Making Sense of "Essence"

A Critical Examination of The Adequacy of the Modern Philosophical Conception of "Essence"

A dissertation submitted in partial fulfilment of the requirements for the degree of
Master of Arts
in
The Department of Philosophy
Rhodes University

Allison Ross

January 2000

ABSTRACT

The idea that some sub-set of the properties of an object captures what it is to be that thing i.e. that it has an essence which is there to be discovered and about which we can be mistaken - is a common-sense assumption that we use all the time. However, philosophers of this century have regarded the realism about essence with skepticism, arguing that we impose essences on things by the way we define our concepts as opposed to discovering them. Essences are supposedly characteristics of our concepts rather than of objects in the world. This was the orthodox view until a group of philosophers of language developed the theory of direct reference. They claimed that proper names and certain other words refer non-connotatively which entails that the real properties of objects are crucial to the establishment of the reference of such terms. It can be shown that the properties involved in reference determination must be all and only the necessary properties of those objects. This discovery has been taken to mark the rehabilitation of the notion of essence, with an object's essence being taken to be that set of properties which it must have in all possible worlds in which it exists. I will argue that the theory of direct reference is correct up to the point at which it assimilates the necessary properties of objects to their essences. I will show that the set of an object's necessary properties cannot fulfill the role reserved for the concept of essence in metaphysical hypotheses concerning the nature of objects. I will go on to show that a sub-set of a thing's necessary properties can fulfill this role and I will suggest that we identify the members of this sub-set by testing their ability to furnish the kinds of explanations we expect from essences. I will demonstrate how this can be done using the Aristotelian idea that the notion of essence is required in order to explain how it is that objects can persist through change.

DECLARATION OF ORIGINALITY

I declare that the content of this dissertation is my own original work except where appropriately referenced. I further certify that this piece of work has not been submitted for examination at any other university and that to the best of my knowledge, it contains no work accepted for a degree or a diploma at any other university nor any material which has previously been published except where due reference is made in the text.

Allison Ross
January 2000

PREFACE

I would like to thank the staff and postgraduate students of the Department of Philosophy for numerous comments and criticisms which have improved the quality of this work. In particular I want to thank Ward Jones for time and effort he put into discussing my work with me and for challenging me into writing a better dissertation than I thought I could. Finally, I want to thank Francis Williamson who has been a diligent and insightful critic of my work, and a generous, patient and supportive supervisor.

TABLE OF CONTENTS

INTRODUCTION

'Essence' is a metaphysical term which is generally used to denote those characteristics a thing must have in order in order to be the thing that it is and not something else. Philosophers have, in the past, made use the notion of essence in the attempt to gain insight into a wide rage of phenomena including the nature of concrete particulars, the possibility of real as opposed to nominal identity, the relation between the mind and the brain and even the nature of value. Given the broad range of philosophical theses in which the notion of essences could play a crucial role, it is important to have a clear idea of what, if anything, an essence is and of how we might go about identifying the essences of individual objects. It is also important that those who wish to use the concept can defend the claims that there are such things of essences and that if there are they are attributable to actual or real objects as opposed to being features of our concepts of those objects. The latter of these two requirements has been adjudged a difficult, if not impossible, task by many twentieth century philosophers whilst very little attention has been paid to the demand for clarity.

In the first part of this dissertation I want to suggest that, as a result of revolutionary developments in the theory of how proper names refer which took place in the 1960s and 70s, the view that objects have some of their properties necessarily can be defended. In chapter 1 I will give a brief overview of the new theory of reference and argue that it is correct. I will then set out the steps required for the inference from this way of construing reference to the claim that objects have certain of their properties necessarily. Now, many proponents of the revolutionized conception of reference are inclined to believe that by showing that objects not only do but must have necessary properties they have defended metaphysical essentialism - a doctrine which entails that some aspects of any given thing have a kind of privileged status when it comes to understanding what the thing is and how it behaves.

In the second chapter of this dissertation I will argue that this belief is false. I will claim that that there is more to a particular property's being essential than its being a property the object must have in all possible worlds in which it exists. I will argue that there are two conditions which must *both* be satisfied by a single set of properties in order to establish that some version of essentialism is correct. First it must be the case that objects have some of their properties necessarily. Second, the properties which constitute a thing's essence must have the capacity to explain the dispositions of that thing both to bring about change in its surroundings, and to be changed as a result of its own activity or that of the objects in its surroundings. The defence of essence via the new theory of reference succeeds in establishing that in at least some cases or for some objects, the former of these two conditions is met. However it neglects the requirement that essence be explanatory. As a result the adherents of

essentialism thus understood end up with an over-inclusive conception of what the essence of any given thing must be. This becomes problematic when they attempt to apply their somewhat simplistic conception of essence in defence of doctrines such as realism about identity or substance/concrete particulars or in theories which attempt to give us some insight into the characteristics and functioning of concrete particulars, continuants and possibly even events.

If I am correct, the modern defenders have achieved only half of what they set out to do. This is a great pity because a defensible conception of essence could be a very useful tool in metaphysics. One has to wonder if there is not some way in which a better understanding of essence cannot be built from the foundation which these philosophers have provided. It is particularly tempting to speculate about the possibility of there being some way in which we can show that some of the properties already identified as necessary can or do fulfill the explanatory role proposed for a thing's essence. It seems almost obvious that there is. We can identify essences by investigating the capacity of different candidate essences to furnish explanations of the ontological characteristics of concrete particulars - persistence through time and change, the unity which distinguishes such things as one thing rather than many and so on. In this dissertation I will argue that we can extract the essence of an object(X) from the set of its necessary properties by formulating questions concerning the persistence or the unity etc. of X and then asking which of the necessary properties of X provide the best answers to these questions. In the final chapter of this dissertation I will demonstrate how I think this can be done using questions concerning the capacities of a range of different objects to persist through a variety of different kinds of change.

In this way I hope to show that if the modern defenders of essentialism are correct about the way that proper names refer (and in chapter 1 we will see that there is good reason to believe that they are) then a complex understanding of essence, one which takes it to be more than just a collection of necessary properties, can also be defended. My project is not uncontroversial. There is a multitude of possible objections to my thesis. For the most part I will attempt to deal with those objections I have anticipated at the points at which I think they arise. There will however be one exception to this approach. In claiming that essences are fundamentally explanatory, I have to assume that we can know that a particular fact, proposition or concept is explanatory regardless of whether or not anyone is ever interested in the associated explanation. This is an unpopular way of understanding explanation, with many philosophers being inclined to believe that no fact, proposition or concept is inherently explanatory - everything depends on the use to which such things are put, i.e. explanatory power is relative to the interests, beliefs etc. of the people who explain and those to whom explanations are given. Chapter 3 of this dissertation will be dedicated to an exposition of the nature of the explanatory

role we expect the concept of essence or individual essences to play and a defence of the view that certain facts, states of affairs, concepts or propositions can objectively be said to posses greater explanatory power than others.

So chapter 1 will defend the reality of necessary properties, chapter two will indicate why the necessary properties of an object are not necessarily essential properties of those objects and chapter 3 will defend the possibility that some facts constitute better explanations of particular phenomena than others regardless of the inclinations and beliefs of an intended audience. All of this will set the stage for a defence of the claim that explanatory significance can be used to distinguish the necessary properties of an object which are essential from those which are not. Chapter 4 will be an attempt at such a defence. The final result will be a framework for a realist conception of essence which is both defensible and intuitively correct.

CHAPTER ONE

DIRECT REFERENCE AND *DE RE* NECCESSITY

In recent times, the general consensus of philosophers about essence, was that the essences of objects are nominal (as opposed to real) - i.e. that a thing's essence depends upon the way in which it is described. Thus, it was believed that things have the essences that our concepts confer on them and that the essential features of an object can change according to how we choose to see it. On such a view, as we come to understand some entity scientifically —- what is actually happening is that we are re-describing it as having a different kind of essence to the one which had previously been attributed to it. At one time we thought of atoms as essentially solid, but this way of thinking about atoms is now considered wrong, the essence of an atom is defined in terms of constituents quarks or super-strings or something. For the nominalist, this is just a redescription in a preferred discourse. This is in direct contrast to the realist understanding of essence which enjoyed prominence in antiquity and right up until the sceptical writings of philosophers such as Descartes and Locke began to dominate. As a result of nominalism, philosophers have tended to believe that there is little of interest or use to be said about the concept of essence. Realists, on the other hand take the notion to be a crucial one and, within the context of a broader realist project, are concerned to find some way of defending the real status of essences. For example, realists consider the original attribution of solidity to atoms to have been *wrong*, and has been replaced by a correct characterisation of atomic essence, which constitutes *progress*. Surprisingly, much of the recent debate about essence has occurred in the context of the philosophy of language. In this chapter I want take a more detailed look at this debate.

At the beginning of this century, Philosophy was dominated by an interest in what the structure and semantics of language could tell us about the way we think. One aspect of the preoccupation with language was the widely held belief that an analysis of language was the key to understanding whatever it is that our beliefs are *about.* It was thought that through an understanding how words refer (i.e. how they are related to the objects of experience which they pick out), we might gain some insight into the nature of the objects of reference - be they phenomena, real entities or illusions. During this period, a single theory of reference commanded almost complete authority. This was the theory, derived from the work of Gottlob Frege and Bertrand Russell, which held that reference is essentially connotative[1]. The thesis that reference is connotative amounts to the claim that that the "meanings" of

1. What I will follow Saul Kripke (See Kripke 1980, p27-29 especially footnote 5) in calling the Russell-Frege thesis is not a theory that the two philosophers shared, in fact they had very different ideas about how reference works. It is even controversial as to whether either philosopher actually held the view that will here be attributed to them. Russell, for example endorsed a referential theory of meaning for logically proper names. He did however, come to believe that most of the terms we consider to be proper names are actually disguised definite descriptions -

words can be analysed into sets of necessary and sufficient conditions (descriptions) which define concepts. These concepts originate in the minds of language users and are imposed on experience by said language users. On this view, reference is simply a matter of a given bit of experience satisfying a particular concept by meeting the necessary and sufficient conditions laid down in its definition. Recently, however, the Russell-Frege (RF) theory of reference has come under criticism and an alternative understanding of reference has been advanced[2]. On this view (often dubbed the theory of direct reference), reference is a relation which holds between objects and signs without requiring mediation by concepts. We do not have to understand a term or know what it means (presuming we understand meaning as Frege did), in order to know to what it refers or in order to use it to refer. In addition, whatever information is carried by the concept associated with a given term, will always under-determine its denotation.

Some of the proponents the direct theory of reference have claimed that much of the scepticism with which metaphysical claims have been regarded during this century, can be attributed to the ascendancy during this time, of the Russell-Frege understanding of reference. They have further claimed that the alternative understanding of reference which they have developed, offers a new and better way of approaching metaphysical issues. Many have begun to explore the implications of the theory of direct reference, in the hope that its truth could provide the foundation for a defence of certain metaphysical theories, assumptions and beliefs which have till now been considered disreputable. This is an ambitious and wide-ranging project. Philosophers engaged in it have a number of things to achieve including demonstrating that the following three claims are true :

1. The indirect (RF) theory of reference can be conclusively refuted.
2. A theory of direct reference can be clearly delineated and defended.
3. The theory of reference defended in 2 does, in fact, entails the metaphysical theses which it is claimed to support.

---

terms in which do not include any directly denoting terms - terms with whose denotations we have direct perceptual acquaintance (See Russell, 1919 and 1905 both in Martinich 1990 p217 and p204-211 respectively). Frege famously distinguished between the sense and the reference of a sign, with the sense of a sign being defined as the mode of presentation of its referent. The sense of a sign can also be thought of as what we grasp when we understand it. It is not however, a subjective thing in the way that a mental image would be. It is the same for all speakers of a language and thus not a feature of individual speakers' minds. Sense determines reference, but Frege is fairly unspecific about how this occurs or what exactly he means by this (See Frege 1892 in Geach and Black 1960 p 56-63 and Kenny 1995 p126129). The thesis referred to above as the Russell-Frege (RF) thesis provides one natural way of spelling this out but whether or not Frege himself was committed to such a view we will never be sure. What will be referred to in this dissertation as the RF thesis convenient way of picking out the set of ideas that philosophers like Kripke have taken Russell and Frege to have in common - in particular, they are assumed to have shared the view that descriptions or concepts are sufficient for determining reference.

2. Most notably promoted by Keith Donnellan, Saul Kripke and Hilary Putnam. See Steven Schwartz (ed.) 1981 Naming Necessity and Natural Kinds for a collection of classic papers on the subject.

One text in which something like this kind of progression can be seen is Saul Kripke's Naming and Necessity[3]. The first two tasks take up the bulk of the three lectures there presented. Kripke succeeds masterfully in demonstrating the weaknesses of the RF thesis and draws some interesting and important lessons from the nature of its flaws[4]. He further, constructs a plausible alternative conception of the way in which names refer. Finally, he attempts to demonstrate the implications of his conception of reference for issues of identity, materialism with respect to the mind and, most importantly for my purposes, for the possibility of *de re* necessity.

In this chapter I will set out the position defended in Naming and Necessity as an example of the way in which metaphysical claims and theses can be extracted from a Kripkean-type theory of direct reference. In section 1.1, I will give a very brief outline of the theory of reference presented in this particular work and I will consider how it can be defended. I will present two different approaches to the defence of the theory of direct reference. The first, Kripke's approach, is a *reductio ad absurdum* of the RF thesis. The second uses a thought experiment to show that it must be the case that something external to the concepts of speakers of a language is involved in the determination of the reference of its terms. As indicated above, my primary interest in the theory of direct reference for the purposes of this dissertation is its claim of to underwrite the doctrine of essentialism - the thesis (roughly) that some of the features/properties of an object are features/properties that it must have in order to be the very object that it is and not something else. My main concern in this chapter will be to show that if the theory of direct reference is correct, then it must also be the case that some objects have properties which are independent of how we choose to conceive of them and that they must have these properties necessarily. An object's possession of necessary properties is generally taken to be a necessary, if not sufficient, condition for its having an essence. So, a defence of the possibility of *de re* necessity (i.e. the metaphysical view that objects can possess their properties necessarily) will constitute a crucial step in any defence of the doctrine of essentialism. In section 1.2 I will discuss how the theory of direct reference and the considerations used to defend it, can underwrite the claim that objects have necessary properties independently of how we choose to conceive of them. In the final section of this chapter (1.3), I will explore the extent to which the theory of direct reference provides us with a reliable means of identifying all and only the necessary properties of objects. I will the question of whether the essence of an object just is the set of properties it has necessarily for chapter2.

3. Kripke 1980

4. He makes particularly interesting claims regarding the relation (or lack thereof) between necessity and *a prioricity* of statements, the possibility of *de re* necessity and the way in which we understand possible worlds - my discussion of these claims will be limited to how they are relevant to or required for his defence of essentialism and will occur largely in the final section of the paper.

1.1 Direct Reference and its Defence

1.1.1 Kripke's Position on Proper Names - The Theory of Direct Reference

The conception of reference presented in Naming and Necessity is primarily an exposition of the way in which proper names refer. Proper names are expressions which refer uniquely i.e. to one and only one object in the world. Although theories of direct reference have been developed for other terms including natural kind terms and indexicals[5], proper names are generally taken to represent the clearest, least problematic case of direct reference. In addition, since any object can be named uniquely, any metaphysical implications of a theory concerning the reference of proper names will be applicable to all so-called concrete particulars. Since we are interested in essentialism and essentialism is a theory about the nature of concrete particulars, for the purposes of this dissertation, we can (like Kripke in Naming and Necessity) restrict our attention to the way in which proper names refer.

Now, according the proponents of the RF thesis, what we ordinarily take to be proper names e.g. Sir Walter, The Sphinx etc., are most often not logically proper names at all. Instead, they are definite descriptions[6]. Their meanings are thought to be given solely by their descriptive content - what Frege calls their 'sense'[7] - which could be grasped regardless of whether they referred to anything or not. We can, for example, make sense of the names Medusa and Hercules although we do not consider either of them to have a referent. Thus according to the RF theorist, reference is a top-down process mediated by the concepts which speakers of any particular language share. Where what we usually take to be proper names do refer, as with Sir Walter or The Sphinx, their denotation is considered to be determined by the associated description. The referent will be whatever uniquely fits said description. For example, imagine that I have in my mind the concept 'The Sphinx', and that what this concept amounts to is something like the following description - "Egyptian limestone monument which is situated near the Pyramids in Gizeh, which represents a creature with the body of a lion and the head of Pharaoh Cephren and which is intended to be symbolic of royal power"[8]. If I understand this description, then I ought to be able to tell which objects fall within the extension of the name by recognising which of the objects of my experience satisfy said description. Those that do are appropriate referents of the name 'The Sphinx'. Note that there is nothing in the above description

5. For a discussion of the direct reference of natural kind terms see Putnam 1973 in Martinich 1990 p308-315. The direct reference of indexicals is considered in Kaplan 1989

6. See Russell 1919 p215-7

7. Frege 1892 in Geach and Black 1960 p57

8. Brewster's Dictionary of Phrase and Fable 1981 p1058

which entails the name can have only one referent. On the RF thesis then, the uniqueness of reference of proper names is merely a reflection of the contingent fact that in some cases of reference there is only one object which satisfies the relevant description.

How then is direct reference different ? Well, Kripke and his fellow direct reference theorists do not deny that the objects to which proper names can often, in practice, refer be uniquely described in terms of their qualitative properties. They do, however, deny that *any* qualitative description can ever:

a. infallibly (across all possible worlds) establish reference to a given individual thing

or b. capture the full meaning of a given proper name

As mentioned earlier, Kripke *et al.* believe that proper names refer directly without requiring mediation by concepts. They think that the names of individuals, places and other specific things have their meanings (or at least part of their meanings) fixed by the natures of their actual referents such that should the referent change so would the meaning of the name. They allow that names carry information, they just deny that it is this information that exclusively determines what they denote and that their denotation is irrelevant to their meaning[9]. According to Kripke, the very fact that we *can* countenance the possibility that someone, say Nelson Mandela, could look completely different to the way he does *or* could (as far as we can tell) have had a completely different history to the one he does *or* could have existed in different spatial and temporal contexts to those in which he does etc. - ie. could have been qualitatively unidentifiable to us as the person we actually identify with his name and still be that very person, tells us that there is more to the relation between a proper name and its referent than the satisfaction of a particular description.

But what then, on the Kripkean view, does the work of establishing a reference relation between a proper name and an object? Here Kripke becomes a bit vague - perhaps because there is little or nothing that can be said other than that reference is direct. The general outline he presents is something like the following : Objects are pre-conceptually apprehended *whole* by speakers of a language. On first encounter they are 'baptised' without first being analyzed into components (qualitative or otherwise). It is simply a matter of stipulation that 'this book' will be called Naming and Necessity or that 'that baby' will be called Nelson Mandela etc[10]. These names are then transmitted along chains of communication within a linguistic community, and it is via an intricate set of social and

9. See Salmon 1982 p11-7

10. The determination of the reference of indexical terms such as 'this' and 'that' is itself a difficult and controversial topic. I do not wish to get involved in that debate here, so I am going to assume that there is some way of unanimously establishing which object such terms refer to. For a detailed discussion of the debate concerning the reference of indexicals see Kaplan 1989.

linguistic conventions that a given proper name comes to refer uniquely to specific individuals for all speakers in the community. Thus, for Kripke and his fellow direct reference theorists, reference is at least partially determined and maintained by a set of events (including the presentation of objects in perception and the establishment of linguistic conventions) all of which are external to the individual speaker's conceptual or psychological states and pretty much beyond his/her control[11]. It is at this point that the we begin to see how a theory about language might gain a foothold in metaphysics. If as the direct reference theorists maintain, the reference of a proper name is fixed by facts about language, society and the world, we may be able to gain access to these facts via inquiry into the conditions under which reference occurs. Of those facts, the most interesting for our purposes will be those which concern the way things are in the world, in particular the way objects must be in order to sustain the referential relation between themselves and proper names.

1.1.2 The Defence of Direct Reference

Why should we accept the theory of direct reference? After all, the popularity of the RF thesis is not unfounded. It has been spectacularly successful in dealing with a range of important logical problems[12]. Given this success, its opponents are going to have to provide powerful arguments in order to persuade mainstream philosophers to give it up. In fact, this is precisely what they have done. The basic form that these arguments take can be understood as follows (I will hereafter refer to this argument as R):

1. The reference of a proper name is either determined solely by its sense or it is not.
2. The sense or conceptual content of a proper name is never sufficient to determine reference.

Therefore 3. The reference of a proper name is not determined solely by its sense.

Clearly, most of the work involved in such a defence consists in arguing for the truth of premise 2, so not surprisingly, it is on this that the theorists of direct reference concentrate. There are at least two possible approaches to defending premise two. On the one hand we could attempt to show that its contradiction must be either false or absurd. Alternatively we can attempt to demonstrate the

---

11. Kripke 1980 p 96-7. Note that this is not all that different from Frege's actual position. Frege does, after all, distinguish sense from mental images on the grounds that the former is objective and public whilst the latter is subjective. The Fregean 'sense' of a name however, remains mental in so far as it is not a feature of the world which it is the function of the name to depict. Fregean sense is essentially intensional.

12. Most notably accounting for the comprehensibility of claims which involve reference to apparently non-existent entities(past, future or mythical) ; and accounting for the way in which two names can have the same referent but slightly different meanings - e.g. 'Hesperus' and 'Phosphorus'; it also explains how it is that when we question the existence of an historical (or otherwise notable) figure, we may not be asking whether some given individual existed but rather whether any one individual had all of the properties generally associated with the referent of a particular name. See Discussion in Kripke 1980 p27-32.

existence of non-conceptual reference determining factors. The first option is the approach taken by Kripke in Naming and Necessity. The latter can be demonstrated using an adaptation of a thought experiment developed by Hilary Putnam. In section 1.1.2.1 I will discuss the arguments of Kripke and in the section that follows I will present a Putnamesque positive argument for direct reference.

1.1.2.1 Kripke's Reductio of the Indirect Conception of Reference

Much of the discussion in the first two lectures of Naming and Necessity is directed at exposing the flaws of the RF thesis. By this means Kripke is attempting to establish that sense is not *sufficient* for the determination of reference. His first stratagem is to assert that if its descriptive content was sufficient for establishing the reference of a proper name then the name and the description concerned ought to be synonymous - which they quite clearly are not[13]. If proper names were synonymous with particular descriptions then claims which predicate aspects of the description of the subject of the proper name ought to be analytic, necessary truths. Consider, for example, what would follow if the name Plato were synonymous with the following description : 'ancient Athenian philosopher, founder of the Academy and author of The Republic'. If this were the case then a claim such as:

i : 'Plato was the author of The Republic '

Would be equivalent to:

ii : 'The author of The Republic was the author of The Republic'.

But this is not what we mean by such a claim. More telling though, is the fact that statement 'i' is neither an analytic nor a necessary truth. It is not analytic because we can conceive of circumstances in which it is discovered that Plato did not, in fact, write The Republic. This particular work might, for example, have been written by Protagoras and only attributed to Plato. Any claim that the man to whom we do, in the actual world, refer when we use the name 'Plato' did write The Republic, must be synthetic - it would convey information of which someone who knew the man and who could refer to him successfully on repeated occasions, might well have no knowledge. That 'i' is not a necessary truth follows from a similar line of argument. There is nothing logically contradictory about a possible world in which Plato failed to write any philosophy at all and instead became an artisan or a poet. So there are possible worlds in which 'i' would be false but no possible worlds in which the same can be said for 'ii'. Consequently it cannot be necessary that someone is Plato if and only if he is the Athenian philosopher who wrote The Republic.

---

13. Here Kripke seems to be arguing specifically against Russell's claim that most proper names are in fact disguised descriptions. It is doubtful that Frege was committed to the view that proper names and their senses ought to be synonymous.

In this way it can be established that there are at least some descriptions of objects which are not synonymous with their proper names. But how far does this go toward refuting the possibility of indirect reference? Surely it remains possible that some as yet unidentified description or some subset of such descriptions of the thing named will, in fact, be synonymous with its proper name. This move faces two problems. First, it is unclear what, if anything, could serve as the components of such descriptions[14]. Second, the claim that some descriptions are more right than others sits uneasily with the conceptualist inclinations of the philosophers who favour the RF view[15]. In the second lecture of Naming and Necessity Kripke argues that no description which uniquely identifies an individual can avoid an appeal to some other proper name, and that any attempt to give descriptions which uniquely fix the reference of the additional proper names cannot fail to be circular[16].

Perhaps the requirement that proper names be synonymous with descriptions is too strong. It has been argued that all the RF theorist requires is that the reference of a proper name is exclusively fixed by a particular description[17]. Kripke acknowledges that the RF thesis can be understood this way[18]. However, even this view seems problematic. It still has the implication that the users of names identify the objects to which they apply by means of an associated description. But, as Kripke points out, speakers can succeed in referring to specific objects and people even when the only description they associate with a particular name is either false or one that does not pick out the relevant person/object uniquely. Kripke uses the fact that some people identify Einstein as 'the man who invented the atomic

---

14. Consider Plato again. Two very different kinds of property suggest themselves :

    a. Properties like that of 'being identical with Plato' - so-called haeccities

    b. Properties such as the property of 'being human'(or animal or living etc) - so-called sortal properties

The problem with sortal properties is that they are necessarily general properties. The role that they play in organising language and concepts demands that each sortal identifies a set of features which it is possible for two or more objects to share. In consequence such properties are unlikely to be able to establish the kind of *unique* reference associated with proper names. The problem with haeccities is that they make use of the proper name which we are attempting to eliminate by giving the description. In the above example, a description which cited 'being identical with Plato' would only be useful if we could already pick out Plato by some other independent means. This is an objection that can be developed against the citation, in any putatively reference-determining description, of properties which themselves involve other proper names. The invocation of such properties simply compounds the problem of explaining how reference occurs rather than resolving it.

15. I suspect that it is impossible to be a conceptualist about objects and still believe that some descriptions of them are more accurate than others . The idea that some descriptions of a referent are synonyms for its name and that other, equally applicable descriptions are not, implies that reference is *not* a purely top-down process - that something other than concepts (which are ultimately stipulative) is doing some of the work. This is precisely the view that the RF theorist would be attempting to defend himself against by making the claim that some descriptions are unique (in principle). In addition, it is interesting to note, that such a defence seems to commit the RF theorist to the idea that only some of a thing's features are crucial to its identity - this is not a position that an opponent of essentialism or an advocate of nominal essences can endorse with much confidence.

16. See Kripke 1980 p68-94. The arguments he develops are specifically directed at the cluster version of the RF thesis but given that the cluster version is supposed to be the stronger version of the RF thesis, if Kripke succeeds in showing that unique reference is impossible in this case I think it is safe to generalise his conclusions as applicable to all versions of the RF thesis.

17. See Devitt 1998 p156

18. Kripke 1980 p 53-9 - especially page 58

bomb' and Columbus as 'the man who "discovered" the Americas' as examples of cases where reference is successful despite the application of a false description [19]. The fact that neither of these two men did the things attributed to them in the above descriptions, does not entail that the names we use for them actually refer to other people. The names 'Einstein' and 'Columbus' even given the above mistaken beliefs, still pick out the very people that we have always believed them to. A case where unique reference succeeds in the face of a general description might be as follows. Someone might know Albert Camus only as 'a twentieth century, French author who won the Nobel Prize for literature'. This description does not identify Camus uniquely - it is equally applicable to a number of others e.g. Andre Gide. Nonetheless, this does not prevent such a person from using his name correctly. When he learns, for example, that Camus was born in Algeria, he does not, due to his ignorance, develop the less specific belief that some/any Nobel prize-winning French author was born in Algeria. His new belief is about the very man to which the name actually refers[20]. Thus even if we do not confuse synonymy and reference, descriptions still under-determine the reference of proper names. This is enough to establish that the RF theory of reference must be mistaken and consequently (by R), that some version of the theory of direct reference must be correct.

1.1.2.2 A Tale of Two Planets - Hilary Putnam's Twin-Earth Thought Experiment

Perhaps the most powerful defence of direct reference yet advanced comes in the form of a thought experiment. The prototype of the argument was developed by Hilary Putnam and consists of a science-fiction thought experiment in which various comparisons are set up between two separate (spatially distant) but otherwise indistinguishable planets - called Earth and Twin-Earth respectively[21]. It is important to note that *both* Earth and Twin-Earth are planets that exist in the actual world (or at least they must both exist in a single world) as opposed to being occupants of different possible worlds. Since both Earths inhabit the same world, otherwise indistinguishable objects one of which is on Earth and the other on Twin-Earth, will nevertheless be clearly identifiable as different because they will have different spatio-temporal locations and will not be made of the same matter.

What does all this have to do with the reference of proper names? Well, Putnam points out that what I have been calling the RF view rests on two crucial assumptions[22] :

---

19. *Ibid.* p p85

20. Adapted from a similar example cited in Kripke 1980 p81

21. Putnam 19 73 in Martinich 1990 p 308-315. Putnam is particularly interested in role played by external factors in the determination of the reference of natural kind terms and develops the argument specifically for these. As I have confined my discussion thus far to the reference of proper names - and intend to do so for the rest of the paper-, I will present an adaptation of the original argument which uses proper names rather than natural kind terms.

22. *Ibid.* p308-9

1. That the meaning of a term determines its extension (sameness of intension entails sameness of extension).
2. That knowing the meaning of a term is just a matter of being in a certain psychological state (as is having a belief) [23].

Thus, distinct senses ought to correspond to distinct psychological states in the subjects who grasp them and conversely, where two subjects are in the same psychological state, they ought to be referring to the same thing. On such an understanding of reference, we ought to be able to tell what someone is referring to by investigating their psychological states when using a particular name in some way. Now, imagine the situation on Twin-Earth. Imagine that on Twin-Earth, there is an individual who is qualitatively indistinguishable from the Mandela of our Earth - who looks and behaves the same and who has the same history as his earthly counterpart - but who is clearly not identical to him (as noted in the previous paragraph - he cannot be the very same person because he is in a different place and is made of different stuff). I will call this individual Twin-Mandela, but the English-speaking people on Twin-Earth refer to him simply as Mandela. In addition to all this, on Twin-Earth I have a twin whose mental biography is qualitatively identical to mine. We speak the same language and from birth, both of us have had exactly the same chain of mental events, associations and images. As I think about Mandela, my twin has exactly the same thoughts - she is in exactly the same psychological state to mine. Our brains could be interchanged and neither of us would notice. Yet, each of us is making reference to a *different* person - I to the Mandela of Earth, and my twin to Twin Mandela. Thus it cannot be true that the Fregean sense of the name 'Nelson Mandela' (which is the same for both me and my twin), determines its reference uniquely. Furthermore, the difference in reference which occurs in this case can only be accounted for by factors external to the minds of the individual speakers since our internal/psychological states are identical. Thus, the Twin-Earth thought experiment allows us to conclude that there are reference-influencing factors which are external to and independent of the concepts of the speakers who use them.

How ought we to understand these external reference-determining factors? Well, we know that that for any given proper name 'a', whatever determines the reference of 'a' is external to the conceptual scheme of any speaker who uses the name i.e. it is non-intensional. We also know that whatever is implicated in reference determination must bear some relation to the referent (X). There a two possible kinds of relation that could exist between X and the reference determining factors (RDF's).

23. See footnote 12 re. the intensionality of Fregean sense.

On the one hand an RDF could be internal to X - in other words it could be an intrinsic feature of X. Alternatively the RDF might be external to (not part of) X and the relation between it and X will be describable as a relational property of X. Either way, we can understand whatever it is that is involved in the determination of the reference of proper names to be potentially predicable of the subject for which the name stands. That is, we can take them to be non-intensional properties of the thing to which the proper name refers. Thus reference is fixed, at least in part by some of the actual properties of the object named.

The picture of reference at which we have now arrived looks very similar to a descriptive theory of reference. Reference is fixed by properties which are predicable of the referent and after all, descriptions are nothing other than a collection of predicates. The important difference between a Russell-type theory of reference and this one, is the different role played by the relevant predicates. In the case of direct reference, the predicates in virtue of which reference occurs only become relevant *after* the first instance at which a proper name is applied to an object and as need not reflect the graspable conceptual content associated with the name[24]. At the time of naming (and subsequently) we need not know the content of the relevant predicates in order to correctly use the name to refer. All that direct reference requires is that some such set of predicates exist and that their content is, in principle, discoverable. It is this requirement that makes the phenomenon of direct reference interesting to the metaphysician.

Such are the arguments for the theory of direct reference. I think they are sufficient to establish the ultimately weak claim that some version of the theory of direct reference must be correct, i.e. that some names have their reference at least partly determined by factors external to the concepts of the speakers who use them. As discussed in the introduction to this chapter, the adherents of direct reference think that if the theory of direct reference is correct, then it has important metaphysical implications. In particular they are inclined to believe that the theory can be used to establish that metaphysical essentialism is justifiable. In the remaining sections of this chapter I will attempt to set out the kind of reasoning which underwrites such claims.

1.2 The Inference from Direct Reference to Necessary Properties Kripke's Essentialism

Both Kripke and Putnam argue from the truth of the theory of direct reference to the truth of certain

24. See Linsky (1977 p110) cited in Salmon 1981 p32.

specific statements attributing particular necessary properties to particular objects and it is this aspect of their arguments that has received the most criticism. I propose to divide this process of inference into two stages. This section will be restricted to demonstrating that quite generally, given the truth of the theory of direct reference, objects which can be named must have necessary properties. I will leave until later the question of how we can work out what those properties are in specific cases. I suspect that much of the criticism that has been leveled at the attempt to infer necessary properties from the fact of direct reference applies only to what I will take to be the second stage of such arguments.

There are two clear stages in Kripke's argument in favour of the view that objects have necessary properties. In the first lecture of Naming and Necessity he sets out to eliminate the anti-*de re* modality preconceptions which were widely held at the time[25]. In the third lecture of the same work he attempts to illustrate the inference from direct reference to necessary properties[26]. In section 1.2.1 below, I will discuss the way in which Kripke dispels some widely accepted philosophical doctrines and thus clears the ground for the inference to necessary properties. In the sub-section that follows (1.2.2) I will consider Kripke's contention that proper names are rigid designators, which plays a pivotal role in the inference from direct reference to necessary properties. In the final section of this part of the chapter (1.2.3), I will reconstruct the steps of that inference.

1.2.1 Groundwork for an Essentialist Metaphysics

Kripke contends that two mistaken doctrines which have been widely accepted amongst philosophers are responsible for the disinclination of philosophers to believe that that objects can have their properties either necessarily or accidentally. First there is the assumption that whatever is necessarily true is knowable *a priori*. Since, the properties of the objects of our experience are clearly not the kinds of things that can be known *a priori*, the assimilation of the necessary to the *a priori* has led some philosophers to conclude that any claims concerning what they are like cannot be necessary. Kripke argues that *a priority* and necessity are fundamentally different notions, the former is an epistemological notion whilst necessity is a metaphysical one. He claims that the assimilation of the two is a philosophical move which requires defence. He goes on to cite a variety of examples of apparently necessary *a posteriori* truths *e.g.* 'Hesperus is phosphorous', 'The Goldbach conjecture is true' etc.[27] and even some contingent *a priori* truths e.g. stick S, by which we choose to define

25 Kripke 1980 p39f

26 *ibid.* p 110-5

27 Kripke 1980 p 36-41

'meter' is 1m long[28]. 'Hesperus is Phosphorus' is necessary because both names refer to the same object Venus, and being self-identical is necessary. It is contingent because it is an empirical fact that the 'star' that appears in the sky in the morning is, in fact, the same 'star' that appears in the sky in the evening. Likewise, the Goldbach conjecture is a mathematical law and as such, is necessarily true if it is true at all. Nevertheless, according to Kripke, there is no way of knowing independently of experience whether or not it is in fact true. Given the common sense acceptability of the *a posteriori* necessity of such statements, the onus is on Kripke's opponents to show that they are not. Until such time it is illegitimate to object to *de re* necessity on the grounds that it relies on the possibility of *a posteriori* necessary truth.

The second mistake discussed in Naming and Necessity involves the notion of 'possible worlds' in terms of which talk of modal concepts is usually couched. Some philosophers have eschewed the idea of *de re* necessity because they think that it commits anyone who invokes it to a stipulation of trans-world identity conditions for objects. Their objection is that that there are no infallible such conditions. Kripke agrees that there are no infallible trans-world identity conditions for objects, but argues that the requirement that there be such conditions reflects a mistaken understanding of what it is for something to exist in another possible world[29]. The idea of different possible worlds is supposed to capture all the counterfactual situations in which a particular object could find itself. It is a mistake however, to understand non-actual possible worlds as places at which we would look through a telescope - places in which our only means of identifying objects are mediated by our senses. Rather we should understand other possible worlds as things that we are *in* just as we are *in* the actual one. This implies that if we have ways of establishing reference that are not entirely dependent on the on the qualitative properties we associate with the things to which we refer in the actual world, then we would have recourse to such means in all possible worlds. So we can know whether the claim that 'X has property F in all possible worlds (ie. necessarily)' is true or false, even when we cannot specify some qualitative criterion by means of which to judge that the object which is X in the actual world is identical to thee object we call X in some non-actual possible world. In consequence, objections to *de re* modality based on the unavailability such of trans-world identity conditions do not go through.

If Kripke is correct on the above two points then he has cleared the path for the inference from direct reference to the claim that objects have some of their properties necessarily. All that remains is to spell out the steps of the actual deduction. To recap, thus far we have established that the sense of a

28. *Ibid.* p32-9 an p 56-7

29. *Ibid.* p 43

proper name (its conceptual content), is not *sufficient* to determine its reference[30]. Since it is not, there must be some non-conceptual factor which is, at least partially, involved in the establishment of any reference relation between an object and its proper name. I argued that we can think of the non-conceptual factors involved in reference determination as non-intensional properties of the referent[31]. So, some non-intensional properties of objects are involved in the establishment of the reference relation between them and their proper names. In order to get from here to the claim that objects have some of their properties necessarily, we require Kripke's notion of *rigid designation*.

1.2.2 Proper Names as Rigid Designators

The notions of rigid and non-rigid designation were developed to capture the different ways in which terms can behave in counterfactual situations (or across possible worlds). It seems to be a fact about linguistic terms - names and descriptions - that they either apply to one specific object in different counterfactual situations (*across possible worlds)* or they apply to different objects in such counterfactuals (*in different possible worlds)*. 'Bill Clinton' is an example of a name that seems to denote the same object in all possible worlds in which it denotes anything at all. Nothing can be Bill Clinton except the very object to which we refer to when we use the name in the actual world. When we make the following claim :

iii : 'Bill Clinton could have been a woman'

we are picking out the same object that is actually a man, and we cannot be picking out anything else. This kind of claim should be contrasted with claims such as:

iv : 'The President of the USA in 1999 could have been a woman'

'The President of the USA in 1999' is a description which applies to different objects in different possible worlds. There is a possible world (or a conceivable counterfactual situation) in which Hilary Clinton was elected to the presidency. In this possible world statement iv would be true. By contrast, there is no possible world in which statement iii can be true because there is no possible world in which the name 'Bill Clinton' designates anything other than the person it designates in the actual world.

If a term picks out the same individual in every possible world in which it refers at all, then it is what

30. See sections 1.1.2.1 and 1.1.2.2 above.

31. See second last paragraph of section 1.1.2.1. above.

Kripke calls a *Rigid Designator*. Note that rigid designation does not require that the referent *exist* in all possible worlds. All that is being asserted is that in possible worlds in which the specific referent of a rigidly designating name does exist, it is always designated by that name and in possible worlds in which said referent does not exist, the name designates nothing. By contrast with rigidly designating names, a *Non-Rigid or Accidental Designator* such as 'The President of the USA in 1999' refers to different objects in different possible worlds - Bill Clinton in this world, Hilary in another and Michael Jackson in yet another.

Kripke contends that all proper names are rigid designators[32]. How might he defend this claim? Well, the fact is that we can and do speculate about what could and could not have happened to a particular individual or object. When we indulge in such speculation, we hold the object in question constant and imagine it to have been different in various ways. We consider what could and what could not have been true of that very object in other possible worlds. It follows that the subject term in the resultant counterfactuals must be a rigid designator(a). Whenever a proper name occurs in the subject position of a counterfactual, the truth value of the resultant proposition depends on the actual nature of the thing designated by the name in the actual world. In other words, all such counterfactuals involve speculation about what could and could not have happened to a particular individual or object(b). Therefore (a and b taken together imply that) all proper names are rigid designators.

1.2.3 Reference Determination and Necessary Properties

Now let us return to the argument from direct reference for the existence of necessary properties. Based on the arguments of the previous sections we are now entitled to make the following claims:

1. All proper names are rigid designators (from section 1.2.2)

2. The reference of a proper name is either wholly or partially determined by non-intensional properties of the referent (from section 1.1.2)

Now, another way of saying that a name refers to the same object in all possible worlds in which it refers to anything, is to say that it refers to that object in all possible worlds *in which said object exists*.

So 3. A proper name refers to a specific object in all possible worlds in which that object exists.

32. Kripke, 1980 p 48-9.

Thus 4. Whatever determines the reference of a proper name must be constant across all possible worlds in which our name refers.

Premises 2 and 4 taken together entail:

5. Some properties of the referent are constant in all possible worlds in which the referent exists.

Properties that an object has in all possible in which it exists are necessary properties of that object. So, we can infer that the properties of an object which are involved in directly determining the reference of its proper name, are properties that the object has necessarily. Thus we have derived the conclusion that objects which can be the referents of proper names must have necessary properties from the theory of direct reference and the fact that proper names are rigid designators.

The above argument might look neat but the conclusion is not uncontroversial. Many philosophers are appropriately suspicious of the claim of what is essentially a semantic thesis to entail metaphysical principles. They have sought the errors in the above kind of reasoning in the kinds of consideration which were dismissed in section 1.2.1 above. They have argued for instance that because it assumes rigid designation, the above argument relies on a particular (apparently implausible) understanding of the notion of possible worlds. The inference as I have presented it above, would not convince anyone who accepts David Lewis' realism with respect to possible worlds which has the consequence that individual objects are confined to a single possible world. Every object, according to, Lewis exists in a single possible world and has similar but non-identical counterparts in other possible worlds[33]. On such a view rigid designation is not a notion that even makes sense. Thus for the Lewisean realist the above derivation does not get off the ground. This is an interesting objection but Lewis' conception of possible worlds is counterintuitive. In addition it seems to have its justification in the requirement that if we are to talk sensibly about things being identical in different possible worlds then there must exist some specifiable trans-world identity conditions for the relevant things. If Kripke is correct, and the requirement of such conditions is avoidable, we have little reason to accept realism about possible worlds. And given that a non-counterintuitive understanding of the possible worlds mechanism is available (i.e. Kripke's[34]), it at least seems reasonable to adopt this alternative view until such time as Lewis version is widely acknowledged to be correct. So as things stand I think we are entitled to the Kripkean construal of modality and thus to his notion of rigid designation. The objection from realism

33. Lewis 1973 reprinted in Loux 1979 p182-9.

34. See section 1.2.1 above

about possible worlds fails and the inference set out above goes through.

A more serious objection has been developed by Nathan Salmon, who argues that arguments like those of Kripke and Putnam beg the question. According to Salmon all such arguments (unless they are trivial) gain their metaphysical import by including premises which assume *de re* necessity from the start[35]. Is this the case for the argument outlined above? If there were any hidden essentialist assumptions lurking within the argument, it would have to be in the form of a presumption made by either premise 1 or premise 2. Let us examine each in turn. Does the claim that proper names designate rigidly assume the existence of necessary properties? Salmon himself is satisfied that it does not. He takes the fact that proper names are rigid designators to be a simple consequence of the theory of direct reference[36]. I think he is correct about this. Provided we accept the Kripkean construal of possible worlds and assuming we take direct reference to be possible, the exposition of rigid designators and the argument that proper names always designate rigidly in section 1.2.2 above, does not rely on anything to do with the properties of the objects designated. How about premise 2? Does the Twin-Earth thought experiment contain any hidden assumptions of necessity? I pointed out at the beginning of section 1.1.2.2 that Twin-Earth is not the kind of hypothetical entity that inhabits another *possible* world and that consequently, the argument does not rely on differences of modality. The crucial judgement that Mandela and Twin-Mandela are different and thus that the referents on different Earths, of the name 'Mandela' are different whilst the intension of the name is identical for myself and my twin, does not assume that either referent has any of its properties necessarily. Hence it cannot be accused of begging the question against the anti-essentialist. So neither premise 1 nor premise 2 of my reconstruction of the inference to necessary properties assumes that objects have necessary properties. Therefore I conclude that the charge leveled by Salmon at Kripke and Putnam, whilst it is effective against some of their arguments, can be avoided.

So much for the inference to necessary properties. Hereafter I will assume that it has been established that objects to which we can refer using proper names do have necessary properties. At this point it is not clear which properties of objects are likely to be necessary or how we might go about identifying them. This will be addressed in the following section.

1.3 Which Properties are Necessary?

35. See Salmon 1981 p182-92 for an argument against the Putnam-style derivation of necessary properties and p196-214 for arguments against the Kripkean attempts to defend the inference.

36. Salmon 1981 p32-41

We now know that objects have necessary properties. We also know that all the properties of an object involved in the determination of the reference of its proper name, are necessary properties of that object. We know this because, proper names are rigid designators so whatever determines their reference must be constant across all possible worlds. We do not yet know how to work out which of a given object's properties are the ones it has necessarily. However, we ought to be able to do so by using a modification of the Twin-Earth scenario which will allow us to identify the reference determining properties of the object in which we are interested.

It was important in our original version of the Twin-Earth thought experiment that there was no qualitative difference between the Earth referent of a given proper name and its inter-planetary twin. The two were only distinguishable as different referents *because* they were on different planets. However, because we now simply want to know which specific properties of some object are reference determining, we can relax this requirement a little. Putnam's original version of the thought experiment allows that the substance we call water and the apparently identical stuff on Twin-Earth differ in chemical composition. Putnam asks whether in a time when we are ignorant of such things as chemical compositions, I and my twin when referring to the colourless, drinkable liquid with which we are familiar, would be referring to different substances. He concludes that we would[37]. Thus, it would appear that specific chemical compositions are reference determining and consequently necessary properties of the substances to which they belong. Unfortunately we cannot construct a parallel argument involving the referent of a proper name because the very difference that allowed us to distinguish specific stuffs on Earth and Twin-Earth as different in the thought experiment above and different objects in section 1.2.2, will ensure that the referents of any given proper name on Earth and Twin-Earth will *always* be different whilst the sense of the name is the same for both myself and my twin.

There is however another alternative. Now that we have established that objects do have necessary properties, we can without fear of begging the question, invoke the apparatus of possible worlds. If an object has a property necessarily, then by the definition of *de re* necessity[38], any object which fails to instantiate that property in any possible world must be a different object. Consider my pet Bengal Tiger who is called Thor. Now consider, an object which looks and behaves exactly like Thor in another possible world (w1), and which I, in w1, take to be my pet tiger and refer to as Thor, but which is not in fact, an animal at all but a god. My mental states, cognitive processes etc when I refer to Thor are identical in *w1* and *wa*. Whether or not these identical thoughts are accompanied by an

37. Putnam 1973 in Martinich 1990 p309.

38. P is a necessary property of X if and only if X has P in all possible worlds in which X exists

identity of referents will depend on whether or not the solitary property with respect to which the referents of *w1* and *wa* differ (in this case being an animal as opposed to a god) is a necessary property of the actual Thor or not. If 'being an animal' is a necessary property of Thor in *wa*, then the objects to which I refer when I use the name 'Thor' in *wa* and *w1* will be different objects. If, on the other hand, 'being an animal' is not a necessary property of Thor in *wa*, then 'Thor' as used by me in both possible worlds will refer to a single object. How does this help us to identify the necessary properties of Thor? Well, if we know that the Thor's of *w1* and *wa* are identical, then it follows that the particular characteristic with respect to which they differ, is a non-necessary property of the actual Thor. Conversely, if we know that Thors of the different possible worlds are non-identical, then we also know that the unique property with respect to which they differ is a necessary property of actual Thor[39]. So provided we have reliable trans-world identity criteria which do not include the property which we are testing for necessity, we now have a means of identifying the necessary properties of objects.

However, as we have seen, the requirement of reliable trans-world identity conditions is a considerable prerequisite. The possibility of such conditions has been the subject of extensive debate amongst philosophers interested in issues of identity and necessity[40]. Once again the sensible option seems to be to accept Kripke's claim that we do not have to have explicit trans-world identity criteria in order to know when something is and is not the same object in different possible worlds. He thinks that the belief that we do, arises as a result of a mistaken manner of understanding the apparatus of possible worlds. To recap, he has argued that we do not discover the way things are in other possible worlds as if we were looking through a telescope at them. On the contrary, we stipulate other possible worlds and in doing so we can *make it true* that X in *wa* is identical to X in *w1*. We can then go on to ask given the fact of trans-world identity whether X in *w1* could fail to have the particular property in which we are interested. This is exactly how he proceeds to identify necessary properties in Naming and Necessity[41].

1.4 Summary

The arguments of this chapter have been fairly wide-ranging and dense. To conclude I simply want to

---

39. Notice, that we can also use this scenario to show that *all* the necessary properties of an object are crucial to the establishment of the reference relation between it and its proper name. We have established that whenever the property which differs from one object to the other across possible worlds is necessary, the objects concerned will be non-identical. But whenever this is the case, we will have a situation in which whether I use the relevant proper name in *w1* or *wa*, my mental states in each case will be indistinguishable and yet I will succeed in referring to different objects on each of the two possible worlds. In other words, *whenever* the property in terms of which our inhabitants of different possible worlds differ is necessary, a single intension will be associated with different referents. And wherever this is the case and those different referents only differ with respect to a single property, we can conclude that the property concerned must be reference determining.

40. See Lewis e1973 and 1986

41. Kripke 1980 p 110-5

summarise the position that I think we have reached and how we have reached it. The guiding aim of this chapter has been to explore the possibility that the theory of direct reference can be used to justify some version of essentialism in metaphysics. In the first section (1.1) of this chapter it was argued that some version of the direct reference theory must be correct. I went on in the following section (2.2), to show how, using the theory of direct reference and the fact that all proper names designate rigidly, we could infer that objects to which we can refer using proper names must have necessary properties. I considered two possible objections to this inference and concluded that neither of them have much force against the inference as here presented. In the final section I considered a way in which we might move from the general conclusion that objects have necessary properties to more specific claims regarding the nature of the necessary properties of specific objects. In chapter two, I will take up a similar theme from a slightly different angle. I will look at the way in which Kripke identifies specific necessary properties and consider whether the resultant sets of properties can appropriately be called essences.

CHAPTER 2

NECESSARY PROPERTIES AND ESSENTIALISM

In the previous chapter it was established via the theory of direct reference that objects have necessary properties. In the final section of that chapter I developed a kind of 'intuition pump'[1] which provides a means of identifying all and only the necessary properties of any given object. Modern adherents of the view that objects have necessary properties have developed a range of different methodologies and arguments by means of which to determine which properties of given objects are necessary to them[2] [3]. When they use their various methods for generating necessary properties, most such theorists take themselves to be identifying the *essential* properties of the relevant objects. In this chapter, I want to question that assumption. I will do so by examining the methodology used by Kripke in Naming and Necessity for identifying what he takes to be the essential properties of objects. I will pay special attention to the kinds of property it generates. Many of the properties that he identifies as essential appear dubious, arbitrary, trivial and sometimes just plain weird. Part of the strangeness of these apparently essential properties may well be accounted for by flaws within the methodology[4]. However part of the problem, or so I will argue, is that their being necessary does not fully capture what we mean when we talk about properties being essential. I will argue that whilst Kripke and his fellow direct reference theorists have provided the basis for a defence of essentialism, they have failed to complete such a defence. I will suggest that this failure is the result of an assimilation of the essential to the necessary - an act which profoundly misconstrues the metaphysical significance of the concept of essence.

The structure of this chapter will be as follows. In section 2.1, I will set out the Kripkean methodology for identifying the essences of objects. Section 2.2 will be devoted to consideration of the properties identified via this methodology as essential. In section 2.2.1, I will discuss the kinds of properties which Kripke in fact considers to be essential to particular objects and in section 2.2.2, I will demonstrate that Kripke's methodology commits him to claiming that all objects have some rather odd essential properties. In the final section of this chapter I will attempt to diagnose at least one of the problems inherent in the picture of essences that results. In the chapters that follow I will go on to suggest a

1. Term borrowed from Dennett, 1984 p 12.

2 Mine as indicated is very similar to the method used by Kripke . See Kripke 1971.

3 Interestingly enough, it is at these methodologies that anti-essentialist arguments are usually directed. That many of the counter-arguments are cogent, seems to indicate that something is going wrong somewhere in the inference from direct reference to specific necessary properties. Furthermore, it looks as if the problem must be at the level of the move toward identifying specific necessary properties.

4 Flaws such as those urged by Salmon in Salmon 1981 p 196-211.

way in which the essences of objects can be extricated from the set of properties that they have necessarily (identified as per section 2.3).

2.1 The Kripkean Methodology for Identifying Essences

In the third lecture of Naming and Necessity, Kripke demonstrates how he thinks we should go about identifying and enumerating those properties of objects which are necessary properties[5]. He does this for both objects - like individual tables and tigers, and for natural kinds - such as water, gold and various species. I will concentrate on what he says about objects. He never explicitly outlines a methodology but one can be gleaned from his discussion of a wide range of examples. He begins by picking our something (X) that is clearly an appropriate referent of a proper name (N) - Hesperus, Richard Nixon etc. and some property (F) associated with that object - 'being present in the sky in the morning', 'having been a president of the USA' etc. He then poses the counterfactual question 'Could X - the very thing to which we actually refer when we use N - have failed to have F in any possible world? or without the possible worlds locution, ' Is it conceivable that X could lack F and still be the referent of N?'. If this is not possible/conceivable then the property F must be one which X has in all possible worlds in which X exists i.e. necessarily[6]. Thus Kripke teases out those ways in which an object could not conceivably be different and takes these to be its essential properties. Clearly, those properties which the object need not have in all possible worlds in which it exists are its non- essential or accidental properties.

Let's take a look at an example. Consider Nelson Mandela. It is conceivable that the man to whom this name refers was not the first president of democratic South Africa or that he was not ever married to Winnie Madikezela. So, 'being the first president' or 'having been married to Winnie', are not necessary features of our actual ex-president. We can conceive of any particular man as having had a different job, historical role or spouse. We can even conceive of Mandela (or any other human being) as having totally different qualitative characteristics including things like skin colour, height and maybe even gender. So none of these things on Kripke's view can be necessary properties of the objects which instantiate them. This excludes a great deal from the realm of the essential. Is there anything left which can be an essential property of Mandela's? We will see in subsequent sections of this chapter that the answer to this question is, in fact, that too much remains within the realm of the essential for the notion of essence to be very useful. However, for the monument consider the

5 Kripke 1980 p110-5. Note that wherever I mention necessary properties in connection with what Kripke says, he is invariably talking about essential properties.

6 Kripke goes to great pains to point out that the questions he is asking concern possibility and necessity rather than change. He is interested in those features of a thing that would not interfere with the rigid designation of a proper name across possible worlds, rather than those features of a spatio-temporally continuous thing which could not change. See footnote 57 in Kripke 1980, p144.

property of being human. A possible world in which *the* very man we know as Mandela, is not a human being *is* inconceivable. He could not be the very thing he is and not be human. Thus, being a member of the species 'human being' appears to be a necessary feature - by Kripke's methodology, an essential property - of Madiba's.

2.2 The Properties a Thing must Have in all Possible Worlds in which it Exists

Most of what Kripke has to say about essentialism comes in the form of a discussion of various examples. He never explicitly *defines* essence or essential properties. However, from the way in which he approaches the identification of essential properties, we can infer that for him the definition of an essential property must be something like the following:

> D : P is an essential property of X if and only if X has P in all possible worlds in which X exists.

Note that exactly the same definition could be given for 'necessary property'. Thus it would appear that for Kripke there is no difference between the two kinds of property.

Note also that in order to know if a certain property (P) is essential to an object (X) we have to know that P is a property of X to begin with. Remember that according to the theory of direct reference we can identify objects uniquely despite not being able to describe them accurately. How then do we discover the properties of an object which may be its essential properties? The important thing to remember here is that essences, on the theory derived from direct reference, are *a posteriori* creatures. Like any other property had by the objects of our experience we discover them by scientific means. Once a property of a thing has been discovered a further conceptual step is required in order to establish whether it is an essential or an accidental property of the relevant object. Imagine we discover two things about water molecules - first that they are made up of two atoms of hydrogen and one of oxygen and second that each molecule has the same total number of electrons as an atom of neon. We ought to be able to work out whether either of these properties is an essential feature of a water molecule by asking ourselves whether water molecules have either property in all possible worlds in which they exist. Interestingly, in a case like this, the required thought experiment would probably produce different results depending on whether it was performed by a lay-man or by a chemist. The former would most likely conclude that the chemical composition was essential but the similarity to neon was not. A chemist might however reason that since the number of electrons possessed by specific molecules and atoms is the same in all possible worlds in which they exist, the similarity between water molecules and atoms of neon will hold across all possible worlds and thus is an essential feature of water molecules.

Already the Kripkean definition of essence begins to look as if it has contradictory implications. However, before launching into a diagnosis if its flaws, lets us gather further ammunition by taking a look at some of the properties which Kripke thinks do satisfy his definition and at some of the rather strange essential properties to which the above definition commits him.

2.2.1 Essential Properties à la Kripke

By a process like the one described above for Mandela, Kripke identifies a range of different properties which he thinks certain objects have necessarily. Kripke himself does not really spell out any underlying principles in terms of which the properties he identifies could be organised[7]. However, in commenting on his work, Charlotte Witt usefully classifies the properties Kripke discusses into three basic types[8]. She calls them properties of origin, properties of substance and sortal properties. I will discuss each of these in turn.

2.2.1.1 Properties of Origin

According to Kripke, if we think about a particular person - say Mandela - we cannot conceive of his being the very person he is, without his having had precisely the parents that he did. According to Kripke, when we ask "Could the very person we know as Mandela have been born of the parents of Walter Sisulu?". The answer, apparently, has to be "no". And, it has to be "no" because this is not what *in fact* happened. This is one example of the supposedly more general truth that all objects have their origins in particular bits of matter essentially. I would not be me (would not exist), according to Kripke if nothing had had ever arisen from the actual gametes from which I, in fact, did arise. The same applies for entities far less complex than human beings. To illustrate this, Kripke asks whether the very lectern at which he stood whilst giving the lecture on which Identity and Necessity is based, could have originated from a different block of wood from the one it fact did, and concludes that it could not[9]. He says that whilst we can conceive of a lectern which is qualitatively indistinguishable from that one but which has as its origin a different piece wood, we cannot think of such a lectern as being identical to the one in question. It is, in the end, only similar[10]. Note that we can imagine the very block of wood from which our lectern is, in fact, made as having quite different superficial features to those it has - it might have been a different colour, it might have come from an older tree

---

7 Kripke 1980 p114 footnote 57, does talk about a principle of the essentially of origin and a similar principle of the essentiality of substance but his brief comments do not amount to a systematic analysis of all the properties he takes to be essential.

8. See Witt 1981 p183.

9. Kripke 1971 in Schwartz 1973 p86.

10. Kripke offers 'something like proof' of this principle See Kripke 1980 p 114 footnote 56. Salmon 1981 p196-211 directs his criticism of Kripkean essentialism at precisely this argument. Arguing, I think effectively, that it only works if we make an essentialist assumption at the outset (see principle V on p 206).

etc. So what is important here is not the specific 'nature' or 'character' of the particular block of wood but rather its being *the* unique bit of matter from which the actual lectern originated.

2.2.1.2 Properties of substance

The second property-type that Kripke thinks will always be necessary to the objects in which they inhere is comprised of properties which specify the *kind* of stuff or matter of which specific objects are composed[11]. Thus, it is not just being made of 'that particular block of wood' that is an essential feature of our lectern. It is also essential that the lectern be made of wood rather than hardened glue or ice. Kripke asks us to consider whether this very lectern, could from the beginning of its existence have been made of frozen water from the Thames. He says:

> One has a considerable feeling that it could not, though in fact one certainly could have made a lectern of water from the Thames, frozen it into ice by some process and put it right there in place of this thing. If one had done so, one would have made of course a different object. It would not have been this very lectern and so one would not have a case in which this very lectern was made of ice or was made of water from the Thames. So it would seem that....in any counterfactual situation in which we would say that this lectern existed at all, we would have to say also that it was not made of water from the Thames frozen into ice[12].

In a similar fashion we could presumably argue that it is a necessary property of Mandela that he is made of protoplasm rather than- say - platinum.

2.2.1.3 Sortal properties

Sortal properties are those which indicate the *kind* to which a given object belongs - properties like 'being human' or 'being carbon' and even properties like 'being inanimate'. They tell us something about the group membership of objects in what appears to be a systematic universe. The classic examples of kinds are the so-called 'natural kinds', kinds or type classifications which appear to reflect the way in which nature is organised. But kind membership is a genuine property of any object which can be thought of a token of some or other type of thing. Thus, the chair on which I am sitting can be thought of as belonging to the kind 'chair' or even to the more general kind 'furniture' etc. Kripke does not distinguish between natural and non-natural kinds when discussing such properties. According to him, 'being a lectern' is a necessary feature of *this* lectern, just as 'being human' is a necessary feature of Mandela. Just as there can be no possible world in which Mandela - the very man - is an angel, there can be no possible world in which our lectern - the very one at which Kripke stood during his lecture - is a raft or a flower pot. The underlying intuition here is that certain kind memberships are mutually exclusive. No individual can be both human and angelic and by analogy nothing can be a lectern, a raft *and* a flowerpot simultaneously.

---

11. Kripke 1980 p114 footnote 57.

12. Kripke 1971 reprinted in 1977 in Schwartz p 87.

2.2.2 The other Essential Properties to which Kripke is Committed

The three kinds of property set out above cover all of the essential properties to which Kripke explicitly commits himself. However, I don't think that he means to be doing anything more with his examples than demonstrate the requisite reasoning. The same procedure could be applied to any property at all. If we do this, we find that there are a number of the properties which turn out to be essential features of objects that we do not ordinarily associate with them at all and which we certainly do not think of as essential to them. For example, Kripke's claims about substance can be expanded. If our lectern is essentially made of wood then it also has the following properties essentially[13]:

1. not being made of ice or not being made of butter
2. not being anything which is necessarily made of some other kind of 'stuff'.

Similarly, the claim that this lectern is essentially a lectern has some interesting entailments. Not only does it become essential that this lectern is 'not a flower-pot', it also becomes an essential feature of our lectern, that it is :

3. not The Taj Mahal
4. not the number B
5. not Pegasus

Any of these properties satisfy the definition of essential properties given at the beginning of section 2.2, as will properties such as :

6. being such that 2+2=4
7. being either red or not red
8. being self-identical

All of properties 1-8 are properties which can accurately be attributed to Kripke's lectern and all of them satisfy his definition (D) of an essential property. However, whilst they are all undeniably properties which we are happy to allow that Kripke's lectern has in all possible worlds in which it exists, they nonetheless seem somehow irrelevant in a way that 'being a lectern' and 'being made of wood' are not. Most of us would hesitate to call such properties essential to the lectern. So if we accept Kripke's methodology, we seem to generate a plethora of essential properties which tell us little or nothing about the objects we are talking about - the potential for insight into which, is I will argue, the main motivation of essentialist metaphysics in the first place.

13. Discussed in both Kripke 1980 p113 and in Kripke 1971 in Schwartz 1977 p87.

2.3 Kripke's Oversight

I think Kripke's attempt to specify the essential properties for certain objects is fraught with problems. Some rather dubious metaphysical assumptions are required in order to get to the point at which it is possible to claim the origin, the substance of which it is made up and the (non-natural) kind to which it belongs, are all essential properties a particular lectern[14]. However, for the purposes of this dissertation, I want to concentrate on a different kind of weakness to be found in Kripke's treatment of essential properties. I think Kripke makes an important error by carelessly assimilating the necessary to the essential. This error is most clearly demonstrated by the capacity of the 'possible worlds' machine to crank out a variety of strange looking essential properties for most ordinary objects. I want to suggest that the common-sense intuition that either such properties are not essential or 'essence' cannot mean what we tend to think it does, is correct. I think that there is an important divergence between what Kripke calls essence and the way in which the notion is traditionally used. If I am right, then the Kripkean understanding of essence ought, at the very least, to be regarded with circumspection by any philosopher who might think of employing the concept in the service of other arguments or explanations.

My primary criticism of Kripke is that the class of properties he identifies as essential is over-inclusive. I want to suggest that a list of the essential properties of any given object should not include the type of property to which (in section 2.2.2) we found Kripke to be committed i.e. properties which provide a negative characterisation of the object in question and properties which are possessed necessarily by all objects which exist at all. Nor ought it to include those properties which a given object has essentially because they are consequences of its other essential properties. Finally a stipulation of the essence of an object should not contain trivial properties, such as being self-identical and (I will claim) being made of the particular kind of matter of which the relevant object is, in fact, made and having the specific origin which it, in fact, has (see section 2.2.1). My objection to the incorporation of any of these kinds of property into the supposed essences of objects is that they are derivative and essences are not. Such properties are properties which objects have *because* of what they are, whereas essences are supposed to specify what the objects that have them are fundamentally or independently of anything

14 Once again see Salmon 1981 p196-211. One such dubious assumption is that objects can be non-circularly distinguished from one another and from non-objects like properties, processes and events without invoking the concept of essence. I think that this is principle which underlies the kind of objection developed Salmon who argues that the metaphysical assumptions made by Kripke at the outset of his argument for the essentiality of substance makes those arguments circular. Kripke does gesture in the direction of a theory of substance by commenting on the inadequacy of understanding things either as bundles of properties or as underlying propertyless substrata (p52). However his comments about not wanting privilege some entities over others tend to indicate that as Witt has claimed, 'Kripke holds that essences of material individuals can be determined independently of the question of the metaphysical constitution of individuals such as persons, artifacts and organisms' (Witt 1980 p 186) . Other assumptions, on which I think Kripke relies, include: the supposition that a thing just is its matter/substance and the conviction that there are such things as non-natural kinds to which things like lecterns can belong.

other factors. In what follows, I will try to spell out in a fairly in more detail and in a fairly untechnical way what is wrong with considering each of these kinds of properties to be essential. In the chapters that follow this one I will move on to defending an alternative view which applies some of the insights drawn from this section.

2.3.1 Consequences of Necessary Truths

Consider necessary properties of mine such as 'being such that the law of non-contradiction holds' and 'being either purple or not purple'. We know these are necessary properties of mine because they reflect states of affairs the negation of which would be a logical contradiction. There are no possible worlds in which it is not the case that the law of non-contradiction holds or in which all objects are not either purple or not purple. So whatever exists, including me, *must* be incompatible with these impossible states of affairs. The problem associated with calling such necessary properties essential is that they just do not seem capable of fulfilling the role which essences fulfill in metaphysical theories. Since the properties I have in virtue of not contradicting any necessary truths are properties I share with everything else that exists, it is hard to see how they could serve to *distinguish* me from anything else that exists. They might be useful in distinguishing me from Pegasus or the Abominable Snowman, however the same thing can be more perspicuously and economically done simply by noting that I exist whilst Pegasus and the Abominable Snowman don't. Neither do such properties explain how I come to have any other properties, or furnish any insight into me, my superficial properties, my causal behaviour or my dispositions to behaviour in particular environments etc.

2.3.2 Negative Properties

From a common-sense perspective, it is difficult to see how properties like 'not being made of ice' can count as essential properties of any object, for example me. First of all, the properties which specify what I am necessarily not, are properties I share with a wide variety of other objects including Kripke's lectern and the Eiffel Tower. However, since we tend to use the set of essential properties that characterises an object to individuate it and properties which I share with a such a diverse range of other objects seem unlikely to be individuative of me. They might distinguish me from certain other things - those that are made of ice in the case of the above example. However this hardly amounts to individuation. Further we expect knowledge of a thing's essential properties to provide insight into its dispositions (or the lack thereof) to behaviour in the presence of specific environmental stimuli (perhaps these are better called forces in the inanimate world). It is hard to see how such insight can be gained from a characterisation of what something is not (except in so far as this allows for extrapolation to a characterisation of what the relevant object is). Think again of me and the necessary property I have of 'not being made of ice'. This property, whilst it might explain why I don't melt at 0°C and possibly

a number of other things about me, it would leave the vast majority of my identifying characteristics, causal behaviour etc. unaccounted for. I believe the failure to decisively individuate or explain stems from the fact that such properties are derivative. They depend for their existence on other necessary properties. Finally, as if these considerations were not enough to make us dubious of negative properties, there are just too many properties of this type for a notion of essence that includes them to be workable.

2.3.3 Consequences of Other Essential Properties

It seems obvious that the necessary properties of objects can stand in a relation of dependence to one another. Examples of such properties abound. Water is necessarily H_2O and *in consequence* its molecules necessarily have the same number of electrons as Neon atoms; Hesperus is necessarily Phosphorus *because* it is necessarily self-identical and finally Socrates is necessarily mortal *because* he is necessarily animate. Thus there are numerous necessary properties for which the following condition holds:

> R : Object x has the property expressed by predicate F (PF) only if it has the property associated with predicate G (PG).

In other words, x's being G would be a cause or a precondition of its being F. Now, where this is the case we can describe x's identity and explain its behaviour without reference to F. This should be enough to make us doubt that *both* PF and PG need be included in a list of the essential properties of the objects concerned.

2.3.4 Trivial Properties

It has long been acknowledged that properties like being self-identical are necessary to all objects - but because of their indexical quality they do not amount to the same property for any two objects. It has even been claimed that such properties can provide the basis for an account of the unique essences of individual objects. However, as Kripke himself would probably argue and as most philosophers are happy to acknowledge, a notion of essence which included properties such as these would defeat the object of talking about essence in the first place. The existence of such properties and the fact that whatever has them, has them essentially provides no insight into the objects concerned. I think that the charge of triviality can be extended to two of the types of property that Kripke discusses in Naming and Necessity[15] - namely so-called properties of substance and properties of origin. Let us consider properties of substance. Aristotle recognised that an object e.g. a table, a mountain, a person, is also always a bit/piece/blob of the stuff of which it happens to be made[16]. Given this fact, it is extremely

15 Kripke 1980 p110-5. See also sections 2.2.1 and 2.2..2 above.

16 Aristotle Physics II,1 193a 9-11.

important to be clear about what it is to which we are attributing necessary or essential properties. Consider the ubiquitous lectern of previous sections. This object can accurately be described as a piece of wood. What exactly are we saying when we say that our lectern is necessarily made of wood? Is it the lectern *qua* lectern to which we are attributing this necessary property or are we making a claim about the lectern *qua* the piece of wood which it also just happens to be? If the latter is the case, then all we are saying when we attribute the necessary property of being wooden to our lectern is that a piece of wood is necessarily wooden -as trivial a claim as any. We can generalise the conclusion reached in the lectern's case as follows):

> C: Being made of substance S is not a necessary property of object X in any interesting sense, but rather a trivially necessary property of S (the stuff) itself of which X contingently happens to have been made.

A similar line of reasoning should show that so-called properties of origin are equally trivial and for the same reasons. Take for example the claim that a given chair necessarily has its origin in the very blob of plastic from which it actually originated. All that we are saying when we make such a claim is that this blob of plastic is necessarily itself. Thus, like properties such as 'being self-identical', properties of substance and origin , whilst necessary properties of the objects that have them, ought not to be included in the sets of properties which we take to be essential to those objects because they are trivial.

2.4 Essence as Essentially Explanatory

At the beginning of the previous section I suggested that ultimately what was wrong with including properties which fall into any of the types discussed above was that such properties are derivative and essences are not. Throughout the previous section (2.3) I also relied on the intuitively appealing claim that necessary properties which do not or cannot play a role either in effectively individuating the objects to which they are attributed or in explaining the ways in which those objects are disposed to behave in particular circumstances, should not be included in the set of properties which we take to constitute a thing's essence. As we will see the two objections are not unrelated. It is because essences are fundamental that they are explanatory and I will argue, that it is a necessary condition of 'essence' that the properties we take to constitute it for any given object can explain certain metaphysical features of that object.

The claim that essences ought to be explanatory has more than intuitive appeal. The philosophers who first talked about essence - most notably Aristotle - originally posited it as a kind of explanatory principle. According to Aristotle, for example, the essence of an object could be defined (even prior to

knowing which properties constitute it) as 'an internal principle of change and staying the same'[17]. This definition might not appear to have anything to do with explanation unless one understands the Aristotelian conception of principles (*archJ*). The word archJ means source or origin, something on which the object which has it is dependent. Now, principles thus understood bear an important relation to the Aristotelian notion of 'cause' (*aitia*). The two notions were considered to be coextensive and consequently to imply each other[18]. Aristotelian causes are generally thought to be best understood as kinds of explanation[19]. Thus, for Aristotle, it was a crucial feature of principles that they are explanatory. So, we can infer that he must have thought of internal principles of change and staying the same as explanatory[20].

So there is an historical precedent in philosophy for the requirement that essences be explanatory. Furthermore, there are modern philosophers who have argued that explanation remains the primary function of the 'essences' of objects. Michael Ayers for example, in extracting what he takes to be the primary characteristics of 'substance'[21] from the philosophical tradition, argues that the essences or underlying natures of substances are the ultimate principles of explanation[22]. What he means by this is precisely that the attribution of essences to objects is required to make sense of the very idea of "enduring things which are capable of affecting other things in a variety of ways and of suffering an indefinite variety of changes themselves"[23]. A different sort of argument for the explanatory nature of essences has been offered by Michael Loux. He points out that we have to accept one of the following three propositions:

1. All of an object's properties are equally important to it's being that very object and not something else.
2. None of an object's properties are involved in making it what it is.
3. Some of its properties are crucial to a thing's being what it is whilst others are not.

He goes on to argue that not only are 1 and 2 implausible but they leave us in an impoverished position as far as our *understanding* of the metaphysics of concrete particulars is concerned. He concludes

17 Aristotle Physics II ,1 192b14-20.

18 See Witt 1989 17-19p.

19 For a very modern exposition of this view see Moravcsik 1974 3-5.

20 That this is how Aristotle's talk of essences is ought to be understood is argued for in Megone 1988 p2-6.

21 Ayers is using the term 'substance' in the traditional metaphysical sense i.e. as the name of a category of being. 'Substance' in this sense is more or less synonymous with what I have been calling the class of concrete particulars or just objects.

22 Ayers 19981p70 and p85-9.

23 Ibid. p85-9.

that we must accept 3 - and consequently essentialism[24]. Given that so much of the argument in favour of an essentialist account of substance relies on the capacity of essences to explain 'everything else', I think it is fair to conclude the explanatory power of essences is a crucial characteristic of them.

So most metaphysicians who invoke the notion of essence, do so for *explanatory* purposes and, as it stands, Kripke's view is unable to furnish much in the way of such insight. In order to get from Kripke's defence of necessary properties to a viable notion of essence, we need some way of dividing the class of an object's necessary properties into two, one containing those necessary properties which provide insight into the potential and dispositions of the relevant objects and the other containing those that do not. In the next two chapters I will attempt to enumerate a means of doing just this.

2.5 Summary

Once again to conclude, I want to summarise the position reached thus far. It was established in the previous chapter that objects have necessary properties. In this chapter I went on to argue that *de re* necessity is a necessary but not a sufficient condition for regarding any particular property of an object as part of that objects essence. I further suggested that an additional necessary condition of essence is that it is explanatory. In the next chapter I will discuss the proposal that essences are essentially explanatory and explore how this might help us to develop a conception of essence more in keeping with the traditional understanding of the notion from the position defended by Kripke *et al*.

24 Loux 1998 p93-127.

CHAPTER 3

ESSENCE AND EXPLANATION

At the end of the previous chapter I suggested that the term 'essence' as it is used in metaphysics, is an essentially explanatory notion. In this chapter, I want to examine exactly what that claim amounts to. In the first part of the chapter I will discuss the conceptual link that exists between essence and explanation. In the following section of this chapter (3.2), I will go on to spell out how this connection can be exploited to develop a metaphysically recognisable conception of essence from the foundation provided by arguments of Kripke and Putnam discussed in chapter 1. It will be obvious from section 3.2 that the success of the attempt to derive essence from necessary properties relies on the possibility of an objective evaluation of explanatory worth/significance. It has, in recent years, been fashionable to deny that there are any virtues which explanations have independently of the interests and attitudes of the people to whom they are addressed. In section 3.3 of this chapter, I will set out the reasons that the proponents of this view give for holding it and then argue that they are bad reasons and that the view they are invoked to support is mistaken. The final section of this chapter will be devoted to a brief discussion of the criteria in terms of which I think metaphysical explanations ought to be evaluated. Having thus established that talk of the objective value of explanations has content, the stage will be set for chapter 4 where, with the help of some examples, I will identify the kinds of properties which make up the essences of objects on the assumption that the essence of an object is that set of its necessary properties which explains it metaphysical characteristics e.g. its persistence through change, its unity, its being at all etc.

3.1 Essence and Explanation as Conceptually Related

The relation between the notion of essence and that of explanation can be approached in different ways. It can be argued that concept of essence is a metaphysical postulate and that the metaphysical postulates we acknowledge are those which furnish the best explanations of the phenomena with which metaphysics is concerned. Thus the notion of essence must, by definition, be invested with significant explanatory powers. The alternative approach is to examine the role played by the concept of essence in the metaphysical theories which have invoked it *i.e.* to provide a functional analysis of the idea. The former approach will identify a practical connection between essence and explanation and demonstrate the more general truth that the power to explain is an inalienable feature of whatever is ontologically fundamental. This is the approach that I will take in this chapter. I am deliberately eschewing the functional analysis because I do not want to beg any questions in favour of the quasi-Aristotelian conception of essence that my dissertation is an attempt to defend.

To understand the relation between essence and explanation it is important to understand the purpose of metaphysics and the nature of its subject matter. Metaphysicians often find it difficult to specify what it is that metaphysics is about and what we can hope to gain by engaging in it[1]. The subject has had a chequered history and has been associated with a range of subject matters including the nature of being, the defence of theism, questions concerning the nature of free will and the idea of mental causation etc.[2]. The kind of metaphysics in which the notion of 'essence' is pivotal is the field of study which takes as its subject matter the nature of being. It concerns itself with question such as 'What is it to be?'; 'What kinds of being are there?'; 'What characteristics do beings have *qua* beings?' and 'How ought we to understand concepts like identity, particularity, unity or wholeness, similarity, difference and any other concept which seems unavoidable when thinking and talking about being?'.

Why bother with such recondite questions? We don't need answers to them to do science and surely science tells us everything we need to know about nature and our place in it. Well, there are at least two important considerations here. First it can be argued that the very existence of metaphysical questions indicates that scientific theories do not (cannot) provide a complete account of what we take to be reality. If the purpose of enquiry is as full as possible an understanding of our 'reality', then no body of knowledge which ignores metaphysical problems can be regarded as complete. Second, this kind of metaphysics is required to justify everything that science assumes. It is foundational to science in the sense that it justifies the presumption that we can sensibly talk about objects, qualities, causes, change etc. and that we can distinguish between what is one thing and what is a collection of things or a part of a thing etc. In addition, it is metaphysical theories which specify exactly what it is that we are talking about when we use these concepts. So the purpose of addressing questions such as these is ultimately *understanding*. Metaphysicians posit whatever is necessary in order to explain the way reality appears to us and good metaphysical theories are distinguished from bad ones primarily by their ability to provide economical yet exhaustive explanations[3]. Metaphysics just is an attempt to provide the information (I) which is of such a nature that for any reasonable (or minimally reflective) audience (A), the citing of I in suitable circumstances, presents A with such information as A must learn in order to understand her most basic assumptions about herself and the universe to the fullest extent that is possible[4].

---

1. This becomes clear on reading the introduction of any introductory textbook on metaphysics. See for example Taylor 1992 p 1-8 or Aune 1985 p3-7 and 10-11.

2. See Loux 1998 p 4-11.

3. Throughout my discussion of essence and its relation to explanation I will be assuming that metaphysical realism is true. This is not a position I can defend here. For the rudiments of a defence of metaphysical realism see Ayers 1991.

4. Adapted from Matthews 1981p354 (definition 13).

Note that in saying what metaphysics is and what its goals are, it is necessary to make use of words like 'basic' and 'fundamental'. I said above that metaphysics was foundational to science (as it is in fact to any form of enquiry). It is important to be quite clear about what metaphysicians mean when they say that the objects of their enquiries are 'fundamental' or 'basic' or 'ultimate' or 'foundational'. Michael Loux has defined the ontologically or metaphysically fundamental as "that which is primitive or underived and which cannot be analysed in terms of, or reduced to more basic entities"[5]. Now, the primitive and the underived stand at one end of a relation of dependence. Consider fundamental entity A. By the above definition of 'fundamental', A must feature in the reduction or derivation of at least one other phenomenon. Suppose, for example, that A is an *analysand* of phenomena B, C and D. B,C and D are then, dependent on A for their existence or occurrence. Where relations of dependency like this exist, they can always be exploited for the purposes of explanation[6]. Because B depends for its existence on A, part of what is involved in understanding or giving an explanation of B's existence must be understanding the relation B bears to A[7].

Thus it follows (somewhat paradoxically) that the things we take to be metaphysically fundamental just are the non-linguistic correlates (the referents) of those beliefs, concepts and conceptual arrangements which are unavoidable and for which there is no conceivable explanation. Here we begin to see another aspect of the connection between metaphysical postulates like 'essence' and explanation. These entities represent the point at which the possibility of non-teleological explanation ends[8]. Why? Because non-teleological explanations are by nature analytic. They consist of accounting for some or other phenomenon by isolating those aspects of the entities concerned *in virtue of which* the phenomenon occurs. Where the entities concerned are fundamental, it does not make sense to talk about aspects of them - they are (remember) unanalysable or irreducible by definition. Another way of making the same point is to appeal to the fact that non-teleological explanations invoke generalisations. For example, when we explain Joe's death by saying he was shot in the chest, we are tacitly appealing to the generalisation that gun shot wounds to the chest are usually fatal. When we explain the rise in the level of the mercury in a thermometer by saying that the temperature in the immediate surroundings must have risen, we are relying on the scientific law which states that metals expands

5. Loux 1998 p11-17.

6. I am aware that this is not an uncontroversial claim. However at this point, I do not propose to defend it. I hope that my arguments against the relativity of explanation later in the chapter will provide the germ of defence of this claim and all other anti-relativist claims.

7. See Kim 1987 in Ruben 1993 p 228-9.

8. The purpose of specifying that the explanations in which I am interested are non-teleological is to beg as few questions as possible. At this point , I want to leave open the possibility that teleological explanations can be given of fundamental entities but nothing I will say in this dissertation commits to the claim that fundamental entities are explicable in teleological terms.

when heated and the additional fact that mercury is a metal. Many explanations take the following form:

G : A explains the fact that B Ns because B is an A and *all As N etc.*

What does all this have to do with the claim that ontologically fundamental entities cannot be explained non-teleologically? Well, because the postulates of metaphysics are fundamental, they cannot be the subjects of the kind of generalisations alluded to above. A fundamental entity cannot bear the requisite relation to some other fact or proposition (a law or generalisation) because it represents the level at which facts and propositions are the most general they can possibly be. For example, say quarks were the most fundamental building blocks of matter according to our best physical theory. We could explain things about objects and even atoms in terms of the characteristics of their constituent quarks but we cannot give the same kind of explanation of quarks themselves. The reason for this is that there is nothing more basic than quarks about which we can generalise. In other words, there are no further generalisations under which quarks fall. As a result there is nothing to which we can appeal for the purposes of explanation.

So postulates of metaphysics like essence, are fundamentally explanatory in the sense that the existence of the concepts/terms with which they are associated in our lexicon is conditional upon their capacity to explain surface phenomena. Further, concepts like essence must be explanatory because they have as their referents fundamental entities. Finally the more fundamental the postulated entity the less likely it is that we will be able to give a non-teleological explanation of it.

3. 2 The Derivation of Essences by Appeal to Explanatory Significance

In chapter 1, I argued that objects do have necessary properties in their own rights. I went on, however to insist that the set of an object's properties that it has necessarily is not synonymous or even co-extensional with its essence. In chapter 2, I listed various kinds of necessary property which seem to have no business being included in essence. I then suggested that the reason that such properties did not fit either the common-sense notion or the technical metaphysical understanding of essence, was that they had little or no role to play in explaining the ontological[9] characteristics of the things to which they can be attributed. In the preceding section of this chapter I have tried to defend that claim. The resultant picture of essence at which we have arrived is this:

9. By 'ontological characteristics' here I mean those features of existent things that they have in virtue of being existing things.

H: The essence of an object is that sub-set of its necessary properties which can play a significant role in providing answers to questions concerning about that object *qua* being.

The kinds of questions I have in mind include things like why the object concerned persists through change; why something is one thing rather than many things; how being a part of something else affects what it means for that object to be at all; what the relationship between the object and its parts or the things of which it is a part is like etc.

I want to suggest that we can make use of the requirement that essential properties must be explanatory to distinguish the essential properties of things from those which are merely necessary. I think we can do this by formulating a metaphysical question to which we have good reason to believe necessary properties can provide meaningful answers and then determining which kinds of necessary properties provide the best answers to it. This is how I propose to proceed in the next chapter. However, before I can do that I need to defend the claim that objectively determinable degrees of explanatory significance can be attributed to certain facts. Relativism with respect to explanation represents a serious threat to my project because it suggests that there is no way in which some features of objects can be singled out as more explanatorily important than others. According to the relativist, explanatory importance is a feature of the perspective brought to a state of affairs by an inquirer. As such, it cannot correspond to anything real and it cannot tell us anything about the state of affairs to which the inquirer addresses himself.

3.3 The Purported Relativity of Explanatory Significance

The thesis that objectively determinable degrees of explanatory significance can be attributed to certain facts is in accordance with our common-sense understanding of explanation. However, it is a view regarded with suspicion by many philosophers. The reason for this is the recognition that explanation is unavoidably pragmatic. Explanations are *essentially* means of bringing about understanding. Understanding, in its turn, is a feature of people rather than of propositions - it is a mental state which is dependent on a whole range of other mental states. Whether explanations are successful, or even appropriate, would seem to depend both on the existing attitudes and beliefs of the persons to whom they are addressed and on what it is that these persons wish to know. This has led some philosophers to claim that explanations are *essentially* relative. According to the relativist, it takes certain attitudes, beliefs and desires on the part of an audience to make any particular utterance or proposition into an explanation - what counts as an explanation depends on what the enquirer wants to know and can accept in the light of the beliefs s/he already has.

If this view were correct then the explanatory power, significance or worth of an utterance or proposition would also be relative to a particular audience in a particular context. We would not be able to say how effective a given explanation will or even ought to be without knowledge of the background beliefs, attitudes etc. of a given audience. Thus according to philosophers sympathetic to explanatory relativism no theory, utterance or proposition can possibly have such a capacity *intrinsically* and consequently to predicate explanatory significance of particular forms of words is to talk nonsense[10]. For the relativist, it is conceivable that some false set of propositions might well be the *best* explanation for a particular person to give to a particular audience at a particular time in history or in a particular context. This is in direct opposition to everything that I have assumed up until now in this dissertation. Remember, I have claimed that explanatory significance automatically attaches to entities which are 'ontologically' fundamental. I have also claimed that we can recognise essences by their wide-ranging explanatory usefulness - a usefulness which stems from their ontological fundamentality and which they ought to have independently of our interests. If relativism about explanation is justified then the essences of objects will vary according to how people choose to think about them - and despite the objectivity of necessary properties established in chapter 1, we will not be able to give a realist, non-nominal account of essence.

Is the reasoning of the relativist justified? Obviously, I think that it is not. Whilst it is undeniable that explanations have pragmatic aspects and that the actual capacity to increase understanding in any context must depend on the persons concerned, this does not justify the inference to relativism. The arguments against relativism do not deny the pragmatic aspect of explanation; they simply insist that their existence does not entail relativism.

3.4 The Objectivity of Explanatory Power

The mistake that the pragmatists make is to assume that whether or not something is explanatory is *entirely* determined by pragmatics. That this is a profoundly counter-intuitive assumption becomes clear when we consider the question of whether, in the right context, any randomly chosen proposition could be taken to explain the truth or falsity of any other randomly chosen proposition. The common-sense answer here is 'definitely not'. To begin with, there is a multitude of examples where differences in explanatory power appear to be constant across all possible contexts. Surely, for example, to say that Chlorine can bond with Sodium because it is like Bromine and Bromine can form

10 See van Fraassen 1980 in Ruben 1993 p308-9.

similar bonds is less explanatory in every conceivable context, than the claim that atoms form chemical bonds in order to fill the outer orbital of their structure with electrons (the maximum no of electrons in one atomic orbital is 8) and since Chlorine has seven electrons in its most peripheral orbital whilst Sodium has one, the two elements complement each other in precisely the right way. Note that an explanation in terms of quarks or super-strings would by the same line of reasoning be even better. I am claiming that the second explanation is better than the first for all audiences. The pragmatist will immediately retort that what I say is quite obviously false. For someone who is familiar with Bromine but not with atomic chemistry, the latter explanation is powerless to enhance her comprehension of the bonding of Chlorine and Sodium whereas the former at least places the phenomenon in a familiar category. How might the realist respond? Well, she would have to argue that we should not make our evaluations of explanations relative to the degree of ignorance of the audience. That is she would have to claim that the best explanation of any phenomenon must be the one which is most acceptable to the person who has as much knowledge as is available about that phenomenon. This seems to be how we proceed in Science - with disputes only taking place at the level of maximal expertise. It also has common-sense manifestation - most of us who do not understand quantum mechanics are quite happy to accept that it offers better explanations of matter and its exploits than anything we do understand - say either atomic theory or Newtonian mechanics. So, there is a precedent in science and in everyday practice for treating explanatory worth as if it were objectively measurable.

Even without the common-sense appeal to the empirical, it can be shown that the pragmatist's premise must be mistaken. Using randomly picked propositions p and q, Philip Kitcher and Wesley Salmon have demonstrated there can always be defined a logical function which will render p explanatory of q regardless of the semantic content of either proposition[11]. This implies that for any two propositions say '2+2=4' and 'Socrates is not the Taj Mahal' there is a context in which the former explains the latter. Thus the assumption that the explanatory status of a proposition is entirely pragmatic appears to have absurd consequences. It would seem that there is some objectively determinable feature(s) of the propositions concerned - something to do with their semantic content - which makes one such proposition explanatory of the truth/falsity of another. Thus, there must some non-pragmatic conditions which determine what can and cannot serve as an explanation. Variations in the degree to which these conditions are satisfied will bring about variations in objectively determinable explanatory significance which one proposition has with respect to another.

The fact is, propositions and their extensions exist within webs of inter-related propositions, objects and events, and what we are doing when we explain something is making a tacit appeal to some objective

11. See Kitcher and Salmon 1980 p312-20.

relation that exists between the *explanandum* and the *explanans*. When, for example, we say that 'the cat is on the mat because the mat is in front of the fire' and take this to be explanatory, we assume that there is a perspicuous relation between the mat's being in front of the fire and the cat's being on it or wanting to be on it. The kind of relation involved is usually implicit in the request for an explanation and can be clarified by consideration of what does and does not count as a suitable response to a given request. Note that whilst there are probably a range of relations to which the inquirer could appeal and that whilst the choice of relation to which the appeal is made will most probably be determined by contextual factors, that a particular kind of relation is there to be appealed to is a matter of fact[12]. Note also, that once the appeal to a particular relation (R) has been made in the process of formulating a question, that the appeal is to R and not some alternative relation (S) is an *objective* part of the explanatory context. An explanation which appeals to S rather than R is just as irrelevant as an explanation which cites a state of affairs which is completely unrelated to the *explanandum*[13]. The most frequently discussed of the potentially explanatory underlying relations is the one that exists between a cause and its effect. But causal relations are not the only such relations to which appeal can be made for the purposes of explanation. I think any relation, as long as it exists, can ground an explanation in a suitable context. Any objectively existing relation between two entities, parts of entities, events, parts of events, facts, propositions, ideas or concepts, can potentially be exploited for explanatory purposes.

So the class of propositions (B...N) which are potentially explanatory of a proposition A is limited to those to which A bears some or other objective or *real* relation. In practice what this amounts to is that it is a pre-condition of something's being an explanation of A that we must have good reason to believe that the relation between A and say B to which the purported explanation is an appeal is real. Thus it is that (our best theory about) the way the world is (or the way the phenomena are, if you prefer) has some role to play in determining what is and is not explanatory - particularly in science and metaphysics. So, for example, the phases of the moon could never explain my behaviour because there are no good reasons to believe that a relation between said behaviour and the moon exists. The pragmatist would probably reply that it is what we *believe* about the world that is doing the work here and we could, after all, believe anything we choose. If, for example, someone believes in a god who has the power and the inclination to exert some kind of causal influence on the universe, then *for that person* it is true that events in the universe can be explained in terms of this god's intentions. I think

12. Van Fraassen (1977 p149) hints at this idea in discussing the relation of explanatory relevance. He says that the appropriate relation will be determined by context but he does not place enough emphasis on the fact that the relevant relations must exist or we must have good reason to believe that they do and as a result fails to acknowledge that there is a sense in which the relations that can serve the explanatory purpose are constrained by the way the world (or the phenomenon) is.

13. See section 4. 3 for a discussion of some examples where the kind of relation invoked requires that certain kinds of explanation be rejected from the outset.

the anti- relativist can allow the pragmatist this much without giving up much ground at all. Consider explanation of events in terms of god's purposes. The anti-relativist can agree that such explanations are indeed acceptable provided, of course, that the belief in this god and her relation to the universe is a reasoned one - the onus being on the theist to prove its rationality.

Thus what counts as an explanation is just as dependent on the way the world is as it is on what people choose to inquire about. Explanations always map objective relations. But what I need to establish in order for the arguments of this dissertation to go through is that there are objective means of *evaluating* explanations i.e. that there are criteria in terms of which the worth of an explanation can be assessed. What I have discussed above indicates that one way in which an explanation can be better than 'another' is by appealing to a real rather than a non-existent relation between the *explanans* and the *explanandum*. However, if explanatory power is just a reflection of the accurate mapping of objective relations, it will be an all or nothing matter. An explanation either appeals to an existent relation or it does not. Hence, if explanatory power were purely a matter of ensuring that there is a real relation between *explanans* and *explanandum* it will not come in degrees. This is counter-intuitive. It seems perfectly reasonable to recognise the possibility that where two facts or propositions bear the appropriate relation to the *explanandum,* one may constitute a better explanation than the other. How can we account for the reasonableness of this supposition?

Robert Matthews develops an interesting account of objective explanatory power in terms of the more general idea of 'powers' as we understand them in everyday language[14]. According to Matthews, the propositions that make up a particular *explanans* - say those we use to explain X - can be ascribed a kind of power: the power to bring it about that X is explained[15]. To ascribe a power P to an entity E, is really just to say that, whenever E is in the presence of an appropriate external stimulus, some specific event Q will result. We are familiar with powers from examples such as the claim that water has the power to dissolve salt. What the attribution of the power or disposition to dissolve salt to water amounts to, according to Matthews, is simply the claim that whenever water is mixed with salt, the salt will be dissolved. How does all this apply to explanations? Well, whatever is explanatory of A, can be described as having the power to bring it about that A is explained. So what? The important thing is that things (objects, propositions etc.) have whatever powers they do *in virtue of* the properties they instantiate. So a particular *explanans* will have its explanatory power in virtue of certain of its properties. Matthews thinks that attributions of explanatory power to propositions are true in virtue of

14 Matthews 1981 p347f.

15 Note that the explanatory power of a proposition or set of them will always be relative to the content of the proposition it is an attempt to explain

properties associated with the content or information they carry. Consider the case where B explains C. This happens when the knowledge that B is a *correct* answer to the question 'Why C?' is produced in an audience A. So B must present A with whatever *information* he needs to acquire in order to understand C. And the intrinsic explanatory power of B will thus be measurable in terms of its ability to provide the requisite information. Note that the information required is not whatever A most wants to hear but rather what A *must* hear in order to have genuine understanding. B has its explanatory power independently of whether or not a given audience believes that it does, in just the same way that '2 is a prime number' and 'e-mc^2' are true independently of whether or not anyone believes them. Now as is the case with inference, information can be more or less relevant, more or less informative, more or less general etc. and as these factors vary, so will the explanatory worth of the propositions which have them. Note that relevance, generality, informativeness etc. are all things we can assess independently of any audience - in exactly the same way that we assess the same qualities in the premises of arguments. Thus we have the germ of an account of the objective explanatory value of one particular proposition in relation to another. Furthermore, because a variety of factors are involved, some of which admit of degrees, we now have the basis for an account of explanatory value within which it makes sense to talk about a given proposition being more or less explanatorily significant.

Thus I think a good case can be made for the at least partial objectivity of explanation and explanatory value. And this is all that is required to avoid the potential objection to identifying essential properties by assessing the capacity of various necessary properties to provide explanations. In addition, it would seem that there are reasonable, unambiguous criteria in terms of which the explanatory worth of a particular proposition or theory can be judged. Since I intend to make use of such criteria in the next chapter, I think will be worthwhile to take a more detailed look at some of the contributory factors to the objective explanatory significance of propositions in metaphysics. This I will do in the next section.

3.5 Criteria of Objective Value of a Metaphysical Explanation

The literature on explanation suggests a range of virtues which purported explanations can have independently of the interests, beliefs and attitudes of the audiences to whom they are addressed. Much has been made, for example, of the role played by a potentially explanatory fact in the most acceptable theory of the cause of the state of affairs depicted by the explanandum[16]. Other criteria which have been suggested include the capacity of the truth of the explanans to increase the probability

16. See Lewis 1986 in Ruben 1993 p 182-8 for the classic exposition of scientific explanation as causal.

that the *explanandum* is also true (so-called favouring of the *explanandum* by the *explanans*[17]); the capacity of an *explanans* to provide the kind and amount of information necessary in order for the explanandum to be comprehensible (hereafter referred to as its informativeness) and degree to which the truth of a particular explanation provides a means of unifying the phenomenological 'data' to which it applies. Let us take a look at each of these in turn.

The call for an evaluation of explanatory significance implies that a comparison needs to be made between alternative *explanans*. Bas van Fraassen recommends that the required comparison be statistical. According to him the best answer to a why-question (all explanations are considered by van Fraassen to be answers to why-questions), will be the one which maximises the probability of the *explanandum* also being true[18]. This is a neat and simple approach to evaluation but it does not really provide much insight into the nature of explanatory value. Part of the point of evaluating the explanatory significance of a particular fact in a particular 'conceptual' / 'linguistic' / 'logical' context is the attempt to gain insight into the reasons that one particular explanans favours a given explanandum more than the alternative *explanans'* do. Van Fraassen recognises that a statistical account of explanatory value may well not be adequate on its own. It is possible, for example, that two equally probable potential *explanans* have identical effects on the probability that the relevant *explanandum* is true. In such cases, he says, the evaluator should carry out a comparison of the competing *explanans'* in other terms. He leaves the matter of the criteria in terms of which such a comparison should be done fairly open. One 'answer' might make some of the others wholly or partially irrelevant, and this would be a reason for preferring the former[19]. Alternatively, according to David Lewis, one answer to a why question can be better than another because it provides more relevant information, less misinformation and less superfluous information[20] *or* because it unifies the data to a greater extent *or* because it best fits what the enquirer wants to hear or what s/he can assimilate given her current level of knowledge for example[21]. Since we are interested in explanatory significance which is not relative to the interests of the audience we can ignore the last of these options. This leaves us with three criteria in terms of which to measure objective explanatory worth - two of which, informativeness and unifying capacity are already under consideration (see previous

17. See van Fraassen 1980 in Ruben 1993 p299.

18. The relevant probabilities will of course be relative to a set of alternative states of affairs or propositions which could have been realised/true but which in fact are not. Any why -question/ request for an explanation picks out a particular set of contrasting states of affairs to which P belongs, any one of which could have been the case but no two of which can be realised simultaneously - the contrast class of the question. Sometimes when there is only one way in which P could be false, the relevant contrast class will simply consist of P and not P.

19. He cashes out what it means for one answer to make another irrelevant in statistical terms - if the probability of P, given the truth of A and A', is equal to the probability of P given only the truth of A' then A' makes A wholly irrelevant.

20. Lewis, 1986 in Ruben 1993 p 193-4.

21 . See Achinstein 1985 p340 on the requirement that goodness of explanations reflect the ends we have in giving them. See Lewis 1986 p193

paragraph). Before we move on to these let us take a brief look at the third option we seem to have for evaluating the objective significance of explanation.

What does it mean to say that one fact or proposition makes another explanatorily irrelevant? I think that there are a number of possibilities. Van Fraassen again makes this a probabilistic measurement. According to him answer A makes answer B irrelevant if the probability of C (the *explanandum*) being true given that A is true, is greater than or equal to the probability that C is true given that both A and B are true. I think the irrelevance is best understood in terms of logical entailment. Where proposition A implies B and B implies C then if B is explanatory of C so is A. Now, when A is cited as an explanation of C, there is a sense in which the given explanation includes B whereas, when B is so cited the resultant explanation does not include A. A has a wider scope than B, and one in which B is included and as a result is arguably the more complete explanation. The claim that the wider the scope the greater the explanatory power (all other things being equal) is controversial. It makes completeness a virtue of explanations. However, I think it can certainly be defended in metaphysics where completeness is an important indicator of fundamentality or ultimacy[22].

Now let us consider informativeness. That an explanation can be informative presupposes that it picks out some non-obvious feature of experience which bears some (in principle knowable but possibly unknown) relation to the piece of experience we are attempting to explain. When is a purported explanation more informative than some other competing explanation? As we saw above, Robert Matthews suggests that an answer is maximally informative if it presents an ideal audience with precisely those facts which such an audience must learn in order to understand the explanandum (P) and does so in such a way that the audience attributes their new-found understanding of P to the acquisition of this information[23]. Note that this does not imply that the explanation which furnishes the most facts is always the better one. It is possible to provide too much information. Take for example a causal explanation of the movement of one billiard ball (B) as a result of being hit by another (A). The actual cause would include the hardness and shape of billiard ball B, the texture and gradient of the surface on which the balls move, the wind resistance and probably a whole host of other things including whatever it was that made billiard ball A move in the first place. To cite all this in an explanation would in most cases be to provide too much information. It is important to avoid telling the audience what they already know, what they are not interested in, or what is insufficient to their

---

for a clear and concise summary of the ways in which an explanation can be more or less satisfactory.

22. See section 3.1 above.

23. Mathews 1981 p352-354

needs[24]. All of which is inherent in - there to be read off from - the original request for an explanation.[25,26]

Moving on to the idea that explanations ought to unify, Philip Kitcher argues that a primary virtue of scientific explanations has always been their capacity to unify or systematise a body of data[27]. But what does it mean to say that a given explanation unifies the data? Unification has been defined as "the comprehending of a maximum of facts and regularities in terms of a minimum of theoretical concepts and assumptions"[28]. This definition makes unification as a matter the capacity to provide fundamental explanations - the kind which cite principles, concepts and facts which represent the absolute base level of what can be understood. Those (hopefully few) facts or principles which can explain everything else or which can be said to systematise the entire range of the phenomena. As I mentioned in section 3.1 principles or concepts or facts which play this role cannot but be explanatory. It can be argued that they are more explanatory or that they offer better explanations because they require a minimal number of assumptions/beliefs to make them true/explanatory and because their scope includes that of most other potential explanations. Kitcher acknowledges that it need not be the case that there is only one explanation which systematises a given set of data[29]. He argues that we ought to consider the systematisation which derives the maximum number of accepted propositions from the minimum number of assumptions to be the best. Even then however, there might be two comparable theories. In such cases he recommends we accept as better, that set of theoretical assumptions/ postulates which exhibits the higher degree of inter-relation between its component propositions and or arguments[30].

There is a range of factors which can be called into service for the purpose of giving an objective evaluation of the explanatory significance of one proposition or set of propositions with respect to another. Above, I have selected three which will be particularly relevant for the evaluations I carry

---

24. Lewis 1986 in Ruben 1993 p 193-4.

25. Van Fraassen has argued that the requirement that the kind of information required is specified in the question implies that the value of an explanation is relative to its audience. Peter Achinstein argues that this is not the case. There is always a fact of the matter concerning what information a particular question is seeking, so although to fully evaluate the worth of an explanation we have to know quite a lot about the context in which it is given, we still have objective access to the problem of whether it does or does not answer the question. See Achinstein 1985 in Ruben p328-33.

26. Peter Lipton (1990 p215) suggests that this can be more accurately done for causal explanations using his method of causal triangulation. Basically some causal factors will be the same for P and other members of the contrast class, what the question will be asking for are those causal factors which are not the same for P and any of the other members of X.

27. Kitcher 1981 p511.

28. Feigl 1970 p12 cited in Kitcher, 1981 p508.

29 . *ibid* p519-529.

30. Kitcher 19 81 p 521.

out in the following chapter. The discussion above is intended to provide a brief sketch of what is involved in 'measuring' qualities like informativeness and the potential to unify a body of data. It will be apparent that I have relied heavily on the more exhaustive discussions of these concepts which occur in the literature cited. However, I hope what has been said here will give the reader a clear idea of what I mean when I talk about these things in the next chapter.

3.6 Summary

In this chapter I argued that the notion of essence is explanatory. The point of doing so was twofold. First I wanted to establish that the capacity to explain certain metaphysical phenomena is a necessary condition for a being essential. Second I wanted to suggest that we can make use of this necessary feature of essences in order to distinguish essential properties from necessary ones. The second part of my project might be considered problematic because explanatory capacity is often taken to be relative to the interests and beliefs of an audience. The majority of this chapter was dedicated to demonstrating that there is such a thing as objectively determinable explanatory worth. There are objective aspects of competing explanations in terms of which we can compare and assess them. In the next chapter, through an appeal to such factors I will demonstrate that there is a class of necessary properties which generally provide better explanations of the persistence of continuants through change. I will then claim, on the strength of what was said in 3.1, that these and only the properties which fall into this class ought to be considered to be the essences of the objects which possess them necessarily.

CHAPTER 4

NECESSARY PROPERTIES AND THE EXPLANATION OF PERSISTENCE THROUGH CHANGE

In chapter one it was established that any item to which we can successfully refer by using a proper name has necessary properties. In this chapter I want to focus on some of the cases in which successful reference does occur and, picking up where I left off in chapter 2, I want to investigate the significance of various types of necessary property that the relevant referents can have[1]. I will assume that there *are* (in some sense) such things as individual objects - objects which are naturally distinct in the sense that their boundaries are given in experience and not imposed by the experiencer. The purpose of this chapter will be to show that some of an individual object's necessary properties are more metaphysically significant than others and to suggest a method for separating the metaphysically significant from the insignificant necessary properties of an object. As indicated in previous chapters, in order to achieve this, I intend to make a direct appeal to the metaphysical questions to which we expect essences to constitute an answer.

In chapter 2, I suggested that a relatively small subset of the necessary properties of any object offer better answers to such questions, than do those properties which it has necessarily but which fall outside this sub-set taken either individually or as a group[2]. I also noted that in order to get from Kripke's defence of necessary properties to a viable notion of essence, we need some way of dividing the class of an object's necessary properties into two - one containing those necessary properties which provide insight into the potential and dispositions of the relevant objects and the other containing those that do not. In chapter 3, I went on to suggest that the very capacity to provide the relevant explanations be used as the criterion for dividing the necessary properties into these two sub-classes. In this chapter I want to follow up this suggestion with a demonstration of how I think the process of isolating essence in this way will work. For the purposes of this dissertation I will focus the capacity of different kinds of necessary properties to explain the survival of objects through some changes and not others.

---

1 The paradigm case of things to which proper names apply, is that of ordinary objects - books, beetles and buildings *etc*. and I will confine my discussion to these.

2 This is a view which has its origins in the work of Aristotle. He thought, logically enough, that the presence of certain properties in the definition of a thing and their capacity to explain various things about that thing were inextricably related. Much of what I have to say in this chapter will be guided by Aristotelian insights into the nature of concrete particulars.

Persistence through change is a very ancient and yet persistently puzzling problem which has a long association with the notion of essence. The problems associated with change are thought to have been first raised by Parmenides who was concerned to understand how changes of any kind are possible. According to him, the ability to make sense of change required us to think of some object as either coming to be of out of nothing or as transforming into another object- i.e. into what it previously was not. Both options require that we be able to comprehend the notion of 'what is not'. Parmenides believed that such a notion is not a possible object of speech or thought[3]. He concluded that the world of change was an illusion and that reality was changeless. Heraclitus considered the same problem and reached the opposite conclusion, namely that the appearances things give of remaining the same or of being unchanging is an illusion and that reality is in a state of constant flux[4]. Aristotle found both of these views intolerably counter-intuitive. He thought that part of the problem lay in the way in which philosophers were construing change. To solve their puzzle, he found it necessary to distinguish between two fundamentally different kinds of change - substantial change and alteration[5]. Substantial change occurs when something changes from one Aristotelian substance (concrete particular) to another - a coffee pot becomes a pile of china dust or Socrates becomes a daisy for example[6]. Alteration occurs where a single Aristotelian substance persists but there is a change in one or more of its characteristics e.g. Socrates goes from being pale and ignorant to being tanned and wise without becoming something else in the process[7]. Aristotle argued that for change to even make sense, there must be some part of whatever changes that remains the same - *i.e.* that only part of a changing thing is ever transformed (Physics 189 a20-191b35). In the case of substantial change, it is the stuff out of which the relevant object is made that remains (Physics 191a19-20 and 192a20-33), whilst in the case of alteration, the *essential* properties of the object persist whilst its accidental properties are exchanged for others. It is with the kind of change which Aristotle calls alteration that I will be concerned in this chapter.

In what follows, I will begin by discussing the reasons that the persistence of objects through certain changes might be considered puzzling (section 4.1). I will then provide a brief defence of the claim that essence and persistence are related phenomena. I will argue that the existence of this relation entails that the properties which a thing has necessarily can furnish explanations of that thing's capacity to

3 Parmenides is generally taken to have argued for the impossibility of change - either of coming or ceasing to be or of the individual properties of things. His arguments were more rigorously formulated by Zeno (of Elea) See Hussey, 1995 p645-6 and 922-3.

4 Hussey 1995 p351-352.

5 Aristotle's response to the Eleatic paradoxes can be found in *Physics* I, 6 and 7 in Ackrill 1987 p86-90.

6 Definition from Megone 1988 p2. Coffee pot example from Wiggins 1980 p27.

7 We can think of alteration as identity preserving change whilst substantial change is the kind of change which involves a change in the identity of the thing that changes.

survive some kinds of change but not others (section 4.2). The bulk of the chapter will be devoted to exploring, via examples, the quality of the explanations that different kinds of necessary properties provide. I will argue that the most appropriate answers to questions concerning the capacity of objects to survive certain changes, cite members of a particular sub-set of the necessary properties of the relevant objects (section 4.3). In the first part of this section I will demonstrate for a specific 'why-question' concerning the capacity for alteration of a particular object, that by citing certain of the necessary properties of an object we obtain better answers to the question than we would by citing others (section 4.3). I will then argue that the explanatory significance of the kind of property identified in section 4.3 can be generalised across most objects and most kinds of change (Section 4.4). I will offer an account of the apparent pre-eminence of type 4 properties and then go on to claim, in accord with the conclusions of 3.1 above, that this difference in explanatory value provides us with both the grounds and the requisite criterion for dividing the class of necessary properties of any object into two and that this distinction coincides with the distinction between those necessary properties of objects that are metaphysically fundamental and those that are not (section 4.5). I will claim that the sub-class of an entity's necessary properties which can provide explanations of that entity's persistence through change comprises its essence.

4.1 Persistence Through Change

Both objects and stuffs are what metaphysicians call continuants. Continuants are things that exist *through* time rather than as a series of temporal stages. Continuants are interesting because most of them have the perplexing (from the perspective of the metaphysician) capacity to persist through change - their characteristics alter, but they remain the very same objects that they were before the alteration occurred and it is not obvious why or how this is possible. The changes involved can be dramatic. Acorns become oak trees and caterpillars become chrysalis's which, in turn, become butterflies etc., and none of this affects the identity of the thing which changes - we have the very same plant or the very same insect etc., both pre and post change. There are however some changes which continuants do not survive - most material[8] continuants would cease to exist if they were chopped up into little pieces for example. Socrates and the result of putting him in a blender would not be the same *anything*. A different kind of change which appears to compromise the identity of an entity is the swapping of membership of one natural kind for that of another. For example, something like a change of species seems mark the going out of existence of one entity and the coming into being of another. Were Socrates to be changed into a daisy, this would signal the end of his existence just as surely as would his death. Note that the impossibility of identity preserving change represented by the

8 I don't know if there are any non-material continuants, but this is not a question I want to beg at this point.

above example is not a matter of a species change causing the thing we know as Socrates to migrate across an arbitrarily defined conceptual boundary. In becoming a daisy the thing that was Socrates becomes something else entirely - something which is discontinuous with Socrates regardless of our concept of what it is to be Socrates.

The obvious answer to questions concerning the possibility of different kinds of identity-preserving change is the one I attributed to Aristotle above, and the one that will be explored in this chapter. Socrates can survive some changes but not others because the latter kinds of changes require a change in his necessary properties. Such answers are intuitively plausible and informative. However, modern philosophers - including Kripke - have questioned the claim that there is a straightforward relation of equivalence between questions of persistence through change and the *de re* modal status of the necessary properties of the objects that persist. Kripke says:

> ..one should not confuse the type of essence involved in the question 'What o properties must an object retain if it is not to cease to exist and what properties can change while an object endures ?' which is a temporal question with the question 'What (timeless) properties could the object not have failed to have and what properties could it have lacked while still (timelessly) existing?' which concerns necessity not time and which is our topic here[9].

In the following section (4.2) I will argue that Kripke is wrong about this and that there is indeed a straightforward and potentially enlightening connection between the properties which a given thing must have in all possible worlds in which it exists and the changes it cannot undergo without a loss of identity.

4.2 Necessary properties and Identity-Preserving Change

Philosophers like Kripke believe that there is an important difference between questions concerning which of properties of a thing can actually change and questions concerning which of its properties a thing must have in all possible worlds in which it exists. Since I have chosen to look an essentially temporal question as a means of extracting the essence of an object from the set of its necessary properties, it is important for my project that no such objections can be raised. So in this section I want to spend some time examining just why it is that Kripke thinks there is a difference between what he dubs the temporal and modal questions and why he is mistaken. First I will try to explain how one might come to believe that the temporal and modal questions are different. I will then move on to discuss why I think such a belief would be false.

9 See footnote 57 in Kripke 1980, p144.

Why might someone think that the temporal and modal questions are distinct and/or unrelated? To recap (from section 2.1), Kripke's method of determining whether something is a necessary property of a given object or not, is to consider whether we can conceive of a possible world in which the very object with which we are concerned could have failed to have that property. If we can, then the property is not a necessary property of the object. He argues that when we consider counterfactual questions concerning whether or not something can fail to have a certain property and still be the very same thing alongside questions concerning whether something can undergo a certain change and yet remain the very same entity it was to begin with, our intuitions will lead us to give different answers in some cases. He concludes from this apparent divergence that the questions to which these different intuitions provide answers must also be distinct. Consider the following statements " and $:

" : Could Socrates have been a daisy?

$: Could Socrates have been turned into a daisy?

Statement " is what Kripke dubs a modal question whilst statement $ represents a temporal question. The answer to question " is clearly 'No'. A positive answer to " would assert that something could be both a human being and a member of a plant species simultaneously - which is impossible. By comparison, the transformation that $ asks us to countenance is something we accept quite happily in myths, fairy tales and bible stories. So the state of affairs which would arise if $ were true cannot be inconceivable. Thus the answer to $ might seem to be 'Yes'. So it does look as if our intuitions point us in different directions when we consider possible answers to modal and temporal questions, and there is a sense in which we reach different answers depending on whether the thought experiments we run are of a temporal or a modal nature. However, I think that before we infer from this that the questions posed in these different kinds of thought experiment must be *seeking* different answers, we ought to consider the reasons for the apparent divergence more carefully. I think that when we do examine the potential sources of our conflicting intuitions in the face of questions like " and $, we can account for the apparent conflict in ways which do not compromise the claim that the properties that a thing cannot lose without losing its identity just are its necessary properties.

In the above case the important question is whether when Socrates is turned into a daisy he retains his identity i.e. he remains Socrates. Clearly if the only possible answer to " is 'no', he cannot. What Kripke seems to fail to see is that whenever the temporal question is asked in the context metaphysics, as opposed to fiction, the condition that the object which changes retain its identity is inherent in the question. And once we insist that the thing that is turned into a daisy continue to have the identity of Socrates, the apparent divergence in our answers to the temporal and modal questions disappears. Whilst we might think that we can conceive of witches turning into black cats, gods taking on human

form or of Lot's wife becoming a pillar of salt, what we actually encounter here is one entity with a particular identity becoming another thing with a different and *discontinuous* identity. So despite the initial plausibility of the kind of transformations we countenance in fairy tales, myths and bible stories, when we seriously consider whether the initial object is the very same thing as the one into which it is transformed, if the transformation involves a change in those properties that the object cannot fail to have whilst remaining the thing that it is *(i.e.* its necessary properties), the only coherent conclusion we can arrive at is that it is not.

The view that the ways in which a thing cannot change without becoming something else corresponds exactly with the necessary properties it has, can also be defended by an appeal to the technical metaphysical concepts of which modality-inclined philosophers are so fond. Part of what it is for x to exist is for x to have an identity, and in considering the ways in which x can change without ceasing to exist, we are actually considering which changes do and do not allow x to retain its identity. If an entity x must be F in all possible worlds in which x exists (i.e. is necessarily F), then any world in which x ceased to be F would, necessarily, be a world in which x ceased to exist. Thus, the distinction between modal and temporal questions concerning properties (or as Kripke likes to call them, modal and temporal essences), collapses.

So not only is there a relation between the necessary properties that objects have and the ways in which they cannot change without ceasing to exist (or becoming something else) but this relation is one of simple equivalence. So we should expect all properties that emerge from Kripke's possible worlds device as necessary to be properties which the object in question cannot lose without ceasing to exist. This suggests that, as Aristotle thought, persistence through change is a matter of a thing's retaining its necessary properties and gaining or losing various accidental properties[10]. A result it is likely that we could appeal to the necessary properties a thing has, in order to explain *the ways* in which it can and cannot change without a loss of identity. In the following section I will explore this possibility.

4.3 Necessary Properties as Explanations

Thus far it has been established that there exists a relation between the ways in which a given thing can and cannot change and its necessary properties which looks as if it could be exploited for explanatory purposes. In this section I want to consider the nature and viability of the explanations that would result. As noted above I believe that such explanations are viable and that some necessary

10 As suggested by Aristotle: *Physics* I,7 (190a12-22).

properties provide better explanations of the phenomena associated with concrete particulars than do others. In this section, I will begin with a discussion of the nature of the 'persistence why-questions' to which the citation of necessary properties might constitute appropriate answers. I will follow this with a detailed analysis of an example which will serve to test the theory that of the necessary properties of an object can explain its disposition or resistance to various changes. What will become clear from the analysis, is that some necessary properties of a given object do a better job of explaining its persistence through various changes than do others. Not surprisingly it will turn out that those properties with superior explanatory powers do not include any of the slightly odd necessary properties listed at the end of chapter 2 (section 2.3) e.g. 'not being made of ice' and 'being such as not to contradict any necessary truths' *etc.*

4.3.1 Persistence Why-Questions and their Answers

We can understand explanations as answers to why-questions. As discussed in chapter 3(section 3.4 p42), the way in which a question is formulated is often a crucial determinant of what does and does not count as an appropriate answer to it. Thus it will worthwhile unpacking the some of the presumptions that are inherent in questions which inquire why something can persist through some changes and not others (hereafter 'persistence why-questions'), before we consider how to posit and evaluate answers to them[11].

What general observations can we make about 'persistence why-questions'? Quite a few, I think. To begin with all the 'persistence why-questions' with which we will be concerned can be taken to have the following form :

Q: Why is it the case that A can/cannot undergo transformation X without ceasing to be A?

The topic (P) of the question will be a proposition which asserts that a certain change is or is not possible for a given object. In all the cases which I will discuss it will be taken for granted (and usually obvious) that P is true. The next thing we need to decide is whether Q is asking:

1. Why can A undergo change X but not change Y whilst retaining its identity ?

or 2. Why it is that A can undergo change X rather than not?

11 In order to do this I will sometimes use technical terminology which although prevalent in a great deal of the literature on Explanation is often associated with relativism because van Fraassen coined most of these terms. See van Fraassen 1980 in Ruben 1993 p294-8.

I will take it that the contrast of interest in my examples, is the latter of these two. Thus the relevant contrast class[12] will consist of only one proposition 'not P'.

I also noted in section 3.4 that explanations and the questions to which they are answers involve a tacit appeal to relations which exist between the *explanandum* and the *explanans*. There must always be some perspicuous relation between the two in order for the one to explain the other. In practice there is sometimes more than one such relation, so it is helpful to specify which relation underwrites any given attempt at explanation. In the case of a persistence why-question, we are seeking a relation between the fact that something can or cannot undergo a particular kind of change and the fact that some of its features are necessary. That certain features of an object are necessary to it *entails* that changes which involve the taking on of characteristics which are incompatible with those necessary features will be impossible. It would appear that this is the relation that underlies the explanatory force of the necessary properties of objects with respect to questions concerning their persistence through change.

So we now know that relevant answers to our persistence why-questions will cite some or all of the appropriate object's necessary properties. Does the nature of Q tell us anything else of interest about the kind of answers we ought to expect and how they should be assessed? Appropriate answers to Q will take the following general form:

E: A can/cannot undergo transformation X without ceasing to be A because A is necessarily N.

This kind of claim, although usually perfectly satisfactory on its own, can perhaps be better understood as an abbreviated argument which goes something like the following:

1. Transformation X would involve the loss or gain of some property which is incompatible with A's being N.
2. A cannot become anything that is incompatible with its being N

Therefore 3. A can/cannot undergo transformation X without ceasing to be A.

The argument form is valid, and I will be assuming, on the basis of the conclusions of chapter one, that the premises are true. So, if the explanations furnished by a given necessary property or some sub-set of the necessary properties of a thing are bad ones it will not be because either of the ‹explanans;

12 A contrast class is the set of propositions which is made up of the topic (P) of particular why-question plus one or more additional propositions which represent non-actual alternatives to P. The question always presumes that P is true and that none of the other members of the contrast class is true. The truth of P is incompatible with the truth of any of the other members of the contrast class.

explanandum› pair is false. Hence we will have to make the requisite evaluations in terms of the criteria discussed in section 3.5 above. As I mentioned there, a popular way of evaluating an explanation is to consider whether the truth of the explanans make the state of affairs depicted in the explanandum more likely. However, this will not work for the kinds of explanation we are considering here. The reason for this is that at the outset, the probability that A can/cannot undergo transformation X is 1. We know, prior to any attempt at explanation, that Socrates cannot become a daisy or mince-meat without losing his identity. Thus no additional fact can make the state of affairs depicted by the explanandum more likely. The addition of the explanans to our set of beliefs will not affect the distribution of probabilities at all. An alternative way of evaluating explanations is to compare the scope, informativeness[13] and capacity to unify the data[14] of a range the potential explanations. The best explanation being the one which maximises scope, informativeness, unification or some combination of the three. It is on this kind of evaluation that I will rely in the analysis of examples that follows.

So much for the form of the questions that will be asked in the examples and the kinds of answer that can and ought to be given to them. Let us now look at an example.

4.3.2 Socrates and the Taj Mahal

Socrates could survive the loss of his youth, his tan and his ignorance but not the loss of his body[15], a change of species or the change from being animate to being inanimate. As I have indicated above, I think that this is because the latter kinds of change would involve the loss of properties that Socrates has necessarily[16]. Let us test this hypothesis with a specific question:

Q1: Why would Socrates not survive being turned into the Taj Mahal ?

It is uncontroversial that Socrates would not survive such a change. It is less obvious, however, why. So there does seem be a *bona fide* need for an explanation here. Can we meet this need by citing some or other of properties which Socrates has necessarily? It would seem that we can. Becoming

13 See Lewis 1986 p193-4 for a clear exposition of the ways in which explanations can fail to be maximally informative- by providing too much information, providing too little information, providing false information, providing contextually information or by failing to correct an existing state of misinformation.

14 See Kitchen 1981 for a discussion of the ability to unify the data as a virtue of explanations.

15 *Pace* Descartes.

16 The alternative would be that the result of a change of body or species would simply stop satisfying the concept we have of 'Socrates'. This does not seem right. If this were the case then the change would cause a failure of reference. The name 'Socrates' would no longer refer to anything as a result of the change. But given what has been established in chapter 1, namely that reference is determined by properties that objects have independently of how we think of them, failure of reference implies a change in the reference-determining/necessary properties of the relevant object. So ultimately some appeal will have to be made to these in order to explain the potential a given object has for change.

the Taj Mahal would involve becoming something which has necessary properties which are incompatible[17] with those of Socrates. So in explaining why Socrates cannot become the Taj Mahal, we need to cite those necessary properties of Socrates which cannot be co-instantiated with the properties which are characteristic of the Taj Mahal. It would not be explanatory to cite the fact that Socrates is necessarily not a daisy because 'not being a daisy' is a necessary property which Socrates and the Taj Mahal share. Thus the necessary properties of an object that can be relevantly cited in effective explanations of the object's persistence or failure to persist through a particular change, will be determined by the kind of change envisioned. However, the nature of the change concerned does not seem to be the only determinant of the suitability of particular necessary properties to the task of explaining persistence or the lack of it through change. There are some necessary properties of Socrates that are incompatible with the necessary properties of the Taj Mahal, and which also constitute minimally relevant answers to Q1, but which are nonetheless unsatisfactory explanations of the impossibility of this particular change. I think that the necessary properties of any object can be usefully categorised into four types :

1. Those that express the fact that whatever exists does not contradict any necessary truths
2. Those which state what the thing in question necessarily is not
3. Those that specify characteristics that the thing necessarily has, but which are not independent in the sense that they are entailed/determined by other necessary properties
4. Those properties which specify characteristics that the thing necessarily has and which do not seem to be derived from another 'antecedent' necessary property of the object concerned.

In what follows I will compare token answers to the Socrates question taken from each of these types. The results of the comparison will show that properties of type 4 generally constitute better explanations of Socrates inability to survive such a change than those drawn from any of the other types. The question (Q1) is 'Why is it that Socrates cannot become the Taj Mahal without ceasing to exist (losing his identity)? We might offer any of the following as answers:

1. Socrates cannot become the Taj Mahal because he is necessarily such that 2+2 = 4 (or such that all necessary truths hold)

2. Socrates cannot become the Taj Mahal because he is necessarily not the Taj Mahal (or not a building, or not inanimate etc.)

17 Whenever I talk of two properties being incompatible I will mean that it is impossible for them to be co-instantiated.

3. Socrates cannot become the Taj Mahal because he is necessarily mortal (or necessarily potentially conscious, or..... necessarily blue-eyed, galactose intolerant etc.)

4. Socrates cannot become the Taj Mahal because he is necessarily animate (or necessarily human etc.)

All four of the above statement types generate tokens which are true, but clearly the members of the last class of statements offer better answers to our original why-question, than do the members of the other three. Let us evaluate each explanation type in more detail.

One might wonder whether properties like those in the examples of type 1 statements given above, are even relevant to the question. After all, 'being such as not to contradict any necessary truths' is not a necessary property of Socrates which the Taj Mahal cannot have. On the contrary, 'being such as not to contradict any necessary truths' is a necessary property that is possessed everything which exists, including Socrates *and* the Taj Mahal. So it is a necessary property that Socrates and the Taj Mahal share. The reason I have included this type of property in this analysis is that there is a sense in which they are minimally relevant answers to Q1. Being in contradiction of a necessary truth is incompatible with *being at all*. There are no possible worlds in which anything which contradicts a necessary truth exists. So, any change which brought Socrates into contradiction with a necessary truth, would not only cause Socrates to go out of existence but, since the Taj Mahal exists, would be incompatible with the result of such a change being the Taj Mahal. Thus there is a minimal sense in which we can say that the type 1 properties of Socrates explain why he cannot become the Taj Mahal. Nevertheless, despite meeting the barest requirements of relevance, unsurprisingly, statements of type are usually not explanatory. With the notable exception of one particular necessary truth (which I will discuss at the end of this section) - there just is no illuminating connection between properties like being such that '2+2=4', 'being Socrates' and 'not being the Taj Mahal'. Furthermore, whatever relation does exist between 'being such that 2+2=4' and 'being Socrates' also exists between 'being such that 2+2=4' and every conceivable property. Because type 1 necessary properties are necessary properties of *everything* no necessary properties of this type will ever meet the requirement that they bring to our attention those aspects of Socrates which he cannot lose without ceasing to exist and which are also incompatible with his being the Taj Mahal. Therefore, type1 necessary properties will almost never be of any assistance in the attempt to give explanations of the ways in which a thing can or cannot change. All answers to persistence type questions which cite this type of necessary property will be equally bad.

The appeal to the fact that Socrates is necessarily 'not the Taj Mahal' or necessarily 'not a building' looks a little more promising, but not much. Answers of this type explicitly cite properties which the

Taj Mahal instantiates necessarily but which are incompatible with some or all of the necessary properties of Socrates. However, the claim that Socrates is necessarily not the Taj Mahal does not tell us anything that we have not already assumed if we think that the question arises at all. More particularly, it does not distinguish those aspects of being Socrates which are compatible with being the Taj Mahal (being a material entity for example) from those which are not. The claim that Socrates is necessarily not a building fares somewhat better on this score. It isolates a feature which the Taj Mahal must have but which Socrates cannot have. However, even this is unsatisfying. Such an explanation would still be incomplete - one wants to know what it is about buildings that is incompatible with their instantiating some or all of the necessary properties of Socrates and why. We would seem to come closer to a satisfying answer to Q1 by citing the fact that Socrates is necessarily not inanimate. This is certainly more informative than the previous two options and the question : "what is it about being inanimate that is incompatible with being Socrates ?" has an obvious answer i.e. that Socrates is necessarily animate. Note however, that in invoking this 'obvious' answer we are making an appeal to another property which Socrates has necessarily - and one that falls into a different category in the above classification. This ought to suggest to us that the specific property to which the appeal is made and perhaps even the category to which it belongs, is somehow more explanatorily fundamental than the negative property to which it corresponds.

Having said this, the fact that we seem to have gotten increasingly closer to a satisfactory answer by selecting different necessary properties within the 'type 2' category, deserves comment. There appears to exist within this category, a natural hierarchy of explanation. Recall the fact that Socrates is necessarily not the Taj Mahal was explanatory because The Taj Mahal is a building and Socrates is not and the fact that Socrates is necessarily not a building was explanatory *because* buildings are inanimate and Socrates is not. Note however, the reverse explanation does not work. We cannot account for the explanatory relevance of being necessarily inanimate in terms which make use of Socrates' necessarily 'not being a building' or necessarily 'not being the Taj Mahal'. So, some of the properties in this category act as the *reasons* that other properties in the same category can explain Socrates failure to survive transformation into the Taj Mahal. The most fundamental reasons of this type will be those properties whose explanatory relevance cannot itself be explained in terms of some other type 2 property[18]. The more fundamental the reason a particular necessary property represents the greater will be its explanatory significance in general.

What about the appeal to the necessity of Socrates being mortal? Statements of type 3, although initially very weird looking are, in a sense, more explanatory than those of type 2. They do after all,

18 I intend, here to be defining the fundamental as the point at which no further explanation can be given.

cite 'positive' or intrinsic properties of Socrates which cannot co-exist with the property of 'being the Taj Mahal'. Take for example, 'being necessarily mortal', presumably buildings, since they are necessarily not alive, cannot die and thus cannot be mortal. Thus, Socrates cannot become the Taj Mahal *because* he cannot become an, *in principle* non-living thing, without becoming something which is not Socrates in the process. Answers of this type are informative, but there is something strange about them nonetheless. Although the information provided is sufficient to differentiate Socrates and the Taj Mahal effectively, the explanation still does not seem to be complete. To try and understand this lack of completeness let us consider an extreme example. Imagine Socrates necessarily has the genetic material which he did in fact have and that included in this genetic material was the genotype characteristic of blue-eyed people. In such a case Socrates would be necessarily blue-eyed. Now being blue-eyed is a property which Socrates must lose in order to become the Taj Mahal - which is an in principal non-eyed thing. But he cannot lose this property and retain his identity. So 'being necessarily blue-eyed' ought to go some way toward explaining the fact that Socrates cannot become the Taj Mahal without a loss of identity. But does it? Only in the most minimal fashion. In order for it to constitute an explanation a whole lot of connections have to be made between the necessary property of being blue-eyed and other necessary properties of Socrates which explain *why* Socrates is that kind of thing that has eyes of a certain colour and the Taj Mahal is not. So the relation between being necessarily blue-eyed and not being the Taj Mahal is neither as perspicuous or as simple as we require our relations of explanatory relevance to be. The same is true of the relation between any type 3 property and 'not being the Taj Mahal' etc. More importantly, once again that fact that type 3 properties are explanatory of Socrates' persistence through change at all, is contingent on the existence of other more general necessary properties of Socrates. Further, whilst the existence of these general properties has to be assumed in order to make explanations in terms of type 3 necessary properties make sense, the converse is not true. General properties like 'necessarily being of the kind of thing which necessarily has eyes' or more simply, 'necessarily being an animal' do not require us to make any additional assumptions in order to explain Socrates' failure to persist through transformation into the Taj Mahal.

From the discussion of the answers to Q1 afforded by necessary properties of types 1,2 and 3 it should be clear that some of the necessary properties of an object imply or entail others but are not themselves implied by any other property the object has. It is the existence of this kind of necessary property that 'type 4' is intended to represent and which I hope to have captured in the token type 4 answers listed above. The properties I have used in the examples are those that I am guessing are most likely to play this role. I may be wrong, about which actual properties are type 4 properties.

However, what the answers I have suggested do have in their favour is that at the level of common-sense they just seem so much more adequate than any of the alternatives.

From the above discussion it should also be clear that those necessary properties which entail but are not entailed by other necessary properties seem to be the ones suited to furnishing explanatory answers to Q1. We saw in connection with type 2 and type 3 properties that an asymmetric explanatory relation exists between some necessary properties of Socrates and others. Socrates' mortality constitutes something like an explanation of his failure to persist through transformation into the Taj Mahal only *because* it indicates that Socrates must also be alive. So the explanatory power of the former necessary property is derived from the explanatory power of the latter. I think in general it is true that whatever explanatory value necessary properties of types 1,2 and 3 have with respect to the persistence of objects through change, is derived from the relations they bear to type 4 properties.

It is interesting that this relation is co-extensional with an inferential relation between necessary properties. The truth of statements that cite properties of type 4 makes statements which cite properties of types 1,2 and 3 true but not vice versa. I know that Socrates is a mortal non-building *because* I know that he is human. If on the other hand, all I knew about the thing called Socrates was that it is a two-legged, non-building I could not infer from this that Socrates is human - he might after all be an ostrich or a Martian. In consequence, these properties do not afford a reason to believe that Socrates is human. Thus there is a unidirectional relation of inferential dependence between type 4 necessary properties those and of types 2 and 3 - properties of types 2 and 3 being determined or entailed by those of type 4. This is one way to account for the apparent explanatory superiority of type 4 properties in the Socrates thought experiment. I think it is the correct one, however, the success of the above argument for the explanatory significance of type 4 properties does not rely on the truth of this particular account.

I noted in connection with type 2 properties that some answers to our persistence why question, such as that Socrates is not a building, seem to be open to further probing - questions directed at acquiring more specific information about the nature of the incompatibility of 'being Socrates' with 'being the Taj Mahal'. I also noted that there is a point at which these kind of probing questions cease to arise. It is no coincidence that this occurs when (and only when) appeal is made to a type 4 property. Only then do we consider ourselves to have a complete explanation. So we could say that explanations of type 4 are generally more informative than those of any of the other types. Furthermore, type 4 necessary properties provide explanations which include the kinds of explanations offered by necessary properties of types 2 and 3 within their scope, but which at the same time explain more. This must surely entail

both that explanations in terms of type 4 properties are better able to unify the data than those in terms of the necessary properties that fall into types 1-3 and that they make explanations in terms of other necessary properties irrelevant. Considerations such as these suggest that indeed, necessary properties of type 4 are explanatorily fundamental.

Thus, for several reasons, it would appear that providing a maximally explanatory answer to Q1 will involve the citing of a type 4 necessary property of Socrates. I want to suggest that this conclusion can be generalised to cover all persistence why questions which have the form of Q. That is, I believe that for all such questions, necessary properties of types 1-3 will never be in a position to provide better/ more explanatory answers than those of type 4. In the section that follows (4.4) I will defend the claim that this generalisation is a justifiable.

4.4 Generalising the Claims to Explanatory Superiority of Type 4 Necessary Properties

Thus far, it has been established that we can answer persistence questions like Q1 in a way which is explanatory, by citing some sub-set of the necessary properties of the object that persists. In the previous section it was also shown that a particular sub-set of Socrates' necessary properties provide better - more exhaustive, more unifying and more informative - answers to Q1 than do the necessary properties of Socrates which fall outside this sub-set (called type 4). It was characteristic of the necessary properties of Socrates which fall within type 4 that, taken together, they entail the other necessary properties of Socrates but are themselves entailed by none of the necessary properties which fall outside type 4 and the existence of this entailment would seem to underwrite their claim to explanatory significance. I also pointed out that type 4 necessary properties seem to be explanatorily fundamental in the sense that the persistence why-question to which the citation of necessary properties is a response continues to arise in ever more specific forms until a type four necessary property is cited. By contrast, once we have cited a such a property no further explanation (at least in terms of necessary properties) seems possible or appropriate.

In this section I will argue that the conclusions I have drawn with respect to the explanatory significance of type 4 properties for Q1 are generalisable across most objects and for any persistence why-question we care to set up. I will do so by briefly considering some examples of persistence why-questions involving different kinds of object (4.4.1) and different kinds of change(4.4.2). I will also consider some purported counter examples to the generalisation I want to make(4.4.3) and argue that they are weak and need not prevent us from making the relevant generalisation within the context of metaphysical enquiry.

4.4.1 Generalisation Across Object Types

In this section, I want to sketch some additional examples all of which will add support to my contention that the conclusions reached here are generalisable across all continuants and possible persistence why-questions. Case 2 will illustrate that an analysis similar to the one set out for Socrates above can be given for persistence why-questions which ask why a particular change *is* possible. In case 3, I will examine an example of a persistence question of which the subject is a substance or stuff and in case 4, I will consider how certain necessary properties of an artefact appear to constrain the ways in which it can and cannot change. In the latter two cases, I will use the negative form of the persistence question as a matter of convenience - as with the above case a similar analysis should be applicable where a positive persistence question is asked. For all three examples, I will only sketch how I think the analysis would go. It will not be necessary to go through the kind of detailed analysis set out in the Socrates case, because in all three cases, it will be obvious that the broad strokes required will be the same as those presented above.

Case 2: Metamorphosis

One example of a remarkable transformation without identity loss is the kind metamorphosis which occurs during the life of most insects. Can we explain this kind of radical by identity-preserving change in an insect by citing one or some of the insect's necessary properties? Take, for example the metamorphosis of a particular silkworm (call him Fred) into a moth. I will take the basic why question (Q2) to be 'Why can Fred become a moth and yet remain the very same thing he was to begin with?' We have the familiar range of answers to chose from:

5. Fred can survive metamorphosis into a moth because he is necessarily such that he does not contradict any necessary truths

6. Fred can survive metamorphosis into a moth because he is necessarily not something which cannot survive such a metamorphosis (or he is necessarily not human, a building, the Taj Mahal etc.)

7. Fred can survive metamorphosis into a moth because he is necessarily an herbivorous, silk cocoon spinner (or necessarily has any other property entailed by the fact that he belongs to the species he does)

8. Fred can survive metamorphosis into a moth because he is necessarily an insect of a particular species.

Once again, I think it is clear that only statements of type 8, come close to constituting an explanation of Fred's survival of such a radical change. Fred survives this metamorphosis because he is an insect of a particular species and it in the nature of such insects that they go from being worms to being

moths. Furthermore, I think the reasons that the other statement types amount to less than satisfactory explanations of the phenomenon or to no explanation at all, are largely the same as they were for case 1. An appeal to the non-contradiction of necessary truths is generally not explanatory in cases where part of what is required in answering the relevant question, is an account the contrast that is implied by the question. What such an answer does tell us, is that there is no necessary law which prevents such a transformation - which is minimally informative. Nonetheless since necessary truths bear the same relation to all existents, such an answer does not identify those necessary properties of Fred the worm which he shares with Fred the moth - and this is what an answer to Q2 must do in order to be properly informative. Explanations of type 6 seem even less explanatory than their counterparts in case 1 were, although, once again we get a progression towards increasingly satisfactory answers to Q2 as we move closer to negations of those necessary properties in virtue of which the change is possible. Explanations of type 7, although they seem to be getting closer, are only minimally informative on their own. Once again, being an herbivorous silk-cocoon spinner will only be an explanation of Fred's ability to survive metamorphosis if we also know that all herbivorous silk-cocoon are insects and that insects by definition are creatures that undergo metamorphosis. We could try offering a long list of Fred's type 3 properties. This list could, however, fall foul of the same problem. In addition, the most accurate, exhaustive such list would be co-extensive with properties of type 8 anyway. In which case a clearer, more economical explanation could be given by citing the relevant properties of type 8. Finally, perhaps the most important consideration of all is that explanations of types 5-7 all depend for their explanatory efficacy on the existence and explanatory power of necessary properties of type 8. By now, it should be clear that explanations of type 8 will both be more informative than any of the others and render them wholly irrelevant.

Case 3: Substances

Now let us consider an example of a persistence why-question which concerns a stuff - sodium chloride or salt. We might ask why a given molecule of salt (Bob) does not survive being dissolved in water (Q3). Again there are a range of possible answers:

9. ....because Bob is necessarily such that it does not contradict any necessary truths

10. ....because Bob is necessarily not insoluble in water (or necessarily not a molecule of alcohol or a compound of carbon etc.)

11. ....because Bob is necessarily a stable compound (or necessarily has the same number of electrons as an atom of Nickel etc.)

12. ....because Bob is necessarily an atom of chlorine bonded to an atom of sodium in a certain way.

By contrast with the complex objects which we have been considering up till now, molecules have relatively few necessary properties. Nonetheless, the fact that necessary properties of each of our four categories set out above can be cited in response to Q3, indicates that those they do have are organised according to the same framework that we identified above, with some properties playing a more significant role in explanations of persistence than others. If we consider statements 9-12, it is clear that the kind of property cited in 12 offers the best explanation. Once again, the first option (9) whilst true, is utterly unenlightening. And once again, considerations of scope, informativeness and explanatory dependence favour 11 over 10. Answers which cite properties like being necessarily not insoluble or not a carbon compound or alcohol molecule, fail to tell us *what it is* about salt that makes its behaviour in water necessarily different to that of an alcohol or any other carbon compound. Answers which cite properties like those given in statement 11 above suffer from exactly the same defects as did those statement which cited type 3 properties above. Most notably, they require the invocation of type 4(12) properties in order to be explanatory. In this case the set of properties which do furnish satisfying explanations is confined to a few specific structural properties. Note that statements of like statement 12, whilst part of an adequate explanation is not complete. For that, we would probably also have to make an appeal to the chemical structure of water and possibly even to the way in which molecules of salt and water interact. Nonetheless, it seems clear that the molecular structure of salt has an important part to play in the required explanation whereas the necessary properties cited in statements 9,10 and 11 do not. Similar arguments can be constructed for the explanatory primacy of the chemical structure of any stuff.

Case 4: Artefacts

Moving on to artefacts, the fourth example I want to look at here concerns the ways in which a sophisticated artefact - a clock - cannot change without ceasing to exist. Unlike any of the entities previously considered a clock (and most other artefacts) can survive disassembly. Given this, can there be any kind of change that a clock would not survive? Indeed there is. It is uncontroversial that if the whole clock was melted in a furnace or put through a metal shredder, it would cease to exist. But why? Let us take this question as our fourth test case - Q4: Why is it that a given clock cannot survive being shredded or melted? Again, we have the usual battery of possible answers:

13. .....because the clock is necessarily such that that it does not contradict any necessary truths.

14. ... ..because the clock is necessarily not shreddable (or not meltable or not something which will never be able to tell the time....etc.).

15. .....because the clock is necessarily something that has the potential to tell the time.

Note that there is nothing which corresponds to the statements which cite type 3 necessary properties (3,7 and11) in this case. This is because artefacts like clocks have simple definitions which cite one specific feature. Unlike members of natural kinds who are defined in terms of their group membership artefacts can be directly defined by a single necessary property. Often but not always, the relevant property is their function. As a result, there is only one type 4 property in this case and this property has no interesting entailments. It follows that there are no likely type 3 properties which might feature in an answer to Q4.

Let us now consider the quality of the explanations furnished by statements 13-15. Yet again, statements of type 15 seem to boast superior explanatory power. As with the corresponding statement above (1,5 & 9), statements of type 13 are uninformative. Statements of type 14 vary in informativeness. To say that the clock is necessarily not shreddable is state the very fact for which we seek an explanation when we pose Q4. At the other extreme of this type answer lies the claim that our clock is necessarily not something which will never be capable of either measuring, recording or registering the time. This comes pretty close to being synonymous with the answer given in statement 15 which I have said provides the best explanation of our clock's failure to persist through shredding. What makes 15 better is simply the fact that is characterises the state of affairs described by both 14(b) and 15 positively and in doing so defines the object concerned - our clock. If we asked why our clock is necessarily not something which will never be capable of either measuring, recording or registering the time, a perfectly natural answer to give would be that measuring, recording or registering the time is what clocks do - and it is this set of activities which defines what they are. What we would be doing here is translating statement 14 into its positive counterpart 15. So the positive definition of an artefact provides a natural end point beyond which it does not make sense to ask why. In addition an explanation which cites the necessary properties which feature in the positive definition will include all other possible explanations within its scope. Thus I think it is safe to conclude that from the perspective of the metaphysician, the necessary properties which feature in the positive definition of an artefact provide the best explanations for the ways in which it can and cannot change whilst preserving its identity.

Before I move on, I would like to remark briefly on something which the application of this kind of analysis to artefacts brings out. It has been argued that the essences of artefacts are merely nominal (as opposed to real). If this were the case, how would it affect what I have said in this last section? Surprisingly, it would not affect what I have said at all. The kind of analysis applied here would elicit the same categories of necessary properties with different degrees of explanatory significance,

regardless of whether we take the relevant properties to be nominal or real. If, as I want to argue, the essence of an object consists of its type 4 necessary properties, what then what would be the difference between nominal and real essences? The difference between these two lies in whether we understand the relevant essential properties to be features of the objects themselves or features of the concepts we use to think and talk about them. The arguments of chapter one of this dissertation were intended to establish that some objects have real necessary properties - necessary properties which are features of the objects themselves. This taken together with the results the above analyses should indicate that some objects have amongst their real necessary properties certain properties which are of a superior explanatory status. If we took essences to be those necessary properties which have maximal explanatory value, then what we would have in this dissertation is the sketch of an argument for the existence of real essence which look a lot more like Aristotelian essences than like the weak Kripkean essences we started out with. In the sections that follow I will try to convince you that we should in fact take the essences of objects to be those of their necessary properties which manifest superior explanatory power.

4.4.2 Generalisation Across Different Kinds of Change

If the trends of the previous section are anything to go by, it looks as if my analysis is generalisable across a wide range objects. But do my conclusions with regard to the explanatory superiority of type 4 properties generalise across questions concerning all possible transformations? In what follows I will consider some different transformation possibilities. I will demonstrate that type 4 properties prove superior in a fairly wide range of cases and will suggest that this provides us with some (but not complete) justification for generalising the conclusions of section 4.3 across different kinds of transformation.

Consider Socrates (again), instead of asking why he cannot become the Taj Mahal, we could have asked why he cannot be transformed into any other thing - a tomato, another person, a mixture of air and fire or mince-meat for example. Let us take a more detailed look at some of these possibilities. Once again the analyses presented here will be mere sketches as I expect the details to be the same as they were in 4.3.2.

Trans-Species Mutation?

Socrates could not become a tomato because he is necessarily not a vegetable or because he is necessarily two-legged or because he is necessarily human. The last of these options seems like the best because it captures everything about Socrates that is incompatible with his being a tomato, thereby making the other possible answers substantially irrelevant; it includes properties such as his two leggedness and his non-vegetableness within its scope and thereby unifies the data; and it is

explanatorily fundamental - we can ask and answer the questions 'Why he is necessarily not a vegetable or necessarily two-legged'. We cannot expect an answer to the question 'Why is Socrates necessarily human?' - there just isn't one.

Incineration

The prospect of Socrates becoming a mixture of air and fire is a slightly more interesting case because we can actually conceive of Socrates being set on fire. At some point in the burning process there would be nothing left but a mixture of air, fire and ash. Would this be Socrates? Clearly not - and the reason why this is true is that mixtures of air, fire and ash are not human and Socrates is and necessarily so. Again there are other possible answers. For example, we could say that Socrates is necessarily not something which survives very high temperatures or not something which is partially invisible. Alternatively we could attempt an explanation by saying that Socrates necessarily has a mind or a means of pumping blood around his body etc. Again the latter alternatives, *along with much else that is relevant,* are captured by citing the fact that Socrates is necessarily human and are all the non-type 4 properties cited are explicable in terms of his necessary humanity.

Note that in this particular context of enquiry there is an alternative type 4 property which might serve the purpose of providing an appropriate explanations just as well as 'necessary humanity' - this is the fact that Socrates is necessarily 'an individual thing' (as opposed to a mixture or collection of things). We can explain Socrates failure to survive incineration (or mincing for that matter) by citing necessary properties of Socrates which reflect the most general category to which he belongs - 'concrete particular' or Aristotelian 'substance'. It would seem that for objects whose type 4 properties include those which demarcate kind membership, there will always be other more general necessary properties which reflect their genera and their ultimate category membership which can also be used to provide explanations of their persistence through some changes but not others. How do we decide whether the more specific kind-membership properties provide better explanations than the more general genera or category membership properties? I think the level of generality of the change in question determines the level of generality of the answer. Incineration is what I would call a very general kind of change. When something burns up, what was concrete particular becomes something other than a concrete particular. And this is true regardless of the kind to which the original concrete particular belongs. An explanation which cites ultimate category membership seems required here in order to fully understand the nature of the transformation we are considering. So in our example, that Socrates is necessarily an individual thing offers a marginally better explanation than the claim that he is necessarily human. By contrast, when something migrates across species its particularity does not change and consequently an explanation which cites ultimate category membership would be inappropriate.

Inter-Personal Transmutation

Consider one last example, Socrates cannot become another person - say Plato or Xanthippe. Why? Obviously, we cannot explain this impossibility by invoking the necessary property of 'being human'. How might we explain this kind disinclination to change? Interestingly enough, the ancients thought that what differentiated individual members of a particular species was their 'matter'[19]. The important thing about the matter of any material body is that it can be replaced during the lifetime of the entity concerned. In modern terminology, those properties we have which pertain to the stuff of which we are made are contingent rather than necessary. Nonetheless, the *de dicto* necessary principle that two spatially separate things cannot be constituted from the same bit of matter still holds. So, Socrates could not become Plato nor Plato, Socrates because although neither of them is essentially different from the other, it remains a necessary (*de dicto*) fact that they are each different chunks of matter. A different way of formulating this claim might be to say that Socrates cannot become Plato because his doing so would contradict a necessary truth - namely that any given time, two spatially different entities cannot have material constituents in common[20]. There is however, another alternative for we who inhabit the age of science. We might think of the genotype of an organism as determining all the other necessary properties it has. Since genotype thus understood would then entail such properties without itself being entailed by any of them, it would classify as a type 4 property. If this is correct then, we could say that the relevant type 4 property that should be invoked in this case is that Socrates necessarily has genotype X and his becoming Plato would involve an alteration in this genotype.

I am not sure about the viability of the genotype alternative, either from a genetic or from a metaphysical perspective. I prefer the ancient view, which entails that there is indeed a species of persistence question to which necessary properties cannot constitute an appropriate answer. So there *is* one kind of persistence question to which my claims about the explanatory significance of type 4 necessary properties do not apply. However, this need not be too much a problem for my thesis. I do not think that the very idea of persistence or the lack of it through inter-personal transmigration is a notoriously problematic notion in its own right. And it is one which cannot be made sense of without giving due consideration to the very difficult and fraught debate over the possibility necessary properties that are *in principle* unique to individual objects. I do not want to enter that debate here.

19 See Aristotle's Metaphysics VII, 8 1034a5-9.

20 Given this, it looks as if my earlier claim that 'necessarily not contradicting a necessary truth' is never explanatory needs to be revised. It would seem that necessary truths can be divided into two different kinds according to what they are about : those few that are about things such as matter, identity perhaps mereology, all of which might be relevant to questions concerning the persistence through change of material objects; and those which are blatantly irrelevant such as necessarily being such that 2+2=4 etc. The former may be thought of as explanatory. However, I suspect that in giving such explanations we are really appealing to the relevant *de dicto* necessary truth of which the *de re* necessary property is a consequence. So once again, these would be cases which fall outside the scope of questions of the form Q because they incorporate a different relevance relation.

However, if it turns out that no properties can be *in principle* unique to individual objects (as I suspect it must), then the differences between individuals cannot be accounted for in terms of their properties. If individual persons are not individuated my means of their properties then it is hardly fair to expect explanations of the inevitable failure of trans-personal mutation to be couched in terms of any their necessary properties, never mind their type 4 necessary properties. Aristotle, of course, believed that there could not be such properties and he has found a modern defender in David Wiggins[21]. If Aristotle and Wiggins are correct then, we can ignore persistence why questions which concern the impossibility of one thing becoming another thing of the same kind. I will just assume that this is the case.

I hope, this brief discussion of some pivotal examples has convinced the reader that we, with a little care, generalise the conclusions of the previous section across most why questions. In other words, that the superior explanatory significance of type four necessary four properties remains intact regardless of the kinds of change we consider with the single exception of changes from one member of a species to another member of the same species - (and even here, no explanation in terms of necessary properties is possible so the question of which types of necessary property provide better explanations does not arise). I now want to move on to consider the metaphysical implications of the much-discussed explanatory abilities of type four necessary properties.

4.4.3 Some Possible Counter Examples

Thus far I have established that my conclusions from section 4.3, although not universally generalisable, apply across a fairly broad range of objects and types of change. In this section I want to approach the question of the degree to which the conclusions of 4.3 can be generalised, from a different angle. I want to consider possible counter-examples. We have already encountered one possible counter-example. However, I hope I have managed to convince the reader that that this case was sufficiently unusual that it ought not to be allowed to de-rail the project entirely. Let us now consider some more serious problems for my view. On reflection, I think there are two possible ways of objecting to the claim that type 4 necessary properties provide better explanations of persistence/ failure to persist through change than any other kind of necessary property. The first would be to deny that citing necessary properties is explanatory at all. In the light of everything that has been discussed in this chapter, I cannot see how such a claim could be very persuasive. The second possible objection however carries more force. In accordance with what was said in the previous chapter, I have focused on objective explanatory worth - my evaluations of the quality of the explanations offered by the four possible types of answer invoked explanatory virtues such as width of scope and

21 See Aristotle, Metaphysics VII, 8 1033b20-5 and Wiggins 1980 117-124

informativeness. But why should I be concerned with these rather than more pragmatic values such as how the explanation given fits in with the desires and intentions and beliefs of prospective enquirers? Could it not be objected that if we were to take the pragmatic value of explanations into account, necessary properties of types other than type 4 would, in certain context, constitute the best/most explanatory answers to a persistence why-question. Consider Q1 again. Now say, for example, the person who seeks an explanation of Socrates failure to persist through a transformation into the Taj Mahal is an astrologer. He not only believes that a causal relation exists between the constellations of the planets and the natures and fortunes of individual human beings, he also believes that the sole dictate of the constellation of the planets at the exact time at which Socrates came into existence, was that whatever came into existence at that time would not be the Taj Mahal. In such a case, surely the best of our four kinds of answer to Q1 to offer such a person would be that Socrates is necessarily not the Taj Mahal - i.e. an explanation is terms of a type 2 property.

The obvious problem with this scenario is that it takes more than that somebody finds it convincing or satisfying for something to constitute an explanation at all. According to the account of explanation we accepted in chapter 3, for any statement (P) to be an explanation, never mind a good one, - the following conditions must be met:

1. P must be true
2. The relation which the explanation presumes must exist (or we must at least have good reason to believe that it does)

Our astrology example satisfies the first of these conditions, we have seen above that it is true that Socrates is necessarily not the Taj Mahal. However, the relation presumed by the astrology example is not the simple logical entailment relation on which our explanations up till now have relied. The relation between Socrates' inability to persist through transformation into the Taj Mahal and his having the necessary property of not being the Taj Mahal, on which the astrological explanation relies, is one mediated by an assumed connection between the constellations of the planets and the natures and fortunes of individual human beings. Firstly, I imagine the astrologer will have difficulty getting metaphysicians to accept that such a relation exists or even that we have rational grounds for believing that it does. Secondly, even if we allow that it does exist, explanations of the kind with which we have been concerned here might still be preferred because the relation on which they are predicated is simpler. Finally, the set of relations that are allowed to underwrite explanations in metaphysics will be circumscribed by a presumption in favour of the view that much of our experience can be accounted for in its own terms and a corresponding suspicion of anything that requires an appeal to some kind of *deus ex machina*.

The above example makes the appeal to pragmatic values seem far-fetched. Are there perhaps less tendentious examples where the context in which a persistence why-question is asked makes the citation of a necessary property from one of types 1-3 a better/more explanatory answer to the question? Perhaps there are. However, I am tempted to claim that the kind of persistence why-questions we have been considering only arises in two possible contexts - those of analytic metaphysics and theology. I would be happy to restrict the generalisation of my claims about the significance of type 4 necessary properties to within the context of analytic metaphysics, should the pragmatist insist that this is necessary. So, my claim can be restated thus: type 4 necessary properties offer better metaphysical explanations of persistence phenomena than do necessary properties of any other type. I can think of no counter-examples apart from the one discussed in section 4.4.1 above.

4.5 The Metaphysical Significance of Type 4 Necessary Properties

Can we now conclude that type 4 properties represent a genuine fundamental metaphysical category such as essences must be? Well, before we do we have to be sure that explanatory significance entail metaphysical significance - or at least that it does in this particular case. In section 3.1 of the previous chapter, I claimed that the connection between explanatory value and metaphysical fundamentally is clear and simple. That postulates of metaphysics, like the concept of essence, just are explanatory principles. If this is correct then we ought to understand type four necessary properties as genuinely delineating a metaphysically significant category. Based on the way in which my analysis has been constructed there are two things that we can say about this category:

1. It is predicable of all objects which can be referred to using proper names
2. It will have as its extension the sets of necessary properties of individual objects which best explain the ways in which those objects can and cannot change.

I want to suggest that the metaphysically significant category or notion that type four properties define is that of essence. Earlier I suggested that there are at least two necessary conditions for being essential. All essential properties are properties the relevant objects have in all possible worlds in which they exist and all essential properties are possessed of the power to explain various metaphysical phenomena associated with the existence and characteristics of concrete particulars or (in technical sense) substances. What I think the conclusions of this chapter suggest is that of all the necessary properties instantiated by objects, only type 4 necessary properties satisfy the second necessary condition for being essential. Unsurprisingly, the picture of essence developed in this thesis looks

remarkably like the Aristotelian conception of essence. The picture I have drawn of essence is not as sophisticated or as intricate as Aristotle's but what I do hope to have shown is that from a modest Kripkean starting point we have the resources to develop a conception of essence which is rich enough to play a meaningful role in metaphysical explanations and theories. Furthermore I hope that it is clear to the reader that if we accept Kripke and Putnam's arguments in favour of the existence of mind-independent necessary properties, we are also committed to the existence of mind-independent essences of the type I have depicted here.

4.6 Summary

It was argued in previous chapters that objects have mind-independent necessary properties but that being instantiated necessarily is not sufficient to make a property essential. It was then suggested that de *re* necessity and the capacity to explain certain metaphysical phenomena might together be sufficient for a particular property's being essential. In this chapter I demonstrated that there is an identifiable, describable type of necessary property which provides complete/superior explanations of one of the above-mentioned metaphysical phenomena - persistence through change. I then suggested that these and only these satisfy both the necessary conditions for being essential.

CONCLUSION

The current tendency to define the essential as a set of properties which some object must have in all possible world in which it exists, has some intuitive appeal as well as the advantage of making essence something which can be discussed and argued about within the confines of modal logic. If this is really what essence is, then essentialism does not fall prey to the objection that we cannot make reliable inferences from or to essentialist claims since such claims are non-extensional. But the greatest contribution made by theories which understand essence this way has been the manner in which they have enabled their adherents to fend off the opponents of realism (as opposed to nominalism) with respect to essence. Such theories have been able to achieve this defensive success, surprisingly enough, because of their links to and origins in the philosophy of language. However, the association of essentialism with the philosophy of language has had disadvantages as well. We have come to think of essence as a logical notion (like necessity) rather than as a metaphysical one. As a result we have failed to fully appreciate the significance of the fact that essences are (or at least can be) *real.* This in turn leads us to ignore the common sense expectation that essences be ontologically fundamental. The result has been an impoverished conception of essence and a doctrine of essentialism which has 'lost its teeth'.

Attempts to employ this new 'understanding' of essence in the service of other theories and explanations have sometimes produced bizarre results — results sufficiently counter-intuitive that they arouse suspicion about all the presumptions on which they are based. So there is a sense in which the conception of essence favoured by the late twentieth century essentialists seems to undermine itself. Such results have led some philosophers to reconsider the original inference from facts about reference to facts about reality. These philosophers suspect that despite their apparent acceptability, such inferences must contain some unjustifiable hidden assumption or some logical mistake. I believe the modern understanding of essence *is* flawed, but that hidden assumptions and logical errors are the least of its problems. In this dissertation I have attempted to show that the fundamental mistake made by the new essentialist is to assimilate 'the essential' and 'the necessary'. I have argued here that *de re* necessity, whilst necessary, is not sufficient for a property's being essential. I have suggested that essence is an inherently explanatory notion and that the requirement that the properties which constitute a thing's essence be able to explain certain metaphysical phenomena associated with that thing, taken in conjunction with the requirement that the relevant properties be necessary properties of the thing concerned, better captures what it means for something to be essential.

To defend my position I have had to argue that objects do indeed have necessary properties independently of how anyone chooses to think of them. To this end chapter 1 of this dissertation sets out the theory of direct reference, defends its claim to be correct and then goes on to show how the metaphysical proposition that objects have necessary properties is entailed by it. Having thus confirmed that objects have necessary properties in their own right, in chapter 2 I explored what is wrong with treating necessary properties as essential and attempted to support my assertion that *de re* necessity is not sufficient for essence. I did so by comparing the consequences of taking the essential properties of an object to be 'the set of properties which it has necessarily' with an understanding of essence which finds application in our speculations about reality and the objects in it. I demonstrated there that the modern understanding of essence is over-inclusive — it has the consequence that a range of properties which most of us would not be content to accept as essential, are in fact necessary properties of their respective objects. So whilst all the essential properties of an object (a) are necessary properties of a, the converse is not true[1].

This need not, however, signal the end of the road for essentialism. At the end of chapter 2, I suggested that a better conception of essence could be produced and defended if we could identify some means of distinguishing those necessary properties that are essential from those that are not. Such a conception of essence would shed the problem of over-inclusiveness whilst retaining the realism associated with the necessary properties of objects that were defended in chapter 1. It was at this point in the project that the suggestion that essences are explanatory came into play. The idea was that we could use the requirement that essences be able to furnish certain sorts of metaphysical explanation as the above-mentioned means of distinguishing those necessary properties of objects which are genuinely essential from those that are not.

This attempt to use the explanatory power of essence to distinguish essential necessary properties from non-essential necessary properties would, however, only succeed if the capacity to provide the relevant explanations is an intrinsic feature of certain of the necessary properties of an object and not others. This is not an uncontroversial assumption. There were two steps involved in defending it. First, in chapter 3, I provided a defence of the general claim that some facts, propositions or concepts can and do have intrinsic explanatory worth with respect to a particular issue, independently of the characteristics of the audience at which a particular explanation is directed. Then, in chapter 4, I set out to demonstrate that there is indeed a small sub-set of each thing's necessary properties which has

---

1 If Pxy means x is a property of y and Exy means x is essential to y and Nxy means x is necessary to y then my claim above can be expressed in logical format as follows :

$$\forall x\,[\,(Px\,a \hat{E} Exa) \rightarrow Nxa] \quad \text{But not} \quad \forall x\,[(Px\,a \hat{E} Nxa) \rightarrow Exa]$$

notably greater explanatory power/significance than the rest and that, predictably enough, this sub-set does not include the weird kinds of necessary properties discussed in chapter 2. In order to show all this, I focused on the capacity of necessary properties to provide answers to questions concerning the persistence of objects through change. Essences ought, however, to be involved in the provision of answers to other questions of a similar ilk including things such as: *'what is it about an object that makes it one thing rather than many*?' and *'what distinguishes continuants from 'beings' that do not persist through time such as events and processes?*'. One way of testing my account of essences as those necessary properties of objects which fall into what I called type 4, would be to test the ability of type 4 properties to provide answers to such questions. If such tests indicate that the view that explanatory significance is an intrinsic feature of essence is defensible, then we will be free to bring explanatory considerations into the picture in our deliberations about the notion. I believe that the result will be a much richer understanding of the concept of essence and one which is much closer than the 'possible-worlds' conception to our common-sense ideas on the subject. Note that such an understanding of essence would be achieved without rejecting the primary thesis of the new-essentialists which is that objects themselves can have necessary properties. So we end up with a conception of essence which is very similar to the way in which it was understood prior to the Enlightenment for which we can provide a defence in the languages of twentieth century analytic philosophy.

Having now admitted that part of what this dissertation attempts to do is to revive a pre-enlightenment conception of essence, I think it is also time for me to admit that the primary influence on my views about how I think essence ought to be understood has been the work of Aristotle. The central thesis of this dissertation was inspired by the Aristotelian construal of essence and the result, not surprisingly, brings our understanding of the concept a lot closer to the way in which Aristotle understood it — particularly in the "Physics". After the "Physics" Aristotle went on to say a great deal more about essences which indicated that his earlier discussions had only really scratched the surface of what is a very complex phenomenon. I want to point out that everything that I have said in this dissertation is similarly preliminary and somewhat superficial. In particular, I have made much of the explanatory role played by the concept of essence, but this is only the beginning of a long and complex exposition. I think there is a still a great deal to be said concerning what it is about essences that makes them inevitably explanatory. In addition, it would seem that the phenomena that essences are explanatory of, are all characteristics of *substances* (in the technical/categorial sense of the term) - characteristics which objects have *qua* being a certain kind of existent. This suggests that the concept of essence is one which only really makes sense when it is woven into the fabric of a theory of substance. I think

that inter-relation of the concepts of essence and substance is an area which deserves more detailed investigation.

BIBLIOGRAPHY

1. Achinstein P 1977 "What is an Explanation" American Philosophical Quarterly **14**(1) 1-15

2. Achinstein P 1984 "The Pragmatic Character of Explanation" PSA ii 275-92 - reprinted in Ruben D-H (Ed) Explanation OUP 1993 Oxford

3. Aristotle "Metaphysics" in A New Aristotle Reader (J Ackrill ed.) OUP Oxford

4. Aristotle "Physics" in A New Aristotle Reader (J Ackrill ed.) OUP Oxford

5. Aune B 1985 Metaphysics: The Elements University of Minnesota Press Minneapolis

6. Ayers M 1975 "The Ideas of Power and Substance in Locke's Philosophy" Philosophical Quarterly **25** 1-27

7. Ayers M 1991 "Substance: Prolegmena to a Realist Theory of Identity" Journal of Philosophy **88** 69-60

8. Ayers M 1998 "Substance" in Routledge Encyclopaedia of Philosophy Routledge London p205-211

9. Brody BA 1972 "Towards an Aristotelian Theory of Scientific Explanation" Philosophy of Science **39** 20-31

10. Carl W 1994 "The Reference and Sense of Names" in Frege's Theory of Sense and Reference: Its Origins and Scope CUP Cambridge p161-185

11. Dennett DC 1984 Elbow Room: The Varieties of Free Will Worth Wanting OUP Oxford

12. Donnellan KS 1962 "Necessity and Criteria" Journal of Philosophy **59** 647-658

13. Evans IH(ed) 1981 Brewers Dictionary of Phrase and Fable Cassell Publishers London

14. Forbes 1985 The Metaphysics of Modality Clarendon Press Oxford

15. Frege 1892 "On Sense and Reference" reprinted in Geach P and Black M 1960 Translations from the Philosophical Writings of Gottlob Frege Basil Blackwell, Oxford p56-78

16. Hookway 1995 Quine Basil Blackwell Oxford

17. Hussey EL 1995 "Heraclitus of Ephesus" in The Oxford Companion to Philosophy (T.Honderich ed) OUP Oxford

18. Hussey EL 1995 "Parmenides" in The Oxford Companion to Philosophy (T.Honderich ed) OUP Oxford

19. Irwin T 1988 Aristotle's First Principles OUP Oxford

20. Johnston M 1990 "Is there a Problem about Persistence I" Journal of Philosophy **87** 107-135

21. Jones WE 1999 The View from Here: A first person constraint on Believing DPhil Dissertation Oxford University

22. Kaplan D 1989 "Demonstratives : An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and other Indexicals" in (J.Almog, J. Perry and H.Wettstein eds) Themes from Kaplan OUP New York

23. Kenny A 1995 "Sense and Reference" in Frege Penguin Books London p126-140

24. Kim J 1987 "Explanatory Realism, Causal Realism, and Explanatory Exclusion" Midwest Studies in Philosophy 12 225-39 reprinted in Ruben D-H (ed.) Explanation OUP 1993 Oxford

25. Kitcher P 1981 "Explanatory Unification" Philosophy of Science **48** 507-531

26. Kitcher P 1998 "Explanation" in Routledge Encyclopaedia of Philosophy Routledge London p520-525

27. Kitcher P and Salmon WC 1987 "Van Fraassen on Explanation" Journal of Philosophy **84** 315-30 - reprinted in Ruben D-H (ed.) Explanation OUP 1993 Oxford

28. Kripke S 1977 "Speakers Reference and Semantic Reference" reprinted in 1990 in The Philosophy of Language (Second Edition) (AP Martinich ed.) OUP Oxford 203-211

29. Kripke S 1977 "Identity and Necessity" in Naming Necessity and Natural Kinds (ed. SP Schwartz) Cornell University Press Ithaca and London

30. Kripke S 1980 Naming and Necessity Basil Blackwell Oxford

31. Lawson-Tancred H 1995 "Ancient Greek Philosophy II : Aristotle" in Philosophy : a guide through the subject (AC Grayling ed.) OUP Oxford

32. Lewis DK 1986 On the Plurality of Worlds Basil Blackwell Oxford

33. Lewis DK 1973 "Possible Worlds" reprinted in 1979 in The Possible and the Actual (M Loux ed.) Cornell University Press Ithaca New York

34. Lewis DK 1986 "Causal Explanation" in Philosophical Papers vol II OUP Oxford - reprinted in Ruben D-H (ed.) Explanation OUP 1993 Oxford

35. Lipton P 1990 "Contrastive Explanation" in D.Knowles (ed.) Explanation and its Limits CUP Cambridge - reprinted in Ruben D-H (ed.) Explanation OUP 1993 Oxford

36. Loux 1998 Metaphysics: A Contemporary Introduction Routledge London

37. Mathews RJ 1981 "Explaining Explanation" American Philosophical Quarterly **18**71-7 reprinted in Ruben D-H (ed.) Explanation OUP 1993 Oxford

38. Megone CB 1988 "Aristotle on Essentialism and Natural Kinds: Physics II,1" Revue de Philosophie Ancienne VI (2) 185-212

39. Putnam H 1962 "It Ain't Necessarily So" Journal of Philosophy **59** 658-671

40. Putnam H 1973 "Meaning and Reference" reprinted in 1990 in The Philosophy of Language (Second Edition) (AP Martinich ed.) OUP Oxford 308-315

41. Quine WVO 1953 “On What There Is” in From a Logical Point of View Harvard University Press Cambridge Massachusetts
42. Quine WVO 1953 “Reference and Modality” in From a Logical Point of View Harvard University Press Cambridge Massachusetts

43. Russell BAW 1905 “On Denoting” reprinted in 1990 in The Philosophy of Language (Second Edition) (AP Martinich ed.) OUP Oxford 203-211

44. Russell BAW 1919 “Descriptions” reprinted in 1990 in The Philosophy of Language (Second Edition) (AP Martinich ed.) OUP Oxford 212-220

45. Salmon N 1982 Reference and Essence Blackwells Oxford

46. Schwartz SP 1977 Naming, Necessity and Natural Kinds Cornell University Press Ithaca New York

47. Taylor R 1992 Metaphysics (fourth edition) Prentice Hall Inc. Englewood Cliffs, New Jersey

48. Van Fraassen BC 1977 “The Pragmatics of Explanation” American Philosophical Quarterly **14**143-150

49. Van Fraassen BC 1980 "The Pragmatics of Explanation” in The Scientific Image Clarendon Press Oxford - reprinted in Ruben D- H (ed.) Explanation OUP 1993 Oxford

50. Wiggins 1980 Sameness and Substance Basil Blackwell Oxford

51. Wiggins D 1995 "Substance” in Philosophy : a guide through the subject (AC Grayling ed.) OUP Oxford

52. Witt C 1989 Substance and Essence in Aristotle : An interpretation of Metaphysics VII to IX Cornell University Press Ithaca New York, and London