition, Frege (1848 – 1925) asserted that propositions are composed of senses, not of the objects themselves. On this view, senses, which are independent of the entities they refer to, mediate between expressions and their reference. This approach to propositions is expressed in the Fregean notion of “thought.”

“On Sense and Reference” (1948), the article in which Frege initially explicates his theory of meaning, opens with an exposition of his core concept, cognitive value, or cognitive significance. He explains that co-referring proper names do have different senses in presenting a person or entity in different ways; consequently they are apprehended differently and often constitute ‘valuable extensions of our knowledge and [statements of the form $a = b$] cannot always be established a priori’ (Frege 1948: 209). This illustrates Frege’s project of investigating language: he is interested in sentences, as opposed to singular words or expressions because only in the context of a sentence does a word have meaning (Frege 1953: 184). Moreover, as a logician, he is concerned with *how* the truth-value of sentences is determined (Dummett 1981: 82). Frege’s theory of language thus approaches semantic meaning in a wider sense because the meaning of an expression contains two equally important semantic values: sense and reference. This account of meaning relies on more than the interpreter’s association of a word with an object in the world; it relies on its mode of presentation of the referent as well as *how* the association of the referent to its meaning is brought about (Dummett 1981: 93). Therefore a Millian conception of proper names is incompatible with Frege’s account.⁶

The sense of a proper name *can* be expressed in terms of a description (Dummett 1981: 98), and, indeed, most contemporary philosophers of language treat senses in this very way. Perry’s interpretation (1977) for instance is limited to this one aspect of sense. However, its function is more far-reaching in order to achieve a fuller sense of meaning. In this way sense is required in order to determine the truth-value of a sentence, as well as the referent or cognitive value of an expression. Moreover, sense constitutes the mode of presentation in which an object is made available to an interpreter, and also expresses a definite cognitive perspective. Sense is an integral semantic feature of expressions, a type of meaning, but it also contains some “extra-semantic” features in connecting reference to cognitive aspects of meaning. This relation between sense and knowledge also illustrates a tacit but strong connection between sense and knowledge in Frege’s account (Dummett 1981: 95). Frege presupposes a certain amount of competency from speakers in order to generate meaning, but he concedes that speakers never attain ideal and comprehensive knowledge (Frege 1948: 210-1). This has also given rise to a number of misunderstandings and attempts to discredit Frege’s notion of sense. Wettstein for instance argues against sense in regard to the cognitive significance assigned to co-referring terms. He does this by attempting to

⁶ “Millian name” refers to a conception of proper names which are conceived merely as labels for individuals and objects and contribute nothing apart from the individuals themselves to the meaning of the expression.

reduce Frege's view to one that requires the interpreter to have excessive knowledge in order to make sense of expressions (Wettstein 1989: 333). But Frege does not claim anywhere that the lack of certain information prevents the agent from identifying the referent (Dummett 1981: 99).

The connection between sense and knowledge is the result of an overarching intention behind Frege's project: he was not interested in semantics itself, he was interested in investigating the possibility of objective thought contents. In other words, Frege was interested in exploring human thought and language served merely as a tool to gain access to the workings of the mind, not in terms of psychology and "thinking" as such, but in the possibility of objective thought and the truth-conditions of sentences expressed through language. Truth-conditions and thoughts are linked in Frege's theory because thoughts are regarded as timeless (Frege 1956: 310), objective, shareable and independent of being thought by a thinker (Frege 1956: 311). A thought is the sense, not the reference, of a sentence (Frege 1948: 215). The referent (the object referred to) of a sentence is a truth-value (Frege 1948: 216). Only complete sentences express thoughts, sentences that are completed by a time-indication (Frege 1956: 309). A sentence uttered may contain a truth value (if it is true) because the objective notion of truth is attached to it. It constitutes a fact and thus a sense can determine a referent absolutely. Frege claimed that reference necessarily required a concept, a sense, and that conceptually unmediated reference was impossible and incoherent.

Frege's notion of sense extends beyond description: it also expresses a cognitive perspective on an object, without which the meaning of the referent cannot be determined. In order to highlight this integral feature, I shall rephrase: Frege investigates the identity of objects not as a relation between objects themselves, but as the relation between signs designating the objects and the objects themselves (Frege 1948: 209). A referent has many senses; they are signified by different signs and expressions. Each of these reveals a different aspect of that same object, containing different information about it. Sense reveals how the referent is thought of and thus enables different perspectives about one and the same object. In this way sense mediates between the sign or word and the object. The strength of Frege's account is that, though the referent cannot be determined without a sense, a sense can be expressed regardless of a corresponding referent. This means that senses can inform us about possible objects and fictional characters and open these up to discussion and contemplation.⁷ In contrast, Russellian singular terms presuppose the existence of the object in question. This adds a semantically insurmountable complication for the incorporation of, for instance, mythical figures or accounts of objects and events of which the interpreter has not sufficient information available.

With this account of mediated reference theory firmly in place and the central role of sense explicated, I will now explore this theory in relation to the use of indexicals and then illustrate how this account is able to accommodate the notion of creativity.

⁷ Evans (1985: 722) elaborates on the benefit Frege's account in regard to incorporate the sense of empty singular terms and sentences containing them. Sentences containing these do not express mere nonsense, but mock thoughts, as expressed in fiction.

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

Sense identified as description throws the incorporation of indexicals into question because indexicals are context-sensitive and their referent is not fixed without a specified context.⁸ Only through the conception of sense as a mode of presentation, as outlined above, does sense achieve a completing sense in indexical use, because the sense of indexicals can be envisaged as a kind of “keeping track of an object.” The interpreter, who is an embodied agent, situated in space and time, is able to grasp the same thought about a particular day expressed in a sentence containing “today” at different times. The completing sense for the indexical “today” is supplied by the agent being able to convert “today” into “yesterday” or “last week” and thus think the same thought again (Evans 1985: 730). This illustrates a dynamic way of thinking that is not only applicable to temporal indexicals, but also to spatial ones. Rather than designating geographical space, “here” belongs to a system of thought by way of which the subject locates itself in relation to its environment (Evans 1985: 153). “Here” expresses an egocentric mode of thought that enables the subject to locate itself in egocentric space in relation to the world.⁹ Rather than locating itself in terms of its own experience, the agent incorporates its egocentric space into the framework of an overall cognitive map, which is a precondition of experiencing egocentric space as *space* in the first place (Evans 1985: 163). This type of self-location is not compatible with a descriptive notion of sense, but it is with a dynamic notion by which the subject in question supplies the completing sense absolutely by virtue of being a spatial temporal being. Just as subjective experience is inseparable from an objective notion of self-location, it is important to note that Frege’s theory of meaning is not separable from an account of the actual practice of language. In other words, his theory of meaning and *how* language is applied are not distinct domains, although they do interact (Dummett 1982: 107).

Frege’s account of mediated reference allows the use of creativity in language because it enables the speaker to refer to the same objects using different senses, depending on what cognitive value a speaker wishes to convey. I can refer to the author of *The Adventures of Huckleberry Finn* as Mark Twain or as Samuel Langhorne Clemens. Both names refer to the same person but the meaning conveyed, attached to each name, is different. The former refers to the man as an author, the latter to a man who lost his father in his youth and pursued a career as a steamboat pilot before becoming a world famous author. Mediated reference theory allows “Kingston” to be used in different senses: in the conventional sense as the capital of the geographical location Jamaica, as well as in the sense of making a statement about the influx of Jamaicans in Great Britain’s Kingston-upon-Thames. On Frege’s account, the shift in the meaning of “to grab” is explicable because, as long as the essence of a word is preserved (in this example a notion of force or the urgency of exchange), it can be applied to a new context, thus constituting a new and creative way of expressing an action or feeling, which is interpreted adequately and institutionalized in the use of language. The example concerning indexical use is also compatible with the application of creativity to language. On this account, pure indexicals like “I” and “you” are

⁸ It is difficult to apply a purely descriptive notion of sense to indexicals, because the sense of “I”, “here” or “now” would indeed not be absolute or have a “completing sense,” which is a precondition for the expression of Fregean thoughts.

⁹ Note the parallels to Kant’s Transcendental Unity of Apperception which states that the continuous process of self-locating experience is the principle and precondition of self-consciousness (Kant’s *Critique of Pure Reason* §15 - § 27)

not determined merely according to a rule-bound procedure, but rather their determination takes into account the broader knowledge of the context, shared by speaker and interpreter. I can understand that "you" refers not to me, the interpreter, but to the speaker, because I witness that it was the speaker and not myself who has made a regrettable mistake.¹⁰

IV: Direct Reference Theory

Kaplan, a student of Rudolph Carnap (1930 – 71), one of the most influential figures in the movement of logical positivism, developed the traditional version of direct reference theory. It is important to understand that though both Frege and Kaplan define their projects as "semantic," their definition of "semantic" differs significantly. Kaplan's project is focused on the uncovering of the semantic rules that govern linguistic practices and it therefore excludes the cognitive elements that feature in the mediated reference theory and that are included in Frege's definition of semantics. Kaplan presupposes a more direct version of the language – reality (world) relation - than Frege's mediated account. Consequently, Kaplan claims to be able to determine reference of singular propositions and words in isolation to the expressions they occur in. This means that the context-sensitivity of indexicals refers to context only for the application of a semantic rule (character), but not for any other, i.e. extra-linguistic determination of meaning. This can be illustrated very well in regard to Kaplan's conception of proper names.

In regard to proper names, Kaplan returns to the Millian view of names, meaning that a name stands in for a certain object or person. On this view, proper names are merely labels for individuals and do not contribute anything else to the meaning of the name or sentence they occur in. This conception omits the particular mode of presentation of the entity as well as the particular way of apprehension on behalf of the interpreter. Kaplan's account is purely semantic because he regards meaning as being satisfied in purely semantic terms.

Kaplan disagrees with the full extent of functions attached to Frege's sense as outlined in the previous section. He assigns sense a purely qualitative and descriptive characterization and therefore claims that sense has no place in a semantic account of indexicals or names, as he conceives it. On the basis of this particular interpretation of sense as merely descriptive, direct referentialists claim that sense cannot account for the role that context plays in the use of indexicals, nor can it explain the *cognitive value* or "*fixed significance*" in non-eternal sentences (Ken Taylor 1998: 261). This argument is put forward in the influential paper "Frege on Demonstratives" by Perry (Perry 1977: 438). Perry defines "context" as 'a set of features of an actual utterance,..., including time, place, and speaker, but probably also more' (Perry 1977: 479) – in sum they are semantic features in the narrow sense, mechanical and rule-bound and thus unlikely to include cognitive features or perspective. The point of the argument is that the Fregean sense cannot determine, for

¹⁰ Frege (1956: 296-7) explains that thoughts can be shared if the relevant information surrounding an event or circumstance is equally available to speaker and interpreter. I claim that indexicals can be identified in the same way as thoughts can be shared.

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

instance, the indexical “today” absolutely, because “today” uttered on different days will yield different referents. Therefore, according to Perry, nothing in Frege’s philosophy can supply a completing sense, an unchanging truth-value attached to an utterance, which Frege himself calls for in order to be able to express thoughts (Frege 1956: 37). The rejoinder to this accusation I have outlined in the previous section, where I showed how Evans makes a case in favour of the incorporation of indexicals into mediated reference theory. Having outlined the concerns of the direct reference theorists, I will move on to explicate their actual account.

At the beginning of this section, I described how direct reference theories explain the reference of proper names. I shall now show how such theories explain the meanings of indexicals and indexical use. According to direct referential theory, the reference of a pure indexical is fixed in a very rule-bound and mechanical way: the meaning of words like “I”, “you”, “today”, “here” or “now” are determined in virtue of linguistic rules, called characters and content, the meaning fixed by character relative to context. Characters are formally described as functions from contexts to content. In an informal sense, characters can be described as linguistic rules that are associated by convention with indexical expressions and demonstratives. These rules, or characters, enable the determination of content in different contexts, for instance “I” always refers to the speaker of an utterance and “here” to the location of the context of utterance or production. I am not aware of Kaplan ever explicating the origins and development of these rules, nor can I imagine how a plausible developmental account might look like, given the mechanical and even clinical approach to language he purports.¹¹

Content designates what is traditionally referred to as the “proposition.” The content of an expression used in a particular context determines a corresponding set of qualities or properties designated by particular words or symbols, also referred to as intensions. When character is taken to a context, an intension is determined in the circumstance of evaluation, and the intension behaves like a function from possible worlds to extensions, and, in this way, the real and actual referent is determined. Character and content are thus equally essential for the determination of reference regarding indexicals. Once reference is fixed for an indexical, this reference stays fixed for all possible circumstances and all possible worlds – in a way once fixed, indexicals behave like the tags that Millian names supposedly attach to proper names.

Kaplan’s theory is directly referential because it allows the referents of objects to be determined directly by way of character without the mediation of sense and by making the objects referred to constituents of the propositions asserted. In other words, no mediation between word and world is necessary, because world or the objects constituting the referents are constituents of propositions. A circumstance denotes actual and counterfactual situations relative to which the content for the referent in question is determined or evaluated. Character can allegedly also bridge this very gap opened up between context of utterance and circumstance of evaluation. For demonstratives, this gap

¹¹ See also Hintikka (1998: 223) for Kaplan not ever giving a clear structural analysis of how precisely characters are determined in the case of demonstratives.

is larger than for pure indexicals because they need not only the determination according to different contexts but in addition they also need supplementation by something else in order to determine reference, though *what* exactly has been subject to debate.¹²

To summarize, Kaplan claims that indexicals can only be treated correctly on a directly referential model, because his theory encompasses merely the relation between world and language, omitting the relation between the mind or mental concepts and the world and the role which language plays in their mediation or the interpreter's perspective. Therefore it is important to note that meaning in his account refers strictly to linguistic meaning, but this is restricted to an application of semantic analysis. Though it is, like Frege's mediated reference theory, a theory of meaning, this notion of meaning is strictly limited to the literal "what-is-said."

Before moving on to an examination of the new direct referentialist account, I will apply my examples to Kaplan's direct reference theory, which highlight the creative elements in language: Kingston is only the capital of the geographical location Jamaica in Kaplan's reading, because any other application of "Kingston" and "capital of Jamaica" cannot yield a true statement. In the same vein, "to grab" will only have the one meaning, very close to the one occurring in a dictionary, because the use of "to grab" in a different sense cannot be accounted for in any other way when employing strict semantic rules. Kaplan's account cannot account for language change. Lastly and most clearly, the example of the use of "you" instead of "I" in a statement like "You are a fool" cannot be used retaining its truth-value on the Kaplanian direct reference theory. The semantic rules determine only "I" as the speaker and therefore "you" can never be used intending to determine the person uttering an expression.

V: New Direct Reference Theory

Both Hintikka and Kapitan are direct referentialists because they are concerned with how direct reference between an expression and an entity in the world is established without a mediating sense. They both share Kaplan's view of the objects of reference being directly constitutive of propositions. The new direct reference theorists differ strongly from the traditional account though, in that they emphasize a broader view of meaning rather than mere reference. They diverge from the Fregean picture in that they question meaning as being absolutely and objectively given. Both Hintikka and Kapitan concede that interpretation of meaning is a semantic issue, but that it is determined by perspective. Meaning in communication, and in particular in indexical use, is marked by an inherent meaning duality (Kapitan 2001: 297), that of speaker and interpreter, and this duality can be resolved by virtue of taking the perspective of the partner in communication into account. Moreover, these accounts expand on the notion of indexical use beyond that of

¹² Early Kaplan (Kaplan 1989a: 481- 564) insisted that demonstrations were needed, but in his later work (Kaplan 1989b: 565- 614) he claimed that intentions were sufficient to supply demonstratives with means to yield referents. Wettstein (1984) has argued that it is neither ostentation nor intention that completes demonstratives with extra- linguistic means in order to complete the determination of reference, but that social institutionalized clues fulfil these functions.

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

communication and the location of objects. On their account, the conception of indexical use enables self-location and self-awareness in relation to other external objects (Kapitan 2006). This ties in with Evans' emphasis on the importance of cognition and cognitive states in relation to the use of language as a way of gaining access to the world through language; though Hintikka and Kapitan claim that this is a direct relation, Evans regards it as mediated through something like sense. In what follows I will back up my claims with the relevant sections from Hintikka and Kapitan's accounts and then I will illustrate how the new referential theories allow for the accommodation of creativity in an account of language.

Hintikka differentiates between public principles used in identification and perspectival identification. The former relies on publicly shared information regarding an entity, the latter relies on the agent's first-hand cognitive relations to the entities in question as the identificatory framework (Hintikka 1998: 205). The two can come together, and indeed need to come together, for successful identification. This is similar to Evans' suggestion that egocentric space be incorporated into public space for successful demonstrative identification (Evans 1982: 151 ff., and Hintikka 1998: 206). Hintikka rejects the application of language universally to all possible worlds and outlines an account of "small world semantics" (Hintikka 1998: 218): semantics needs to be developed in a way which can be applied to a specific and actual context in isolation of other applications of the language in question without reference to other "small worlds" (*Ibid.*) The core point Hintikka intends to capture is that demonstratives receive their meaning through comparisons (Hintikka 1998: 221), but not to comparisons across all possible worlds, but across all possible alternatives applicable to the "small world" in question.¹³ To summarize, Hintikka claims that the application of language is always localized and particular to a certain actual situation, therefore it is the particular perspective of an agent that picks out an object demonstratively as the one to be determined, rather than the universally applicable force of character (Hintikka 1998: 223).

Like Hintikka, Kapitan emphasizes that language is not merely a means of communication, but one of self-location and thinking as well (Kapitan 2001: 293). Whereas the interpretation of an indexical expression is guided by the meaning of the indexical or demonstrative in context of the occurring utterance in order to determine the referent, the indexically mediated thought cannot be parasitic upon the speaker's antecedent production, but relies therefore on a different process, guided by different principles of identification.¹⁴ Kapitan

¹³ For instance the meaning function associated with "red" picks out the class of red objects from each situation or scenario, not only from each possible world. These applications are self-sufficient, not getting their meaning from the semantics of some wider global world by relativization or from comparisons with the semantics of other scenarios (small worlds) (Hintikka 1998: 218).

¹⁴ For a comprehensive list of differences regarding interpretative and executive interpretation please refer to Kapitan (Kapitan 2006: 11- 14): (1) interpretation is utterance- reflexive, execution is not, (2) tokens are causal inputs of interpretive identification, but outputs of executive identification, (3) interpretation is subsequent to utterance, execution is antecedent, (4) execution of indexical reference has room for a creative employment, interpretation lacks this element, (5) indexical interpretation is always non- autonomous, but executive interpretation can be autonomous if an indexical is the only means of asserting a state. This means that the indexical "now" might identify a time relatively to an unspecified time of "then", (6) both interpretations are guided by different principles: "today" necessarily entails different cognitive perspectives on the day in question for speaker and interpreter.

names this type of identification *executive identification* (Kapitan 2001: 294) which features elements of “creative manoeuvre” that interpretative identification lacks (Kapitan 2001: 296). It is worth highlighting Kapitan’s point here: the identificatory use of indexicals is affected by a meaning duality, that is, by executive and interpretative meaning. This duality is accompanied by a systematic coordination at the deepest level in communication (Kapitan 2001: 297) in which both parties are required to utilize both meanings in addition to the context-dependency which is typical for indexicals. Whereas Kapitan endorses a standard utterance-reflexive account for the interpretation of indexicals, such as Kaplan’s, he claims that executive identification cannot be explained on this account. In sum, without an account of executive interpretation an account of communication and language is not adequate.

Whereas interpretative identification is guided by the identification principle ‘One identifies an item only in terms of what is unique to it, that is, by means of one or more of its distinguishing features’ (Kapitan 2001: 298), it seems that executive interpretation does not need individuating modes, which are modes of presentation that identify referents, because the speaker can conceptualize indexical thoughts without specifying what he is thinking about (Kapitan 2001: 299). But in executive interpretation, one must still allow for an account of discriminatory awareness which provides the link between a thought and the object it concerns (Kapitan 2001: 300), and this awareness necessarily occurs prior to communication, intention and demonstration. Kapitan claims that the mechanism able to account for discriminative awareness is the position within one’s perspective (Kapitan 2001: 301), or, in other words, the self-location of the speaker within his specific perspective *in relation* to a particular context. Therefore it can be asserted that perspective itself is a content of awareness (Kapitan 2006: 32-3) or, in other words, that perspective is constitutive of awareness which factors in executive identification as part of communication. Executive meanings, the meaning creatively produced by the speaker, ‘function as forms or ways that items are apprehended and cognized’ (Kapitan 2006: 17). Kapitan’s description of “form” bears an uncanny similarity to Frege’s notion of sense, in that it allows for creative choice on behalf of the speaker to present the entity in question in a particular way.

Despite the direct-referential underpinnings of these accounts, it is interesting that their Fregean features enable them to accommodate creativity in a way that the more strict direct referentialism of Kaplan’s account cannot. Applying the three instances of creativity used throughout this paper to Hintikka’s account, the following assertions can be made. Firstly, a creative process on behalf of the agent in terms of bringing public and perspectival identification together is involved. Secondly, because Hintikka is concerned with localised and actual situations rather than with globally valid possible worlds, creativity can be employed in the identification of entities, as the application of my above examples illustrates. Kingston-upon-Thames can be named as the capital of Jamaica in a particular situation (here with racist intention) rather than only being used in making a global assertion of an all possible world scenario (which would look only at the actual geography of Jamaica). On a “small world” reading, the change of the meaning of “to grab” also makes sense: in particular situations “to grab” can be applied in the sense outlined in my above examples, because it creatively expresses in a particular situation the force of desire rather

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

than the global meaning of “seizure or acquisition by violent or unscrupulous means” and it is therefore easily explained how local or “small world” applications of a term can thus enter the global use of that term and thus contribute to the gradual change of its meaning. Finally, on Hintikka’s account, no character or semantic rule is needed in order for the interpreter to figure out the referent of “you” in the exclamation of “You are a fool!” – “you” in this particular situation is identified perspectively as the person who has made a silly mistake, and if it happened to be the speaker, as in the example, then he is determined as the referent as “you.” The creative use of “you” instead of “I” adds to the meaning of the expression, which in this case can be read as the cognitive perspective of the speaker, expressing reproach. To briefly summarize Kaplan’s view: the classic puzzles regarding identity of objects and individuals *are* clearly instances of our creativity and epitomize the choices we face in expressing the same referent in different ways, each time referring slightly differently, conveying different notions *about* the referent, thus attaching a different meaning to it.

VI: Conclusion

The purpose of this essay was to assert that the necessary element of creativity in the everyday use of language can be accommodated in Frege’s account of mediated reference as well as in that of the new direct referentialists. I illustrated this claim with the application of three examples. I outlined the strong parallels between Frege’s notion of sense and the perspectival accounts of language by the new direct referentialists. The most striking ones are, firstly, that expressions, and in particular indexicals, can have more than one meaning; secondly, that a cognitive perspective is assumed in the use of language which is constitutive of the meaning expressed; and thirdly, that language has more than a communicative function: it is fundamental for self-awareness and self-location. These functions of language need in fact to be presupposed in order for language to fulfil its communicative function. I concluded that Kaplan’s traditional theory of direct reference is unable to account for the use of creativity and that therefore, though able to determine reference in most cases, the direct reference approach does not achieve the same depth of meaning than the mediated reference theory or the new direct referentialists.

Bibliography

Braun, David (2007) “Indexicals,” published in the Stanford Encyclopaedia of Philosophy: <http://plato.stanford.edu/entries/indexicals/>

Davidson, Donald (1986) “A Nice Derangement of Epitaphs,” in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, (ed.) Ernest LePore, Blackwell (Cambridge)

Dummett, Michael (1981) “Some Theses of Frege’s on Sense and Reference,” in *Frege’s Philosophy of Language*, Duckworth Press (London)

Evans, Gareth and J. E. J. Altman (1973) “The Causal Theory of Names,” in *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 47, pp. 187 – 225

Swantje Lorrimer

Evans, Gareth (1982), "Demonstrative Identification," in *The Varieties of Reference* (ed. John McDowell), Oxford University Press, (Oxford)

Evans, Gareth (1985), "Understanding Demonstratives," in *Collected Papers*, Oxford University Press (Oxford)

Frege, Gottlob (1948), "Sense and Reference," in *The Philosophical Review*, Vol. 57, No.3, pp. 209- 230

Frege, Gottlob (1953), *The Foundations of Arithmetic*, (transl. J. L. Austin, Blackwell (Oxford)

Frege, Gottlob (1956), "The Thought," in *Mind*, Vol. 65, No.259,

Hintikka, Jaako (1998), "Perspectival Identification, Demonstratives and "Small Worlds,"" in *Synthese*, Vol. 14, pp. 203- 232

Kapitan, Tomis (2001), "Indexical Identification: A Perspectival Account," *Philosophical Psychology*, Vol. 14:3, pp. 293- 312

Kapitan, Tomis (2006), "Indexicality and Self- Awareness," in *Consciousness and Self-Reference*, Uriah Kriegel and Ken Williford (eds.) MIT Press (Cambridge)

Kaplan, David (1989a) "Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals," in *Themes from Kaplan*, Almog, Perry, Wettstein (eds.) Oxford University Press (Oxford)

Kaplan, David (1989b) "Afterthoughts," in *Themes from Kaplan*, Almog, Perry, Wettstein (eds.) Oxford University Press (Oxford)

Lycan, William (2000) *Philosophy of Language*, Routledge (London and New York)

Macquarie Dictionary (1982) Macquarie Library Pty. Ltd (Sydney NSW)

Perry, John (1977), "Frege on Demonstratives," in *The Philosophical Review*, Vol. 86, pp. 474- 497

Perry, John (1995) "Indexicals and Demonstratives," in *Companion To The Philosophy of Language*, (eds. R. Hale and C. Wright, Blackwell Publishing (Oxford)

Taylor, Ken (1998) "Language and Context," in *Truth and Meaning: An Introduction to the Philosophy of Language*, Blackwell Publishing, Oxford

Wettstein, Howard (1984), "How to Bridge the Gap Between Meaning and Reference," in *Synthese*, Vol.58, pp. 63- 84

What Can Kaplanians Learn from Perspectivist Accounts of Creativity?

Wettstein, Howard (1986) "Has Semantics Rested on a Mistake?" in *The Journal of Philosophy*, Vol.83, No. 4, pp. 185- 209

Wettstein, Howard (1989), "Turning The Tables on Frege or How is it that "Hesperus is Hesperus" is Trivial?" in *Philosophical Perspectives*, Vol. 3 Theory of Mind and Action Theory, pp. 317- 339

Wettstein, Howard (1988), "Cognitive Significance without Cognitive Content," in *Mind*, New Series, Vol. 97, No. 385, pp. 1-28

Harry Potter¹ and the Time Travel Paradox: an analysis of predestination and ontological paradoxes as encountered in J. K. Rowling's *Prisoner of Azkaban*.

Steph Rennick
BA Candidate
Macquarie University

1: Introduction

It has been said that in any book or film, once you introduce time travel, all bets are off.² This is due in no small part to the multitude of paradoxes associated with such an endeavour. These have been explored in countless books and films, and are the focal point of numerous theories and analyses. This paper is an examination of the paradoxes evident in *Prisoner of Azkaban (PoA)*, and an analysis of some of the common contemporary theories of time and time travel as a possible method of resolving them.

A "paradox" is defined by some to encompass only those scenarios or lines of reasoning that lead to an explicit contradiction.³ While this narrow, formal definition is useful in some circumstances, here I employ the term in a looser sense to include those situations where a clash in our beliefs or intuitions arises from an apparently sound piece of reasoning. That is, whereas in most cases arguing from true premises would lead us to true conclusions, a paradox has a conclusion that strikes us as 'manifestly false, unacceptable or undesirable.'⁴ If the conclusion can be neither embraced nor avoided, we are forced to re-examine any attendant assumptions or premises as candidates for rejection.⁵ Whether it concerns an explicit contradiction or merely an intuitional clash, paradoxes 'demand philosophical attention because they invariably point to confusions or inconsistencies in our concepts and reasoning.'⁶ With this wider definition in mind it is possible to examine a range of perplexities that arise when considering time travel.

In *PoA*, two particular types of paradox are evident: the predestination and the ontological. The former, also called a causal loop or closed time loop,⁷ occurs when a

¹ Rowling, J. K. *Harry Potter and the Prisoner of Azkaban* (1999, Bloomsbury, London). HARRY POTTER is a registered trade mark of Warner Bros.

² Penny & Greg, *Podcast #23: Going back in time* (March 8, 2007) <http://www.hpprogs.com> - Accessed 20/07/08

³ Such as in paradox 1 below: Buckbeak is both killed and not killed.

⁴ Cave, Peter *Can A Robot Be Human?* (2007, OneWorld, Glasgow) p. xiii

⁵ Dupre, Ben *50 Philosophy Ideas You Really Need to Know* (2007, Quercus Publishing Plc, London) p. 110

⁶ *Ibid.*, p. 110

⁷ Lewis, David "The Paradoxes of Time Travel," in French & Brown (eds.) *Puzzles, Paradoxes and Problems* (1987, St Martin's Press, New York) p. 238

Harry Potter and the Time Travel Paradox

time traveller becomes caught in a loop of events that dictates or “predestines” his journey into the past; such as an archaeologist who travels back several millennia in an attempt to discover the origins of a recently discovered human skeleton, ‘only to die and become that very skeleton.’⁸ The ontological paradox, which often accompanies the predestination paradox in science-fiction,⁹ occurs when an item or piece of information is delivered from the future to the past, only to become the same thing that was originally passed back, with no discernible origin. In his *Paradoxes of Time Travel*, Lewis presents a scenario in which a time traveller returns to the past to talk to an earlier version of himself.¹⁰ During this conversation, he tells his younger self how he built his time machine. His older self had been told (by himself) when he was younger, and preserved the information via the ‘causal processes that constitute recording, storage, and retrieval of memory traces.’¹¹ His younger self possessed the knowledge because he had been told the necessary details by his older self in the course of the conversation. The problem lies in discerning the origin of the information, and perhaps more poignantly, in ascertaining why the series of events occurred in the first place.¹²

Multiple theories of time and time travel have been proposed, depicting the former in a myriad of ways, and affording the time traveller varying degrees of influence. Arguably the most successful of these, in terms of coherency and resolution of paradoxes, is the 4-D View of Time. This paper examines the work of Lewis and Smith, as well as considering mutable universe theories presuming a branching time stream or parallel universes. It considers specifically the paradoxes found in *PoA*, but these are analogous to many of the objections raised against time travel theories, and the success of the latter in dealing with these perplexities is reflective of their effectiveness as a whole. While many of the paradoxes in question are answerable, those concerning closed causal loops (see below) are the most difficult to circumvent, and even the most robust theories do so only with limited success.

II: The Plot

The time-travel phenomenon and relevant paradoxes we are concerned with here lie in the latter chapters of *PoA*.¹³ Harry Potter, famous teen-wizard, is having a rather bad day. Hagrid’s hippogriff, “Buckbeak,” has apparently been executed, and our hero is attacked by Dementors and possibly about to die. On the verge of unconsciousness he sees what he believes to be his dead father, who casts a Patronus charm to save him. He wakes up in hospital after his miraculous rescue, only to find out that his

⁸ French, P. A. & Brown, C. (eds.) “Time Travel,” in *Puzzles, Paradoxes and Problems* (1987, St Martin’s Press, New York) p. 208 cf. Heinlein, “– All You Zombies –” (1958) which includes a very complex causal loop of this kind, involving a young man who sets into motion the circumstances by which he becomes his own mother, father, son and daughter.

⁹ cf. Heinlein, “By his bootstraps,” (1941) for a particularly fine example.

¹⁰ Lewis, *op. cit.*, p. 237

¹¹ *Ibid.*, p. 236

¹² *Ibid.*, pp. 236-7

¹³ Specifically pp. 231-317 (Chapters 16-22)

newly-discovered godfather is about to be killed. On the advice of his headmaster, Albus Dumbledore, he travels three hours back in time with his friend Hermione, and saves the ill-fated hippogriff. In the hopes of seeing his dead father, he watches the Dementors attacking, only to realise that it was himself he saw, and proceeds to cast the life-saving Patronus (thanks to an odd bit of logic). He then uses the hippogriff to free his godfather, and returns to the hospital just as his earlier self is using the Time-Turner to depart. Thus ill-fortune is averted, several lives are saved, and the novel has a dramatic climax. Voila!

III: Paradoxes

If only things were that simple. There are three particularly evident paradoxes that arise from this turn of events:

1. Buckbeak appears to have both been killed and not killed.
 2. Harry knew he could cast the Patronus because he had "already done it."¹⁴
 3. Harry survives to save himself because he is saved by himself.
1. In the "original" set of events, Buckbeak seems to have been killed by the executioner Macnair, as Harry hears the swish and thud of the axe, and Hagrid's tears at the event. In the "new" set of events, Buckbeak is saved and thus does not die. Thus, both p (Buckbeak was killed) and $\sim p$ (Buckbeak was not killed) appear to be true, and through *reductio ad absurdum* we have a contradiction, and so a rather tricky paradox. Such paradoxes have been used to demonstrate that backwards time travel is impossible, but in this case the time travel itself is a given, and thus a new resolution must be found in order to circumvent it.¹⁵
 2. & 3. There is both an ontological paradox and predestination paradox operating simultaneously here: Harry must survive in order to be able to travel back in time to save himself, which leads to his surviving and travelling back in time to save himself.¹⁶ Thus his time travel appears to be "predestined" by the circumstances of his survival, and his journey merely fulfils his part in the history he had already witnessed. The causal chain is dramatically altered, with the main events becoming both causes and effects of each other (including his survival, which was predetermined by the time travel, so which came first?).¹⁷ The ontological paradox occurs when we consider how it was that Harry saved

¹⁴ Casting a Patronus charm requires a great deal of skill and concentration, something Harry has had difficulty mustering in prior confrontations with Dementors. He had no such lack of confidence this time however, as he had already witnessed his feat, and had no qualms about repeating it.

¹⁵ cf. Saunders, Niki "Godel was Right," in *Cogito Vol. IV No. 2* (2007, Sydney) pp. 21-26

¹⁶ Depending on the theory of time, this cycle could continue *ad infinitum*, or cease once the timelines converge. This is not made clear in the novel, and does not contribute to our analysis of the paradoxes, so I will put it to one side here.

¹⁷cf. Harrison, J. "The Inaugural Address: Dr. Who and the Philosophers," in *Proceedings of the Aristotelian Society, Supplementary Volumes Vol 45*. (1971, Blackwell Publishing) pp. 1-24; Lewis, *op. cit.*, p. 238

Harry Potter and the Time Travel Paradox

himself - how he managed to cast the Patronus. He attributes this feat not to any great aptitude or magical phenomenon, but to confidence born of previous experience - he had already seen himself successfully cast the spell, so had no reason to doubt his ability to do it again. The problem lies in discerning the origin of this confidence, a pursuit which leads to difficult questions. At first, it is all too easy to ask "but how did he do it the first time?," but of course, the first time is the only time, his earlier self witnesses a feat which he later performs, which his earlier self witnesses. The Patronus, it seems, is only summoned once, as once is enough to save Harry.

"It seems" appears frequently in the preceding paragraphs, as the problems are not as they appear at first glance. What follows is an attempt to resolve the paradoxes raised by J. K. R., and in the words of Robert A. Heinlein, "paradoctor" them.¹⁸

IV: Finding a Solution

How one attempts to circumvent these paradoxes depends on the time theory employed. Presuming that time travel is possible, the most basic level theories can be divided into two categories: those that presume a mutable universe, where the past can be changed by the time traveller; and those that presume an immutable universe, where the role of the time traveller ranges from a mere observer, to an actor with limited capacity for influence.¹⁹

1: Having the Capacity for Change: A Mutable Universe

There are several different theories of time under which a time traveller might have the capacity to change the past, but some of the most popular include the idea of a "branching" universe, or parallel universes. These vary from conceptions of vast branching tree-like structures, with each branch corresponding to a scenario's possible outcome, to the traveller diverting the "stream" of time, thus creating a new history different from the original. These theories are often employed to combat the grandfather paradox, where the time traveller journeys into the past to kill his grandfather, thus ensuring his non-existence, and by default, that his grandfather lives - if the traveller is never born, how can he travel back in time to commit murder?²⁰ Under a single-stream analysis of time, the grandfather would be both killed and not killed, which is a contradiction. Under a branching or parallel universe analysis however, the paradox-creating act results in the creation of a new universe or branch in the time continuum, in which the grandfather is dead and the traveller never

¹⁸ Heinlein, R. A. "- All You Zombies -" (1958) <http://ieng9.ucsd.edu/~mfedder/zombies.html>- Accessed 15/07/08

¹⁹ There are multiple theories that contest the possibility of time travel, however, in order to examine the problems travelling backwards in time may pose, it is best to assume that such travel is possible, and thus the question relevant. For an opposing view, see Saunders, *op. cit.*

²⁰ cf. Lewis, *op. cit.*, pp. 237-42 for a lengthier version and analysis of this paradox.

conceived. The grandfather is still alive in the original universe, and the traveller's existence ensured.²¹

Smith describes the action of such a time traveller as not *changing* the past, but *avoiding* it.²² That is, if we assume parallel universes, while the time traveller can prevent his own birth by killing his grandfather in one universe, he cannot prevent the survival of that instance of the grandfather he knew as he grew up, who bought him ice-cream or told him stories. That is, time is not rewound and recorded over; where there is a bifurcation in the time line, there can be no real change, argues Smith, only *avoidance*: he wanted *his* grandfather to die, not 'some analogue of [him] in another universe or temporal dimension.'²³ A time traveller's ability then, presuming a mutable universe, lies not in changing that which has already happened, but rather opening another set of possibilities that lie parallel to the original events, either temporally or spatially, allowing two apparently contradictory events to occur simultaneously. In universe or dimension *A*, the grandfather and time traveller coexist, while in universe *B* (which has either always existed, or spawns as a result of the time traveller's actions, depending on the particular theory), the grandfather is dead, and the time traveller ceases to exist.

How does this match up with the paradoxes encountered earlier in *PoA*?

1. Buckbeak seems to have been killed and not killed.
2. Harry knew he could cast the Patronus because he had "already done it."

Similarly to the aforementioned *grandfather paradox*, Harry's travelling back in time and saving the hippogriff may have caused a divergence in the time "stream," causing Buckbeak to die in one universe, and live in another, thereby averting the paradox. However, this analysis fails to account for (2), Harry's confidence in his ability to cast the Patronus due to his prior accomplishment of the act. Because the problem with this scenario lies not in two opposing states being simultaneously true, but rather in the existence of a loop with no discernible origin, there seems to be no catalyst for branching, and no guaranteed resolution to the perplexity whether parallel universes exist, or not. This is also the case for

3. Harry's surviving in order to save himself.

A catalyst for branching seems not to exist, and in this the predestination paradox can be considered the opposite of the grandfather paradox. In the latter, the reason for travelling in time (to kill the grandfather) is eliminated, as in the new universe, neither the grandfather nor the time traveller exist. In the predestination paradox, however, as is the case in (3), the events necessarily require the time travel to take place. There

²¹ Lewis, *op. cit.*, p. 242

²² Smith, Nicholas J. J. "Bananas Enough for Time Travel?" in *The British Journal for the Philosophy of Science*, Vol.48 No.3 (September 1997, Oxford University Press) p. 366

²³ *Ibid.*, p. 366

Harry Potter and the Time Travel Paradox

may be room to argue that if Harry - after being saved - chose not to travel in time, a branch might occur, where the earlier Harry in the new branch is not saved, and thus dies; but on an intuitive level this seems to cause problems for the whole loop of events. The limitations of mutable universe theories are not limited to the world of Harry Potter - while a mutable universe does not prevent the possibility of backwards time travel, it fails to resolve common time travel paradoxes, such as the aforementioned self-discovering skeleton.²⁴

Under the branching or parallel universe account of time travel then, we are left with something of a resolution to paradox (1), a dubitable explanation for (3), and no convincing method of circumventing (2).²⁵

2: Lacking the Capacity for Change: An Immutable Universe

The idea that time is immutable arguably does not agree with the commonly conceived notion of time as a thing that flows or moves, or that we move through.²⁶ Such an account supposedly eliminates the possibility of paradoxes occurring, as changes cannot be made that lead to a contradiction: for example, you could never kill your grandfather if you went back in time, because it is known that your grandfather was not killed, and that you were born.²⁷ If nothing can change, then you must be born, and your grandfather must live long enough to ensure this is the case. A common view that falls into this category is the 4-dimensional notion of time, where time is a dimension 'strictly analogous to the three spatial dimensions.'²⁸ Under this type of theory, of which Lewis proposes a particularly coherent version, the universe is depicted as a 'four-dimensional manifold of events.'²⁹ To discern the capability of this theory to combat the aforementioned paradoxes, it is necessary to examine where it differs from the conventional or "commonsense" theory of time, and whether it appears to relate to the world J. K. R. envisioned when developing the scenario in question.

a: The 4-D View of Time

Theories employing the 4-D View of Time, generally speaking, make the following claims:

1. Supporters of the "naïve" or "commonsense" view that time moves, or that we move through it, are unable to provide a coherent explanation of this motion. If

²⁴ French & Brown, *op. cit.*, p. 208

²⁵ While there are several mutable theories of time that do not involve branching/parallel universes, they have great difficulty in combating the grandfather paradox, and thus for the sake of brevity, have been excluded.

²⁶ French & Brown, *op. cit.*, p. 209

²⁷ Lewis, *op. cit.*, p. 238, Lewis proposes "commonplace reasons" as preventing the time traveller who attempts to change the past, thereby preventing a contradiction. See also Smith, *op. cit.*

²⁸ French & Brown, *op. cit.*, p. 209

²⁹ Lewis, *op. cit.*, p. 232

motion is considered to be a change in location with respect to time (that is, object A is in place x at time T_1 , and then at place y at T_2), what is it to say something moves through time? 'Presumably, to be in different *temporal* locations at different... *somethings*?'³⁰ The problem arises in trying to determine what the *somethings* might be - not places, since we can supposedly move through time while stationary; nor times (which would just equate to being in different temporal locations, which was already a given). Thus the notion of what it would mean to travel through time is incoherent.

2. Different temporal stages of me have different properties. That is, just as there's a part of me between my collarbone and bellybutton, so too is there a part of me between my fifth and sixth birthdays.³¹ The collection of parts, both spatial and temporal, makes up a person.³² The five year old and the twenty year old are both *parts* 'of the enduring object which is me,'³³ each with some variations in experience, memories, and other qualities.³⁴ This allows changes to occur during an individual's lifetime, without dictating that you are no longer the same person you were several years ago. If this paper should alter your opinion on paradoxes or Harry Potter, then there will be a difference between the temporal stage of you that started reading at the beginning, and the one which finishes - but both belong to the same person.³⁵

3. There can be facts about the future. While many versions of the "naïve" view claim that there are no facts yet about tomorrow (because time has not arrived there yet), Lewis appeals to the *principle of bivalence*, which states that every proposition is either true or false, and cannot be neither, thus there must be "true" facts about the future.³⁶

The 4-D view would have us conceive ourselves as stretched over both time and space.³⁷ In H. G. Wells' *The Time Machine*, the time traveller argues that a cube with only length, breadth and thickness cannot exist, because if it doesn't last for any time at all, it cannot have a real existence. Thus 'any real body must have extension in *four* directions: it must have Length, Breadth, Thickness, and - Duration.'³⁸

b: Immutability and the 4-D View

Even with a clear understanding of the 4-D theory of time, it may be necessary to ask, why does such a view dictate immutability? Lewis defines change as 'qualitative

³⁰ French & Brown, *op. cit.*, p. 209

³¹ *Ibid.*, p. 211

³² Lewis, *op. cit.*, p. 232

³³ French & Brown, *op. cit.*, pp. 209-210

³⁴ Lewis, *op. cit.*, p. 232

³⁵ adapted from an example by Lewis, *op. cit.*, p. 232

³⁶ Dummett, M. A. E "Truth," *Proceedings of the Aristotelian Society, New Series* Vol. 59 (1958-59, Blackwell Publishing) p. 143 cf. Ryle, Gilbert "It was to be," in French & Brown (eds.) *Puzzles, Paradoxes and Problems* (1987, St Martin's Press, New York) pp. 300-307

³⁷ French & Brown, *op. cit.*, p. 210

³⁸ Wells, H. G., *The Time Machine* (1961, Garden City, New York) pp. 193-4 Oddly, Wells mixes this part of the 4-D view with the Naïve View that our consciousness moves through time.

Harry Potter and the Time Travel Paradox

difference between temporal parts of something,³⁹ and thus anything lacking temporal parts cannot undergo change. He argues that numbers fall into this category, and so do events at any moment in time, since they 'cannot be subdivided into dissimilar temporal parts.'⁴⁰ In such a view, the future is already determined, and time travellers move about on the static continuum. Thus, should a time traveller decide to travel to the past, and save a life or prevent a birth, he has already been there and done so. To illustrate this, and combat objections to such a view, Smith proposes a scenario in which every object 'has written upon it the date on which it will cease to exist,'⁴¹ as a result of a prescient god or time traveller coming back from the future. One cannot preserve things beyond their expiry date, or destroy them prematurely; any attempts fail. The fact that Smith's God has seen an object survive as long as it has, is because it has survived that long - that is, because my attempts to destroy my pen have failed until 2003, the date 2003 is written on it; the date being there doesn't *prevent* me from destroying it beforehand, but rather is there *because* I couldn't or didn't destroy it beforehand; it was always destroyed on the date in question. 'Future events do not have to conform to the date. It is the date that conforms to future events.'⁴² There is no "second-time-around,"⁴³ the way things are at any moment in time, is the way things always were at that moment in time. If you were born, then you could not have prevented your birth by killing your grandfather: any attempts you might make would fail, and they were always part of the timeline. 'A complete chronicle of events occurring at the time to which she travels describes [the time traveller's] arrival and her actions, before she departs.'⁴⁴ That is, anything you do while time travelling in an immutable universe you always did - you cannot change the past, and by travelling in time, your influence is limited to necessarily preserving the timeline and thus supposedly avoiding contradictions.

c: Harry in 4-D?

The universe in which *PoA* is set could conceivably fall into this category, as both the book and film version depict Harry and Hermione intending to change history upon arriving in the past, only to discover that they were performing actions they had previously witnessed - such as casting the Patronus.⁴⁵ This is an example of the predestination paradox in effect: they are fulfilling parts that existed in the history all along, and thus the time travel becomes necessary to ensure consistency. However, it can be argued that the time travel prevents any such paradoxes forming, as seeing the actions of their future selves but not travelling back in time would result in an inconsistency. Their backwards time travel does not allow them to change anything in

³⁹ Lewis, *op. cit.*, p. 232

⁴⁰ *Ibid.*, p. 232

⁴¹ Smith, *op. cit.*, p. 376

⁴² *Ibid.*, p. 377

⁴³ *Ibid.*, p. 365

⁴⁴ *Ibid.*, p. 366

⁴⁵ *PoA*, p. 300

the past, and thus is self-consistent and supposedly devoid of paradoxes.⁴⁶ Any actions they performed must have always been performed, and thus the time travel predestined. In other words, they could not have chosen not to go back in time, because they had already done it.

But is the journey really devoid of paradoxes? Does the 4-D view eliminate the three identified earlier?

1. Buckbeak seems to have been both killed and not killed.

This paradox is perhaps the most easily averted. In a mutable universe, it can be circumvented via branching theory, and in an immutable universe, it is avoided by predestination and consistent time travel. That is, we can assume Buckbeak was never killed. Earlier Harry heard the "thud" of the axe and Hagrid crying, and presumed the execution had taken place. Later Harry saves Buckbeak, and witnesses Macnair swinging his axe into the fence in frustration and Hagrid crying tears of joy at the hippogriff's escape. Under this analysis there is only one possible version of this event - the one in which Buckbeak is saved - as it is the only way the time line remains consistent. Thus time travel is a necessary element in the scenario, and Harry is fulfilling a "fated role;" he has not changed the past, only precipitated the future he remembers.

2. Harry knew he could cast the Patronus because he had "already done it."
3. Harry survives to save himself because he is saved by himself.

Many supporters of the immutable universe find no problems in such examples, as the predestined nature of the time travel is part of the solution.⁴⁷ That is, Harry is fulfilling the role of both victim and rescuer, roles that he must have always fulfilled in order to maintain consistency in the singular time line. While examining the external sequence of events, this seems plausible - Harry's attack and Harry's rescue occur simultaneously, and the time travel is a requirement of the scenario playing out as witnessed - if Harry did not go back in time, he would have been unable to save himself as he does. Thus when we ask the question - *how did Harry survive in order to return to the past?* The answer seems obvious, *he was saved by his time-travelling (or "later") self.* But if we examine the series of events from Harry's personal point of view, in terms of his own internal chronology, things go awry.

V: External vs. Personal Time

Time travel 'inevitably involves a discrepancy between time and time,⁴⁸ because the separation in time between the departure and the arrival of the time traveller is

⁴⁶ cf. Friedman, Morris *et al.* "Cauchy Problem in Space-times with Closed Time-like Curves," *Physical Review D* 42 (1915)

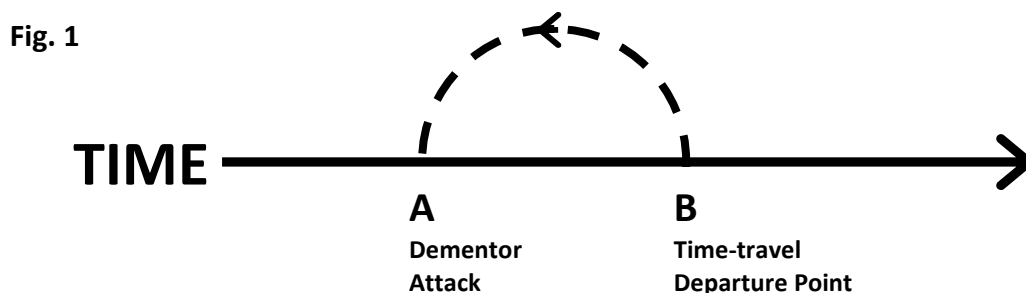
⁴⁷ cf. Penny & Greg, *op. cit.*; Lewis, *op. cit.*, pp. 237-42

⁴⁸ Lewis, *op. cit.*, p. 231

Harry Potter and the Time Travel Paradox

different to the duration of his journey - the trip itself might take an hour, as measured by the traveller's watch or the aging of his cells, but he may be transported centuries into the past.⁴⁹ To account for this, Lewis distinguishes between time itself, or *external time*, and *personal time*, which he defines functionally (not operationally) as that which 'occupies a certain role in the pattern of events that comprise the time traveller's life.'⁵⁰ That is, certain regularities hold in the course of an individual's life - infantile stages precede adult stages; memories, knowledge and experience accumulates; hair and nails grow; wrinkles appear - and these can be mapped out against external time.⁵¹ In a time traveller's life, these stages do not correspond to external time in the normal way, but the temporal stages still exhibit the same regularities. Lewis describes *personal time* as that which plays the same role in the time traveller's life that time plays in the life of the average person, and thus we can employ our usual temporal vocabulary when discussing him.⁵²

If we consider Harry's personal timeline, the reason the analysis clashes with our instincts becomes clear: something not yet accounted for in the sequence of events must have occurred for Harry to have survived. In the external time line, Harry is saved by himself due to his return to the past, allowing for two Harrys to simultaneously exist for a period of time:



N.B. Between points A and B, two Harrys exist simultaneously until "early" Harry goes back in time and "later" Harry continues to exist from where he left off and the original time line.

In Fig. 1, Harry travels back in time at B, to save himself in the aftermath of A. When "saved" Harry reaches B on the timeline and returns to A, the "later" Harry continues to exist where he left off on the original time line.

However, if we examine Harry's personal chronology, an immutable universe results in a problem as shown in Fig. 2.

⁴⁹ cf. Harrison, *op. cit.*, pp. 1-2

⁵⁰ Lewis, *op. cit.*, p. 233

⁵¹ *Ibid.*, p. 233

⁵² *Ibid.*, p. 233. This allows for phrases such as "soon he will be in the past" to be uttered without contradiction, as the "soon" refers to his personal time, and "the past" to the external timeline. Cf. Heinlein's "- All you Zombies -" where the protagonist speaks of his "thirty subjective years of time-jumping."

Fig. 2

Harry hears the axe and Hagrid's cries. He assumes Buckbear is dead.



Harry is attacked, sees the mysterious rescuer and falls unconscious.



Harry wakes up in hospital.



Harry travels back in time to save Buckbear and Sirius.



Harry casts the Partonus to save himself.



Harry returns to the hospital and the timelines converge.

In Harry's personal timeline, there is a gap between his falling unconscious and waking up in hospital. If he has not yet gone back in time to save himself, how was he saved?⁵³ If one focuses solely on external time, this is not a problem; at temporal location *A* Harry is saved by himself because he travelled to the past at *B*. There is backward causation involved certainly, but there is no apparent contradiction - Harry survives. He could not have chosen not to travel in time, because that would have led to a similar case as with Buckbear - Harry would have saved himself but not survived to save himself; he would have survived and not survived simultaneously. Therefore, his time travel was necessary to preserve the external timeline. However, when one shifts focus and considers Harry's personal chronology, an immutable universe does not seem to account for the gap. Harry's surviving in order to travel back in time and save himself seems to involve an intuitive clash.⁵⁴

This is not overly concerning for J. K. R., who could create a number of possible scenarios to explain the discrepancy,⁵⁵ but it is potentially a problem for the philosopher. Turning to the Novikov *self-consistency principle* or Lewis' "commonplace reasons" as an answer does not seem particularly convincing, under which Harry must have always saved himself because the past cannot be changed, and he saves himself in the series of events we are privy to in *PoA*.⁵⁶ In a mutable but non-branching universe, this would be

less of a problem: the very first time Harry is saved (with regards to his personal time) the rescuer may have been someone else altogether, or the circumstances different. This "first time" could then be wiped clean thanks to changes made upon his return to the past and the victim/saviour loop taking its place.⁵⁷

⁵³ It is tempting to add "the first time" to the end of this question, but that just leads to further issues. However, this is essentially where the problem lies.

⁵⁴ It is almost as if there's a hole in the tapestry of time, which Harry (in terms of his personal chronology) later goes back and repairs. In the external time line, the duration between the hole's creation and its repair is non-existent, but in Harry's personal timeline, it exists for several hours.

⁵⁵ For example, Dumbledore could have saved Harry the first time, only to have Harry go back in time to cover his interference, thus "recording over" the past, and creating the loop.

⁵⁶ cf. Lewis, *op. cit*; Friedman, Novikov *et al*, *op. cit.*: These are common responses to the grandfather paradox - where the grandson will be unable to kill his grandfather due to some commonplace reason e.g. the gun sticking, the person in question not being the real grandfather etc in order to maintain consistency.

⁵⁷ The fact that Harry does not remember any such event occurring is not problematic: firstly, he was unconscious, so unlikely to remember what really happened, and secondly, memory loss is often a side-

Harry Potter and the Time Travel Paradox

Perhaps the 4-D view or other immutable universe theories could be altered to resolve this type of conflict between external and personal time, as Lewis has attempted with regards to causality contradictions.⁵⁸ Some might argue that it does not need to be changed, as this is a fairly obscure scenario, and thus is unlikely to ever be a problem, should time-travel take place. However, we still have one paradox to resolve, before we can determine the effectiveness of the 4-D view in combating such problems:

2. Harry knew he could cast the Patronus because he had “already done it.”

This sets up a loop similar to that previously mentioned of the time-traveller who tells his earlier self how to build a time machine. The problem there lay in discerning the origin of the information passed from one to the other, and in *PoA*, it lies in determining the source of Harry’s certainty. Lewis describes this as a ‘closed causal chain,’ where some of the causal links are conventional and some reversed.⁵⁹ For example, Harry muttering an incantation (cause) results in the Patronus (effect), as opposed to Harry casting the Patronus (cause) which takes place later in his personal time than him being saved (effect). Each link in the chain has a causal explanation, as it is caused by other events on the loop, but the loop as a whole in both this case and that of Lewis’ time-traveller, is apparently inexplicable. Lewis describes this as “remarkable,” but does not seem at all perturbed by it, unlike many philosophers, who consider the existence of such a paradox to be a deal-breaker for the possibility of time travel.⁶⁰ He argues that this is not so different to many other inexplicabilities we accept, such as the lack of a cause for God, the Big Bang, or ‘the decay of a tritium atom.’⁶¹ If these are possible, then why not the inexplicable loops that could arise if one travelled in time? Smith argues that these puzzling cases fail to prove time travel ‘impossible or incoherent, or even improbable,’⁶² as they rely on the assumption that objects or information very rarely come from nowhere. His exception involves mishearing - one person might mishear what another person whispers, mistakenly believing that the latter said something wise and original, which neither ‘would, in fact, have ever thought of.’⁶³ He argues that if things regularly started popping out of nowhere, we would not question where the idea or object originated, we would ‘simply accept it without raising an eyebrow.’⁶⁴ Smith believes that causal loops are inexplicable not because they do not conform to current theories, but rather because they involve phenomena regarding which there is no existing theory, and thus should not be immediately rejected as impossible.⁶⁵

effect attributed to time travel, especially where the past has been changed (cf. *Dr. Who*, Christopher Pike’s *Time Terror*)

⁵⁸ cf. Lewis, *op. cit.*, p. 236; Lewis “Causation,” in *The Journal of Philosophy* Vol. 70 (1973) pp. 556-567

⁵⁹ Lewis, *op. cit.*, p. 236

⁶⁰ Lewis, *op. cit.*, p. 236; Saunders, *op. cit.*

⁶¹ Lewis, *op. cit.*, p. 237

⁶² Smith, *op. cit.*, p. 371 (footnote 11)

⁶³ *Ibid.*, p. 371

⁶⁴ *Ibid.*, p. 371

⁶⁵ *Ibid.*, p. 370

Harry's reasoning for his ability to cast the Patronus, a feat well beyond most wizards of his age, is that he had done it before. But he hadn't. The Patronus he sees and the Patronus he casts are one and the same, just seen from two different points-of-view with regard to his personal chronology. The Patronus that gave him the confidence to cast is identical to the Patronus he cast because they are one and the same thing. Is it enough to merely shrug and say, like Lewis, "Strange!"⁶⁶ or must we write off time travel altogether? Can we accept that this is an oddity, but not an impossibility?⁶⁷

VI: Conclusion

Thus far, there has been no adequate, definitive answer. Unlike (1) and (3), (2) is truly a paradox, and seems to stand up to each theory of time - no matter which you employ, he gains his confidence to repeat an event from witnessing the event; but what appears to be two, is really one and the same, and he accomplishes a feat with the confidence supposedly derived from that very same event. Whether time is a single stream or a plane, mutable or immutable, branching or cyclical, the end result is essentially the same - Sirius and Buckbeak go free, and Harry and Hermione survive - only the mechanics are different. But it is in the mechanics that the answers to the paradoxes reside. In a mutable universe, we have an answer for (1), and problems with (3). In an immutable universe, we have answers for (1) and (3). But (2) is paradoxical under any of the common theories of time. What this means for the possibility of time-travel varies: for some it is proof that such an endeavour is impossible, for Lewis, Smith and science-fiction authors it is merely an amusing and mind-bending oddity.⁶⁸ Indeed, it is strange, and more than a little frustrating. Smith's argument regarding inexplicability and mishearing whispers may account for objections concerning the origins of ideas, but it seems far less convincing when trying to account for physical objects, or even in the case of Harry and his logic-defying burst of confidence. To claim, however, that "Things do not just pop out of thin air!" is unsatisfactory also, as not having experienced a phenomenon does not necessarily guarantee its impossibility: as Smith proposes, 'it is no better than arguing that one has never seen humans fly...even as the Wright brothers set up in the neighbouring field.'⁶⁹ Time travel itself is an example of such a phenomenon,⁷⁰ and the mind-bending perplexities involved are not enough to write it off as an impossibility, at least for some. Theories concerning an immutable universe, specifically Lewis' 4-D view of time, seem the best equipped to deal with the multitude of paradoxes inherent in the issue of time travel. The exception lies in the inexplicable closed causal loops, like (2) above, where an idea, object or sensation is apparently without origin. It seems insufficient to say merely that a pursuit as wacky as time travel is bound to involve perplexities and

⁶⁶ Lewis, *op. cit.*, p. 237

⁶⁷ *Ibid.*, p. 231

⁶⁸ *Ibid.*, p. 231

⁶⁹ Smith, *op. cit.*, p. 368

⁷⁰ That is, it would be inappropriate to claim that objects popping out of thin air is something I've never experienced, thus it is impossible, if I'm simultaneously claiming that time travel is possible, though I've never experienced it.

Harry Potter and the Time Travel Paradox

phenomena never before experienced or considered, but a more satisfactory response, it seems, has yet to pop into existence.

Regardless which theory of time one prefers, the problems posed by predestination and ontological paradoxes permeate even the best defences. Without a time machine, or at least a magic wand, the consequences of these paradoxes are hypothetical at best, but still we strive to solve them, a pursuit that haunts us with its intricacies, leading even the most self-assured philosopher to occasionally (and perhaps unfortunately) be left with little more than "Strange!"⁷¹

Bibliography

Cave, Peter *Can A Robot Be Human?* (2007, OneWorld, Glasgow)

Dummett, M. A. E "Truth," *Proceedings of the Aristotelian Society, New Series* Vol. 59 (1958-59, Blackwell Publishing) pp. 141-62

Dupre, Ben *50 Philosophy Ideas You Really Need to Know* (2007, Quercus Publishing Plc, London)

French, P. A. & Brown, C. (eds.) "Time Travel," in *Puzzles, Paradoxes and Problems* (1987, St Martin's Press, New York) pp. 207-211

Friedman, Morris, Novikov, Echeverria, Klinkhammer, Thorne & Yurtsever "Cauchy Problem in Space-times with Closed Time-like Curves," in *Physical Review D* 42 (1915)

Harrison, J. "The Inaugural Address: Dr. Who and the Philosophers," in *Proceedings of the Aristotelian Society, Supplementary Volumes* Vol 45. (1971, Blackwell Publishing) pp. 1-24

Heinlein, R. A. "— All You Zombies —" (1958)
<http://iengg.ucsd.edu/~mfedder/zombies.html>- Accessed 15/07/08

Lewis, David "Causation," in *The Journal of Philosophy* Vol. 70 (1973) pp. 556-567

Lewis, David "The Paradoxes of Time Travel," in French & Brown (eds.) *Puzzles, Paradoxes and Problems* (1987, St Martin's Press, New York) pp. 231-243

Penny & Greg, *Podcast #23: Going back in time* (March 8, 2007)
<http://www.hpprogs.com> - Accessed 20/07/08

Rowling, J. K. *Harry Potter and the Prisoner of Azkaban* (1999, Bloomsbury, London).

⁷¹ Lewis, *op. cit.*, p. 237

Steph Rennick

Ryle, Gilbert. "It was to be," in French & Brown (eds.) *Puzzles, Paradoxes and Problems* (1987, St Martin's Press, New York) pp. 300-307

Saunders, Niki "Godel was Right," in *Cogito Vol. IV No. 2* (2007, Sydney) pp. 21-26

Smith, Nicholas J. J. "Bananas Enough for Time Travel?" in *The British Journal for the Philosophy of Science*, Vol.48 No.3 (September 1997, Oxford University Press) pp. 363-389

Wells, H. G. *The Time Machine* (1961, Garden City, New York)

Book review

A New Platonic Theism

Tim Smartt
BA Candidate
University of N.S.W.

*There is no theology that has not stood in Plato's shadow;
Be it in the Judeo-Christian West, or in the Islamic East.*¹
– Walter Burkert

Platonic theology is far from straight forward. There is no single dialogue which focuses exclusively on the topic, and nowhere does Plato systematically inquire into the nature and activity of the Divine. The most extensive passage on theology in Plato is *Laws* Book X, and this limits its inquiry to the rationality of theistic belief. It focuses only on arguments for the existence of God and says little about the kind of existence this is and even less about the way in which the world and God might interact. Despite this, theology is a central issue for Plato and it is raised at some point in connection with almost all of his other major areas of inquiry.² In his earlier dialogues, Plato seems to adopt a theology which was either compatible with the Greek civic religion of the time or a now familiar conception of the Divine as the powerful and transcendent God of classical theism. Then, in his later post-*Parmenides* work, more attention is given to the World Soul, and the role of the Forms is occasionally replaced by the Demiurge or God. In these later dialogues, theological questions are increasingly developed within Plato's philosophy of nature and movement. What this leaves for the Plato scholar attempting to talk accurately about Plato's theism, or the philosopher of religion attempting to think philosophically and theologically along lines sketched by Plato, is a patchwork of different questions and thoughts, raised in different dialogues and within the context of larger investigations. Daniel A. Dombrowski, in a recently published study of Plato's religious thought, *A Platonic Philosophy Religion*, tries to join the dots into a coherent picture of Plato's theology which challenges traditional interpretations, arguing that they venerate Being as opposed to Becoming, and so offer a simplification of Plato's theological thought.³

Dombrowski offers a philosophy of religion that seeks to be both Platonic and panentheistic.⁴ He sees his task as bringing together the process theism of the likes of

¹ Burkert, Walter. *Greek Religion*, trans. John Raffan, (Cambridge: Harvard University Press, 1985), p. 321.

² For example, significant questions about theology are raised in relation to virtue and morality in *Euthyphro*; in relation to politics in *Statesmen*, *Laws*, and *Republic*; in relation to aesthetics in *Ion* and *Symposium*; and in relation to metaphysics in *Phaedo* and *Timaeus*.

³ Dombrowski, Daniel A. *A Platonic Philosophy of Religion* (Albany: State University of New York, 2005)

⁴ Panentheism (lit. "all in God") can be understood as a reaction to classical theism, and as distinct from pantheism. The crucial difference is between how God relates to the world, and what this means for the being of God. I offer the following workable definition: panentheism affirms that while God and the world are ontologically distinct and God transcends the world, the world is *in* God ontologically. This is in contrast to the

Charles Hartshorne and Alfred North Whitehead with particular aspects of Plato with the two-fold purpose of outlining a new philosophy of religion whilst moving closer to properly understanding Plato's theology.⁵ We will first look at Dombrowski's four major claims, presenting the central theses of this new work, and then offer some initial critical reflections on this work's contribution to contemporary Plato scholarship.

Dombrowski states, 'the goal of the present book is to provide a counter-balance to previous treatments of Plato's thoughts on God that overstate the case for his ontolatry; that is, for his worship of being as opposed to becoming.'⁶ Dombrowski is calling for the perspectives of several process philosophers to be taken seriously, while maintaining that it is his project to outline 'a disciplined rational account of the existence and nature of God on the basis of the evidence supplied in Plato's writing.'⁷ Throughout his book, he maintains that the process perspective allows us to see new things in Plato's dialogues themselves, or at least to see them more clearly than they were seen before. He begins by acknowledging that no thinker has been more influential on the Western conceptions of God than Plato, and that there is certainly something bold in his claim that most interpreters of Plato's theism over the centuries have either not noticed or have underemphasized the dynamism of his theism. Dombrowski confines his study to the later dialogues in Plato. He argues that post-*Parmenides* is when theological questions come to occupy more space in Plato's thinking. He is not attempting a systematic treatment of all of Plato's religious or theological claims, but is hoping to show that by seeing some unity in the theological teaching of Plato's later dialogues the reader will come to reassess the direction and emphasis of all of Plato's theology.

Dombrowski's first argument is his call to take the World Soul seriously.⁸ The World Soul appears in four of Plato's later dialogues and Dombrowski argues that it is systematically overlooked and Plato scholars must re-engage with it.⁹ Dombrowski rejects the view held by many but exemplified in Richard Mohr, that the World Soul is an *odd* part of Plato's cosmology.¹⁰ Mohr treats it as something which is on the fringe of Plato's thought and so as something which can be explained away with little consequence. Dombrowski rather argues that the World Soul ought to be seen as being at the centre of Plato's cosmology, and of much relevance to contemporary philosophical theology. In defending belief in the World Soul and attempting to make it intelligible, Dombrowski largely relies on the metaphysics of process thinker Charles Hartshorne, and the classicist Friedrich Solmsen's account of Plato's theology, each of whom he thinks complements and strengthens the other.¹¹

unqualified distinction between God and the world of classical theism, and the total collapse of God and the world into each other of pantheism. Pantheism maintains that *in some sense* the world is a part of God.

⁵ See particularly, Alfred North Whitehead, *Process and Reality* (New York: Macmillan, 1929) and Charles Hartshorne, *The Divine Relativity* (New Haven: Yale University Press, 1948), *The Logic of Perfection* (LaSalle: Open Court, 1962)

⁶ Dombrowski, op cit., p. 1.

⁷ Dombrowski, op cit., p. 12.

⁸ Dombrowski, op cit., pp. 15-32.

⁹ These dialogues are *Statesmen*, *Philebus*, *Timaeus*, and *Laws*.

¹⁰ Mohr, Richard. *The Platonic Cosmology* (Leiden: Brill, 1985), p. 171.

¹¹ Solmsen, Friedrich. *Plato's Theology* (Ithaca: Cornell University Press, 1942)

Solmsen places the World Soul firmly within Plato's theology and wider thought. The World Soul and the Demiurge are to be seen together as a dyadic conception of the Divine. They are two aspects of the one God that deal with separate functions - the World Soul being concerned with movement and life, and the Demiurge being concerned with order, design and rationality. To this Hartshorne adds that the universe is a society or an organism of which one member is preeminent, just as human beings or animals are societies of cells of which the mental part is preeminent. This organism is the World Soul and the preeminent member is the Demiurge. The prime function of the Divine body, the cosmos, is to furnish the World Soul with awareness of, and power over, its bodily members. So although there is no special part of the cosmos recognisable as a nervous system, every individual becomes, as it were, a brain cell directly communicating with the World Soul.

At the heart of Dombrowski's rejection of Mohr and his agreement with Solmsen and Hartshorne is the argument that the World Soul in Plato necessarily means he has a view of God that is organically inclusive. He concludes, 'We are not in God as marbles are in a box or as an idea is in a mind. Rather, if we are to take the Platonic idea of a World Soul seriously, then the sort of panentheistic inclusiveness to be considered is the organic inclusiveness of bodily pain in a whole animal.'¹² Accordingly, Dombrowski argues that Plato ought to be seen as a cosmological monist as opposed to a cosmological dualist. God is in and throughout the world, rather than transcendent, pure spirit and wholly other to the natural world. This is not a pantheistic view, that God *is* the world, but a thoroughly *panentheistic* view – that God is *in* the world.

Dombrowski's second major claim is that the well-known definition of being as power found in *Sophist* has profound consequences for Plato's philosophy of religion.¹³ In this dialogue, Plato raises the idea that being is the power to affect or to be affected by others.¹⁴ Dombrowski applies this ontological principle to the Divine and argues that it is illogical to speak of anything as being omnipotent. Relying on Whitehead and Hartshorne, Dombrowski follows Plato's thinking regarding being as power to ontological conclusions. Beings are instances of dynamic power, and so to exist as a thing is to always have the power, however small or great, to affect or to be affected. To lose this power is to lose ontological status as a thing. So if being is power then any relation between two things in which one of the things was rendered wholly powerless would be a relation in which that first *thing* was absolutely *nothing*. Omnipotent power is thus unintelligible power over the powerless. God can exercise great power over individuals and things, but cannot be said to be omnipotent as an individual cannot be utterly coerced if being *is* power.

By applying Plato's thought to the Divine, Dombrowski does not see himself as saying something negative about God, but as saying something positive and unique about Divine agency in the world. Divine agency is persuasive rather than coercive. This is seen in *Timaeus* when the Demiurge orders the world *ex hyle* (from matter already in existence),

¹² Dombrowski, *op cit.*, p. 32.

¹³ Dombrowski, *op cit.*, pp. 33-50.

¹⁴ See *Sophist* 247e, tr. Nicholas P. White in John M. Cooper (ed.) *Plato: Complete Works* (Indianapolis: Hackett, 1997)

rather than omnipotently creating the world *ex nihilo* (out of nothing).¹⁵ Dombrowski quotes Whitehead here to make the important point that this primordial ordering of the world by the Demiurge is the ultimate act of Divine persuasion, and the victory of persuasion over force.¹⁶ The Demiurge converses with the world, rather than dictating to it, and Dombrowski concludes that this is how we ought to think of God's relating to the world, in a persuasive rather than coercive way. It can be seen that this very much follows from thinking about the kind of omnipresence and intimacy of the World Soul within the world Dombrowski outlined earlier. Before moving on, it is of interest that Dombrowski also makes the claim that this idea of being as power is reflected in the Socratic Method and the structure of the Platonic dialogue. The way that Socrates and an interlocutor influence one another supports the idea that being is the ability to be affected by or affect another. It is an example of one doing something with (not to) one's dialectical partner. If one extends this analogy to the relationship between individuals and the World Soul it suggests not only that the World Soul is intimately involved with individuals and the world as argued above, but also that each individual, by virtue of *being*, has the power to influence and move the World Soul, and so it can be said that the Divine Being becomes and changes.

Dombrowski's third claim is his central one.¹⁷ He takes up the task of outlining the type of process dipolar theism that he is attempting to attribute to Plato, that is, outlining the nature of God according to Plato. This is his main argument against the charge that Plato's theism is a type of ontolatry, which he has previously stated is among the central purposes of the book. Dombrowski takes two well-known elements of Plato and argues that these are not as complete or simple as scholars often render them. First, in *Republic* Plato quite clearly says that God cannot change.¹⁸ These passages are often assumed to be a simple assertion about the immutability, impassibility, omnipotence and indivisibility of the Divine along lines congruent with the classical Aristotelian theism of the likes of St. Anselm and St. Thomas Aquinas. Where Dombrowski breaks with the majority of the scholarship is in asserting that this is not Plato's last word on God, and that this is in fact only a fraction of what Plato thinks about the nature of God. Second, Plato is famous for his dipolar categorical scheme, in which form is contrasted to matter, permanence is contrasted to flux, being is contrasted to becoming. And yet, Plato ends up with, as we have seen, a cosmological monism wherein the World Soul includes everything. Dombrowski takes these two tensions in Plato and argues that they are not contradictory but make sense through dipolar theism. Put simply, dipolar theism is the position that what are often thought to be contradictions, such as activity and passivity or being and becoming, are mutually interdependent correlatives in the Divine. God's being and his becoming form a *single reality*, and there is no reason to leave these two poles in a paradoxical state. Dombrowski states his argument clearly once more through the thoughts of Hartshorne when he says

¹⁵ *Timaeus* 30-31, tr. Donald J. Zeyl in John M. Cooper (ed.) *Plato: Complete Works* (Indianapolis: Hackett, 1997)

¹⁶ Whitehead, Alfred North. *Adventures in Ideas* (New York: Free Press, 1961)

¹⁷ Dombrowski, op cit., pp. 65-80.

¹⁸ *Republic* 380e, 381b-e, 382e, tr. G. M. A. Grube, rev. C. D. C. Reeve in John M. Cooper (ed.) *Plato: Complete Works* (Indianapolis: Hackett, 1997)

A New Platonic Theism

'there is no law of logic against attributing contrasting predicates to the same individual, provided they apply to diverse aspects of this individual.'¹⁹

Dombrowski's point is that there is a way to reconcile the exaltation of both being and becoming, especially in the later dialogues, in terms of Hartshorne's distinction between Divine *existence* and Divine *actuality*. The former concerns that *fact* that God exists, whereas the later concerns *how* God exists. If one were to ask whether God existed, the answer would be given in terms of an unchangeable, self-sufficient, invariable, indissoluble being that always exists and is always God. However, if one were to ask about the mode of God's existence, the answer would be given in terms of becoming and change. In relation to the tension between the Demiurge and the World Soul in *Timaeus* and other later dialogues, it can be seen that they are to be conceived of as different perspectives on the one reality, as Solmsen has already shown above.

To bring these first three strands together, Dombrowski argues that Plato is to be understood as putting forth a theism that is panentheistic and dipolar. It is dipolar in that it stresses that excellences are found on both sides of contrasting categories, and panentheistic in that God is not to be thought of as completely removed from the world as in the case of Aristotelian theism, nor completely identified with it, as in stoic pantheism. Rather, God is world-inclusive, in that God cares for the entire world and all feelings are felt by God as the Divine animal. Plato's theism advocates a cosmological monism and organic inclusiveness of the Divine knit within the world. It is process theism as it sees God as needing to change in order for God to be called perfect, but this is never at the expense of God being greater than any and all others. Plato's theism is neo-classical in that it agrees with traditional theists such as St. Anselm and St. Thomas Aquinas who described God as the supremely excellent, transcendent and the greatest conceivable being. But Dombrowski argues Plato's theism is distinguished from classical theism as it posits that God must necessarily change and develop to be considered *perfect* in a world of flux.

Plato's theism is a middle way between a number of extremes. It seeks to be a middle way between Aristotelian transcendent monotheism on the one hand, traditionally found to varying extents in Abrahamic religions, and stoic pantheism on the other. It seeks to be a middle road between the overconfidence of the theology of the Homeric poets of Greece, and the scepticism of atheism. To both of these extremes Plato's process theism says humans are capable of working their way towards an adequate conception of God in which he is neither far away nor simplistically immanent.

Dombrowski's fourth point concerns how we are to think about the so-called "mystical" elements within Plato's writings in light of this theism.²⁰ Dombrowski argues that Plato ought to be seen as a very theocentric thinker, and one who understands the *telos* of human life as to become as much like the Divine as possible. Dombrowski acknowledges that there is division amongst Plato scholars as to whether or not Plato was a mystic. Indeed the term

¹⁹ Dombrowski, op cit., p. 70. See Hartshorne, *Philosophers Speak of God* (Chicago: University of Chicago Press, 1953) pp. 14-15.

²⁰ Dombrowski, pp. 95-112.

"mystic" is a problematic one to pin down. This aside, Dombrowski is critical of a trend in scholarship which is reluctant to refer to Plato as a mystic at all on the grounds that they see it as undercutting the worth of his thought.²¹ That is, if being a mystic is to despise the senses and hate the world, one would not want to attribute this to Plato as it severs his commitment to reason and seems to prioritise a sort of unintelligible experience. Dombrowski argues that this is an unnecessary scepticism, and herein lays one of his most important contributions. He argues these scholars are right in one sense; it would be a mistake to read Plato as advocating escape from this world. Plato ought to be understood as advocating *transforming* the world, or at least transforming our attitudes to it. He is not a "mystic" in the sense that he places any inherent value in denying the senses, or that his philosophy leads to flight from this world; quite the opposite, Plato is for changing this tangible, sensible world for the better. Dombrowski argues the way that Plato believes we do this is by becoming as much like the Divine as possible, keeping in mind that Plato has a specific conception at work for thinking about the Divine's actuality and relation to the world. This becomes a more complex prospect than many readers of Plato imagine it to be in light of what Dombrowski has argued is Plato's conception of the Divine. Dombrowski does advocate a kind of mysticism as the activity of the Platonic philosopher, but this is qualified in light of what he has shown Plato thinks about the Divine, and so what he believes is the object of the philosopher's contemplation.

On first reflection, this seems entirely plausible, as Plato's great tracts *Republic* and *Laws* are concerned with how to bring about excellent political and social organisation in the present time. In *Laws* the guardians help the civilians become more like God by applying their knowledge of the good to society and educating people. Citizens, by obeying the city's laws, can make progress in becoming like God. So guardians are not fleeing the world, Dombrowski argues, but fight to help citizens reach likeness to God. Dombrowski also appeals to *Philebus* as putting forth the idea that we become like God when we are intelligible causes of change in the world, when we cause good to come about in the world.

Dombrowski reaches the familiar conclusion that Plato advocates that there is something Divine in the philosophical life, that the intellect is the spark of Divinity in humanity and that by exercising intellect, one can assimilate to God. Scholars often note that whereas Greek ecstatic religions had their bodily experiences induced by sex and wine by which they claimed to attain something like salvation from this world, Plato believes it is by engaging the mind in philosophy and intellectual dialectic that a kind of salvation can be attained. Platonic philosophy and many forms of Greek religion pursue the same end – salvation from this world – but they differ in their means of reaching this, the religions with their bodily excesses and Platonic philosophy with its dialectic.²² But Dombrowski argues that this is not an apt way of talking about the goal of Plato's philosophy. He argues that God is our saviour by giving us intellect in this world and we should not try to escape from it. Dombrowski strongly states 'we are redeemed by God *in* this world, not *from* it.'²³ This

²¹ Dombrowski singles out Julia Annas *Platonic Ethics, Old and New* (Ithaca, Cornell University Press, 1999)

²² See for example Michael L. Morgan, "Plato and Greek Religion," in Richard Kraut, ed., *The Cambridge Companion to Plato* (Cambridge: Cambridge University Press, 1992), pp. 227-247.

²³ Dombrowski, op cit., p. 109.

A New Platonic Theism

thinking is quite obviously couched in religious language, but one cannot help but think Dombrowski is suggesting Plato means much more by this thought than simply religious doctrine – that much of Plato’s thought is being summarised here and it has to do with giving us an outlook on *this* life that is fundamentally committed to life in *this* world. Dombrowski concludes that Plato is a mystic of sorts, but that in Plato meditation on the Divine is never opposed to action; rather it is a springboard for excellent action in religion, politics, art and education. This leaves us with a very rich picture of what it means to speak of Plato as a mystic.

A New Platonic Theism? - Initial Reflections

Dombrowski makes a good case. All four theses are firmly rooted in the Platonic dialogues, and he often presents lengthy exegesis to assure the reader of the accuracy of his arguments. Compared with Mohr, Dombrowski attempts to engage with the thorny and odd parts of Plato’s theological thoughts, and this is to his credit. His work does an excellent job of explaining these parts, such as the World Soul and its panentheistic implications. He particularly presents convincing arguments that panentheism is logically entailed by such a concept as a World Soul, and the most accurate way to systematize relation between the Demiurge, the World Soul and a world *of* souls, particularly in *Timaeus*, is through a panentheistic idea of inclusiveness. Dombrowski has presented Plato scholars with a new and thoroughly argued treatment of this important part of Plato’s thought.

If we take Dombrowski’s central theses to be true then this will affect the way we approach many other matters in Plato. One particular way that I think Dombrowski’s work could be applied to wider contemporary debates in Plato scholarship is the fresh perspective it brings on how to think about Plato’s duality, or “Two-Worlds View.”²⁴ It is usually assumed by both those who have had little contact with Plato and those who have read and reread his work that Plato is the archetypal metaphysical, epistemological and ontological dualist. It is assumed that he thought through and articulated the gap between body and mind, matter and spirit, the world of things and the world of the forms, with a depth and conviction that no thinker had done before, and in a way that would set the questions of Western Philosophy for centuries to come.

Ontologically, the divided line in *Republic* points us toward the apparent ontological heterogeneity we see in the world around us, and the different ontological status of all that exists, with visible and physical things being at one end, and the intelligible and *more* real things of the world of the forms being at the other end. This correlates epistemologically with how the philosopher can claim real, penetrating knowledge, with belief and opinion correlating ontologically with the things down the bottom end of the divided line, the things of matter, and true and genuine knowledge corresponding to knowledge gained from the

²⁴ For a summary of the current debate on Plato’s “Two-World’s View,” and an interesting aesthetic response to this challenge in Plato, as distinct from this theological response I am arguing for (along lines sketched by Dombrowski), see Eugenio Benitez, “Philosophy, Myth and Plato’s Two-World’s View,” in *The European Legacy*, Vol. 12, No. 2, 2007, pp. 225-242

apprehension and contemplation of the forms and the *more* real things of the intelligible realm, of which the physical things were only shadows.²⁵

The myth of the fish and the pond, or the myth of the true Earth, in *Phaedo* offers a metaphysically dualist picture along similar lines. At the conclusion of the *Phaedo*, and at the end of Socrates' life, Socrates is confident in the face of death, and explains to his hearers his confidence in not only the immortality of the soul, but the existence of a better and more real world. Socrates recounts a detailed, though somewhat muddled, picture of a person's existence as being like a fish in a pond. The fish takes the pond around him to constitute all of reality, and indeed the fish knows no other reality. But if the fish would swim to the surface and poke his head out, he would see a whole other world, a world that was the real world, of which his pond was only a very small and unimportant part, and which his pond aped. This is what reality is like, explains Socrates. There is a whole other world, a bigger and more real world which we only glimpse in this life like a fish glimpsing a mountain or the night sky by looking up through the hazy surface of his pond.²⁶ This myth is exemplary of what is assumed to be Plato's metaphysical dualism and his "Two-Worlds View" – his commitment to the existence of a better world beyond this world, and a more real existence which one moves on to after death.

Dombrowski does not deal specifically with either the divided line of *Republic*, and its assumed ontological and epistemological dualism or the myth of the fish and the pond in *Phaedo*, and its assumed metaphysical dualism. But given what we have outlined of Dombrowski's arguments, these passages become less straightforward than traditionally thought. Given Dombrowski's thesis about the reconciliation of conflicting duality through a dipolar Divine being, this whole way of thinking about Plato begins to look less convincing, or at least less like *all* of what Plato had in mind. When one thinks about God as being intimately *in* the world, Plato's "Two-Worlds View" is fundamentally reoriented. Questions are raised such as why would one want to escape the world that God is in? To which better world will she go, if God is in this world as a soul is in a body? If cosmological monism is true, and this *whole* cosmos is the body of the Divine World Soul, then there is indeed no other world to escape to which would not also be the body of the Divine! And if the philosopher seeks to become like God, and through using her intellect assimilate to God, this does not entail flight from the world, but engagement with the world and commitment to it, as she is not assimilating to a transcendent deity, but to the deity which is the Soul of this World and is deeply entangled with it. These questions do not of course refute that Plato had a "Two-Worlds View," but certainly serve to muddy the issue for those scholars wishing to conclude that Plato was simply a dualist without thinking through the nature or emphasis of this. The post-*Parmenides* Plato was increasingly occupied with theological questions as being inseparably entangled with metaphysical and ontological questions. These then affect ethical and political questions as they always have in Plato. So by Dombrowski redefining his theology, he fundamentally shifts our perspective on Plato. Dombrowski's treatment of Plato's theology in his later dialogues has a ripple effect on how we read many passages, and ought to lead us to question the assumptions with which we approach Plato.

²⁵ *Republic*, 509d-511e

²⁶ *Phaedo*, 108e-114d

A New Platonic Theism

For Plato students and scholars, this book will provide an excellent resource. It does provide a thorough-going account, rooted in the later dialogues of Plato and strongly supported by the scholarship of modern process philosophers, for understanding Plato as coming to hold a theism which is panentheistic and dipolar. If one is convinced by Dombrowski's arguments from Plato's later dialogues that he ought to be understood as putting forth a panentheistic theology one will come to understand this one part of Plato's thought better and be stimulated to think through how this might effect many other areas of Plato's thought. I have attempted to sketch some initial lines about how this might be so in regards to his "Two-World's View." If Dombrowski fails to convince you from the dialogues of Plato's leaning towards panentheism then this work will still be of value as it points to the main sections on Plato's disparate theological patchwork, attempts to take these seriously, and then offers a unifying and organising system. His works challenges the unconvinced student of Plato to continue to think through these passages and ask oneself afresh what is driving Plato to think like this, and what he is trying to get at through these theological reflections.

Tim Smartt

Bibliography

Annas, Julia *Platonic Ethics, Old and New* (Ithaca, Cornell University Press, 1999)

Benitez, Eugenio "Philosophy, Myth and Plato's Two-World's View," in *The European Legacy*, Vol. 12, No. 2, 2007, pp. 225-242

Burkert, Walter *Greek Religion*, trans. John Raffan, (Cambridge: Harvard University Press, 1985)

Cooper, John M. (ed.) *Plato: Complete Works* (Indianapolis: Hackett, 1997)

Dombrowski, Daniel A. *A Platonic Philosophy of Religion* (Albany: State University of New York, 2005)

Hartshorne, Charles. *The Logic of Perfection* (LaSalle: Open Court, 1962)

Hartshorne, Charles. *Philosophers Speak of God* (Chicago: University of Chicago Press, 1953)

Hartshorne, Charles. *The Divine Relativity* (New Haven: Yale University Press, 1948)

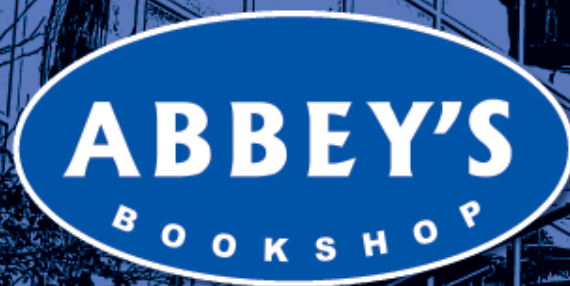
Mohr, Richard *The Platonic Cosmology* (Leiden: Brill, 1985)

Morgan, Michael L. "Plato and Greek Religion," in Richard Kraut, ed., *The Cambridge Companion to Plato* (Cambridge: Cambridge University Press, 1992)

Solmsen, Friedrich. *Plato's Theology* (Ithaca: Cornell University Press, 1942)

Whitehead, Alfred North. *Adventures in Ideas* (New York: Free Press, 1961)

Whitehead, Alfred North. *Process and Reality* (New York: Macmillan, 1929)



Australia's biggest and best selection of Science and History

Also Philosophy • Religion
Politics • Sociology • Anthropology
Literature • Literary Criticism

and the complete range of the Loeb Classical Library

**SPECIAL
ORDERS
WELCOME**

Australia's biggest **foreign language** bookshop upstairs
includes a large range of new and classic fiction

131 York Street Sydney (Opposite QVB) **9264 3111**

also order at **www.abbey.com.au**

