Jungian Archetypes: A Step Towards Scientific Enquiry

Kyah Leonie McDonald

University of Huddersfield

Contents

Jungian Archetypes: A Step Towards Scientific Enquiry

Abstract

The Jungian archetypes and transformations are often branded as pseudoscientific and difficult to ontologically define. This has led to problems in testing and a scarcity of papers attempting to scientifically explore this phenomenon. This paper explores why that has historically been the case and critically examines our current scientific standards and how they might be applied to the archetypes. The paper remains critical of a purely positivist paradigm, but concludes that some testability is necessary for the ethical application of the theory; its continued relevance in Psychology; and to ease current tensions in the field so that each side (those who champion the material and those who champion the experiential) may benefit from the other in refining this theory. I finish by presenting an account of the archetypes and transformations that focuses on the observable aspects of the phenomenon, and how they may be explained by recent findings in Neuroaesthetics and image schemata, while not denying the existence of the experiential aspects, e.g. the numinous quality of the archetypes, or their role in individuation. It is hoped that this will enable testing in some capacity and position the archetypes and transformations as an experiential and material phenomenon, being neither held on a precious, untestable pedestal, nor reduced to a set of fragmented parts, disconnected from their history.

Introduction

The following thesis will explore the concept of the archetypes and their transformations. With the increasing relevance of religion and climate crisis in the political landscape (Kashima 2016, cited in Saroglou, 2016), the theory of the archetypes may prove to be a valuable tool to understand our world and give our lives deeper meaning. However, this will only be the case if they can be shown to be sufficiently evidenced. At present, there have been very few scientific studies which attempt to test the phenomenon of the archetypes, and Jung's work is often written off as pseudoscientific due to the many barriers that have stood in the way of scientific enquiry.

This work will therefore aim to outline the archetypes as they were initially conceptualised by Jung, exploring some examples, from the collective through to the individual representations. With this outlined, I will discuss the issues with ontology which plagued Jung's work since its conception and continue to cause contention today. For example, researchers such as Laughlin and Tiberia (2012) have stated that "The archetypes are either inherent systems of neural circuitry, or they do not exist. It is as simple as that" (p.140). Meanwhile, developmental researchers such as Knox (2010) have argued that "Domain-specificity not only does not have to be pre-existent, it cannot be pre-existent" (p.523), i.e. that symbolic thought cannot be determined by our genetics. By offering potential explanations of why Jung appears ontologically vague (Palmer, 1997) in his initial conception of the archetypes, it is my hope that it will be possible to gain deeper insight into the nature of the archetypes (and the appropriate methodology for their study), which is one of the key issues which stands in the way of further scientific enquiry.

With ontology (and some of Jung's background) established, I will further explore why this work matters, and attempt to provide a breakdown of the original theory of the archetypes and collective unconscious to make it more accessible to the casual reader. This in itself is a valuable step, as at present, Jung's work can seem esoteric to the extent of dissuading engagement, and the frequent personification of concepts may in itself be unappealing to the modern scientific reader (Storr, 1973). Additionally, the field itself is rife with different uses of each relevant term, often making it difficult to discern what phenomenon is actually being explored in each new article. Most importantly though, this breakdown will, hopefully, create a clearer path for establishing clear hypotheses regarding the original hypotheses, which can be theoretically explored and tested.

Following this, I will give some background on falsification, why it matters (and why it does not) and use the breakdown of ideas to explore how Jung's theory stands up against modern ideas of falsifiability. After establishing the shortfalls of the theory, I will then explore modern conceptualisations of the archetypes and transformations (particularly that of Knox, 2004, and Merchant, 2009) to see if these can be resolved, and finally offer a re-conceptualisation of Jungian theory which is underpinned by the subjective but encompasses modern psychological research to explain the observable phenomenon. This will be used to suggest a clearer route to scientific enquiry, while aiming not to lose the heart of the original

theory. It is my hope that this will contribute at least one piece towards the puzzle that is the archetypes.

Jung and the Psyche

The Collective Unconscious and the Archetypes are concepts originating from the theory and life's work of Carl Gustav Jung. The collective unconscious is a broad concept which can simplistically be described as a theoretical substratum of human consciousness which gives rise to universal ideas and patterns of behaviour (known as archetypes). However, this explanation is not satisfactory when trying to understand, for example, the role of archetypes in the supposedly compensatory function of dreams, or how an individual's relationship with the archetypes may be used to understand their complexes or delusions. In order to fully grasp Jung's notion of the collective unconscious and the archetypes, it is first necessary to understand some of Jung's ideas about consciousness and a few of the concepts that Jung borrowed from the natural sciences, used analogously to outline psychic functioning[1].

Within his clinical work with patients, Jung observed an estimated 80,000 dreams (Von Franz, 1964), wherein similar symbols (spontaneous images with underlying meaning) appeared across different individuals. In some cases, these symbols seemed to have esoteric roots, as in mythology or obscure texts which he believed to be outside of the patient's realm of knowledge. One notable (and slightly outlandish, but resultantly memorable) instance of this is the solar phallus example, wherein Jung discusses a schizophrenic patient he encountered at Berghölzi Mental Hospital. Jung describes the patient tilting his head by the window and urging Jung to stand with him and do the same, that he might see "the sun's penis" (1936/37/1975, p.51, *CW* 8, para. 105), which he claims controls the wind. At the time, Jung dismissed this as the man being "crazy" (1935/1977, p.41, *CW* 18, para. 85). Four years after the event, Jung discovered a book on mythology by Albert Dietrich, which included an exploration of a ritual for the God 'Mithras' (Jung, 1936/37/1975). The origins of this God remain contested, but it is notable that Mithras has many associations with the sun, including his birth supposedly coinciding with that of the official sun God (Natalis Invicti),

[1] The word 'psychic' here is used to mean 'pertaining to the mental world', as it was used by Jung throughout his works.

being depicted as dining with Sol (Roman sun God), and being assigned at the seat of the equinoxes framed by Cautes and Cautopates for the solstices (all, of course, with relation to the position of the sun. In the Mithraic cult, this bore symbolic relation to life, death and the journey through immortality) (Beck, 2000). The book included a ritual which explicitly included breathing in the rays of the sun and bringing oneself into a mental state where they would become able to see the sun from an aerial position. From here, it would be possible to see a tube running from the sun which controlled the winds. Upon checking the publication date, Jung found that this information had been published in 1910 (Jung, 1936/37), four years after the institutionalisation of the patient (Merchant, 2009). This led him to conclude that it was not possible that the patient had come across this image before and merely replicated it. Jung determined it was highly unlikely that such a similar image would appear as a matter of coincidence, and that there must be an underlying basis for the similarity that he believed superseded typical conscious experience. It was a collection of experiences such as these which led Jung to generate his initial hypotheses regarding the collective unconscious and the archetypes.

This initial insinuation brought a great deal of criticism and accusations of Lamarckianism (believing that one can genetically pass on traits inherited during a lifetime) to Jung's work. It is important to note here that at least by the end of his career, Jung was not suggesting that the archetypes were specific images that could be inherited, as the above might imply. By contrast, Jung eventually appeared to lean towards a more general notion of the archetypes, i.e. in Jung's foreword to Harding (1948/1977) he explicitly states that an archetype is not "an inherited idea but rather an inherited mode of psychic functioning" (p.518, *CW* 18. para 1228). The archetypes should therefore be thought of as a mode of psychic functioning which gives rise to patterns of thoughts and behaviour. These patterns appear to share commonalities in both function and often appearance.

Jung posited that the psyche has a tripartite structure comprising of consciousness, the personal unconscious and the collective unconscious (McGuigan, 2009). Consciousness is the store of individual experiences that can be recalled at will, explored and re-experienced. The personal unconsciousness similarly arises from individual experiences but cannot be readily recalled, or re-experienced, while the collective unconscious is the root of the archetypes, which are a set of tendencies or instincts (which can be applied to thought or behaviour)

which do not arise from personal experience and are transcultural and resistant to temporal change. Akin to the personal unconscious, these cannot be consciously accessed but only be observed through their secondary manifestations i.e. the archetypal 'transformations' (Laughlin & Tiberia, 2012). 'Transformations' here will be used to describe the symbol itself, arising from the underlying archetypal structure. Transformations are the way that the archetype manifests in a particular image or concept which usually surfaces through dreams, fantasies or visions; even the most idiosyncratic personal fantasies can hold universal qualities. For example, Jung worked with a psychotic patient, Miss Miller (pseudonym), who experienced visions and chaotic dreams and created literature based off the experience (Jung, 1952a/1967). In one of her works, 'Songs of the Moth', the moth's search for the sun is made a metaphor for her own desire for God and as is highlighted by Jung, this comparison of God with the Sun is a common theme, extending beyond her own time and geographical location; it can also be found in Mithraic liturgy, Egyptian symbolism and the poetry of Goethe and Nietzche (Palmer, 1997). Here, for example, the transformation is the Sun (being the direct symbol) as a representation of the 'God' archetype.

In explaining the relationship between these structures, Jung makes use of the concept of enantiodromia (Palmer, 1997), a Heraceitean concept which suggests that 'opposites generate opposites' (largely akin to Newton's third law, 'for every action, there is an equal and opposite reaction') and that all opposites eventually meet (similar to the 'as above, so below' concept outlined in the Kybalion (Three Initiates, 1908)). On the basis of this assumption, Jung suggests it must therefore be that everything we hold in our conscious minds must have an equal and opposite in the unconscious. Although Jung went on to call Freud's Totem and Taboo (1913) a "fanciful assumption" (Jung, 1953a/1977, p.440, *CW* 18, para. 1074) later in his career, there is a similarity between the two core assumptions of their work which is relevant here and most concisely illustrated in Totem and Taboo. Freud (1913) suggests that if a taboo exists, then it must be the highest desire of the individual or the group, for if there is no desire for an object, then there is no need for prohibitions against it. The development of a 'taboo' or scorn towards any object or action is thus an internal preventative action towards a pre-existing want. The standpoint is that a 'block' against a desire is generally formed of either complexes in the self, or social convention. In Jungian terms, the desire for a prohibited object or action would be made unconscious, and cause the conscious mind to apply equal and opposite pressure to repress it, which causes a lack of overall psychological balance as a

high level of energy is being expended in a single area, e.g. a virgin who wishes to appear chaste but who has a powerful desire for sex, may develop an outward obsession with portraying sex as contemptible, or a paranoia that others believe their virginity is in question. The sexual obsession itself will make itself known, simply in a way that is considered acceptable for the conscious mind, and since the true function of this behaviour remains unfulfilled, this obsessive behaviour will continue. These blocks, for Jung, are the primary cause of neurosis. However, he differs from Freud in that he suggests that these desires are not primarily sexual (Jung, 1953a/1997), and that problems can arise across whole the spectrum of human experience. This led to Jung's belief that his task as a clinician was to create a *unio mentalis* (Jung, 1964), more commonly referred to in Jungian psychology as individuation (achieving mental balance, where neither conscious or unconscious content are causing undue stress on the other). Most typically, this was achieved by making unconscious content, conscious.

In situations where conscious boundaries were lowered, Jung believed it was possible to see the cause of neurosis in the unconscious manifestations, such as in fantasies, delusions, and dreams. Across his observations of these, he noticed recurrent themes, inclusive of figures, environments, and situations. Jung found that these themes ran parallel to many primitive ideas, myths and rites (Jung, 1964), which he believed to be indicative of a kind of cultural psyche that extends beyond the individual. To explain the recurring patterns, Jung hypothesised the existence of archetypes and the collective unconscious.

The Archetypes and Collective Unconscious

The way the archetypes emerge is multi-fold, but most noticeably, they are recognisable as repeated ideas or situations which appear across time and culture. The archetype is more akin to a template, a repeated pattern of functions and properties which can assume many different forms. To illustrate this, we will explore the Mother archetype, shifting from collective to individual representations and the different forms the archetype can take. The Mother archetype typically emerges in aspects of life related to fertility, nurturing, containment and other common maternal associations (Jung, 1952b/1967). This following exploration will allow us to look concretely at the core phenomenon that Jung was attempting to explain with his theory of the archetypes.

There are a variety of collective representations which may be a branch of the core Mother archetype, but the one we will focus on here is that of the Earth Mother. Following the Neolithic revolution (circa 10,000 BCE), wherein hunter-gatherer societies transitioned into agriculture and farming, there are many accounts of the idea of the Earth Mother which appear across different times and cultures. The early Earth Mother concepts centered around ideas of fertility and nourishment of the earth, but also ruthlessness, violence and death which in turn brings new life, reflecting the everyday man's experience of harvesting and hunting. The pervasiveness of this notion of Earth as Mother is observable in many early religious and cultural beliefs; Armstrong (2005) provides a list of figures which act as collective representations of this notion - in Syra, Asherah or Anat; in Sumer, Inanna; In Egypt, Isis; in Greece, Hera, Demeter and Aphrodite. The idea of the Earth as the mother and ultimate source of all life extended to such a degree that the fifth mythical king of Athens, Erechtonius, was even said to have been born from sacred soil rather than a physical womb (Armstrong, 2005). Jung (1952b/1967) outlines shared rituals and beliefs which stem from this idea, such as European ideas of a King being considered as marrying the land itself upon coronation, and their luck being reflected in the yield of the harvests. Jung (1952b/1967 highlights the story of Domalde, a mythical Swedish king who was executed on his third year for crop failure, and also outlines the old Chinese custom of emperors ploughing a furrow upon their ascension; their rule is also considered to be a union with the land, as a feminine body. Interestingly, Armstrong (2005) notes that the earth was not always regarded as female in China and Japan and only became so when the dawn of farming and the maternal role of women aligned in their history, which appears to indicate some connection between physical circumstance and the development of the archetypes, or at least their expression. We can therefore see that regardless of location, the Earth as Mother concept was an idea that repeatedly emerged and gained cultural popularity.

As a nod towards being temporally enduring, as Jung determined the archetypes must be, we need only glance towards the familiarity of the term 'Mother Nature', a highly common phrase in Western society, to see that the association with the Earth and the feminine mother has not largely changed – and interestingly, nor has the general principle of agriculture. Seeds are still sewn into the earth, crops are still harvested, words like 'fertile' are still used to describe good land, and we depend still upon the whims of the nature to determine whether or

not the food will grow. This is to a lesser or greater degree depending what part of the world one lives in, but the fact itself has not largely changed since the Neolithic era. The continued idea of humans as ‘born from the earth’, akin to Erechtonius, may also be suggested by the symbolic significance of the scythe carried by the typical figure of death. Lakoff (2015) discusses this figure in his lecture on the embodiment hypothesis, and suggests that the psychological significance of the scythe, as a common agricultural tool, is that we continue to view ourselves as plants in the life cycle (which he notes is also evident in phrases such as 'young saplings' or 'sewing seeds'), with death as the figure that would cut us down in our prime, like a farmer at harvest. This demonstrates that despite thousands of years passing in between, the concept of Earth as Mother persists today, indicating what Jung considered a temporal resistance to change.

Although the Earth Mother is a widespread manifestation of what Jung termed the Mother archetype, it is by no means the only one. The Mother archetype is thoroughly illuminated and elaborated on by Jung himself (1954a/1975), and in this explanation, Jung states that the archetype can manifest itself in anything from the literal mother (or stepmother) figure, to more abstract symbols like the mythological Demeter, the religious Virgin Mary, or physical symbols such as water, churches or ovens. At first glance, the links between these objects seem tenuous but Jung goes on to state that traces of the Mother archetype can be found in anything that possesses the following traits: it inspires devotion or awe, has connotations of fertility and fruitfulness, is protective or hollow (in reference to the womb), or is found to be helpful or nurturing. Due to the essential element of duality in Jung’s theories (internal/external, positive/negative), the mother archetype also has a negative set of manifestations which Jung (1954a/1975) states may manifest in figures akin to the Goddesses of fate, such as the Moirai in Greek, or Norns in Norse mythology. These three are female entities which preside over life and death, sometimes as witches or other unearthly beings such as giantesses which hold immense power over mortal life. The negative mother archetype can also be found in any ‘devouring’ or entwining animals such as dragons, or serpents and can also be symbolised by deep waters or the grave. The last two of these returns us again to the idea of containment or a hollow, only this womb brings death rather than life. Jung himself recognises that this list is not exhaustive and is explicit in saying that these transformations are not always related to the literal mother herself, but the idea of a mother figure.

This concept will be further explored below to illustrate how archetypes and their transformations may manifest in individual representations as opposed to the collective representations like the folklore or mythology mentioned above. However, this can often blend as the individual may subconsciously borrow from the collective, or alternatively the representation may be altogether unique to their life and experiences. The form that an archetype takes can depend on a wide variety of factors. Although collective representations (such as folklore) may share some uniformity, the same is not true always within the individual (Jung, 1954a/1975), as the archetype is subject to sociocultural and personal elaboration prior to expression. Jung (1954a/1975) also makes a point of the stating that since the archetypes are not the images in themselves, the way they manifest is a matter of "infinite variety" (Jung, 1954a/1975, p.81, *CW* 9 pt.1, para 156).

Merchant (2009) outlines one of the individual experiences with the negative mother archetype, drawn from Jung's (1931/1991, cited in Merchant 2009) work on the structure of the psyche. An army officer, who was a patient of Jung's, was displaying a set of psychogenic symptoms - a heart pain, a choking sensation, and a sharp pain in his left heel. Prior to the onset of these symptoms, the man had been in a relationship, but the woman had then left him for another man. Jung eventually determined that the heart pain and choking sensation were the result of unconscious affect - the man's broken heart, and the sensation of a 'lump in his throat' caused by repressed tears, both of which disappeared upon expression. This did not account for the sharp pain in his heel, and so the therapy continued. Subsequently, the man dreamed that he was bitten in the heel by a snake, and that he was instantly paralysed by this attack. During the discussion of this dream, the man revealed that his mother had pampered him which in turn led to a femininity of character that did not serve him well in school, nor in the military where he had sought to subdue this side of himself. As Merchant (2009) points out, the snake is a common biblical motif, representing corruption and downfall, and a loss of potential from what could have been. Merchant (2009) also indicates that it is highly likely that the man would have heard the Genesis story in early life, and so the unconscious, finding this an appropriate symbol for his feelings of betrayal from his mother and girlfriend, draws on this to describe how he has been "lamed" (Jung 1931/1999, cited in Merchant 2009). In this instance, we arguably see the unconscious mind drawing on a common collective

representation to express itself. However, the symbols that the mind utilises can be altogether more personal in nature.

One such case of this can be found in the Miller Fantasies, a collection of writings wherein Jung (1952b/1967) analysed the fantasies and what he described as 'visions' of Miss Frank Miller. In one of these visions, Miss Miller was centered in a dream-like city. This city held clear importance to Miss Miller and was described as being ardently longed for, and which is likened to heavenly Jerusalem by Jung. Here, the figure that stands in for the mother archetype is not a figure at all, but a setting - a place which contains 'crowds' of citizens and provides them with care and sustenance. The city thereby serves the function of the mother archetype as a container and provider of life, although the transformation itself is quite distinct from those laid out above. Jung (1952b/1967) also draws a parallel between this city and the Jerusalem and Babylon of the Old Testament, as these are also spoken about like women and referred to with phrases such as 'harlots' or 'the lady of Kingdoms'. This trend indicates the continuity of thought between the individual and the collective, which characterises the collective unconscious.

Jung's perception of reality as the interplay of psychic factors underpins the significance of the archetypes in day to day life; since the world we live in is a matter of conscious perception, interactions with others are really interactions with ourselves, and our relationship with different parts of ourselves can shape our life experiences and our attitudes towards others. To utilise the archetype explored above, a person who has repeatedly internalised the negative mother as their conscious attitude may deeply resent or be dismissive towards anything which shares abstract properties with the mother archetype. Equally, it seems possible by extension that a nation-state which has developed a poor relationship with the Mother archetype may take reckless attitudes towards agricultural policies and the continued vitality of the earth itself. However, in accordance with the duality assumptions outlined earlier, they may also have a private desire to be aligned with the abstract properties of the mother archetype, e.g. the desire to feel protected and feel devotion (which would emerge in dreams or fantasies) which would need to be fulfilled in order to establish mental balance. The issue this potential for dual outcome produces is that of difficulties in falsifying the theory, i.e. providing potential conditions under which the theory could be shown to be false as opposite outcomes could still both be used to support the theory with ad hoc reasoning to

explain the outcome; the only thing that would not be expected is apathy. At present, there has been little work to determine which factors would determine which of the outcomes would emerge, and indeed, there may be a multitude of hidden variables involved in shaping this.

Science, Ontology and the Archetypes

What has been presented thus far largely focuses on the experiential aspect of the archetypes as opposed to their potential material basis. This is primarily due to the fact that Jung's writings were generated from idiographic practises with clients rather than through experimental methodology. When exploring the original theory, there is therefore much more information to present on the subjective aspect. However, Jung was keen to highlight that he did believe that there was a biological basis to the archetypes. In 1948(/1977), he wrote:

> "Of course, this term [the archetype] is not meant to denote an inherited idea but rather an inherited mode of psychic functioning, corresponding to the inborn way in which the chick emerges from the egg, the bird builds its nest, a certain kind of wasp stings the motor ganglion of the caterpillar, and eels find their way to the Bermudas. In other words, it is the "pattern of behaviour". This aspect of the archetype, the purely biological one, is the proper concern of scientific psychology." (p.518, *CW* 18. para 1228)

Jung is often charged with the accusation that he is ontologically vague (Palmer, 1997), which makes a biological basis arguably hard or impossible to determine. The discourse surrounding this likely emerges because Jung stresses that the archetypes are not reducible to biological processes and that they cannot be channelled into concrete reality (Jung, 1953b/1977). As Mill's (2012) highlights, Jung's implicit supernaturalism has drawn much criticism, particularly considering the apparent God-like implications of a structure which is supra-individual, transcends time and culture and is all too frequently personified, implying that the archetypes have a kind of autonomy independent of the individual. In reality, these descriptions are more likely to be a reflection of Jung's attempts to capture both the

subjective experience and the objective reality of the archetypes in a single sentiment.

In modern psychology, the tendency to view subjective and material as a dichotomy makes this position hard to reconcile in the current paradigm; an object is considered concrete or abstract rather than both. For example, it is more common to see *love* primarily as an abstract concept and to focus on its affective significance rather than its biological basis, or vice versa, it is more common to view a chair primarily as material object rather than having affective significance. This mode of thought can make it difficult to place the archetypes; abstract, or material? If Jung viewed the archetypes as purely metaphysical as implied above, this would make them problematic for the modern psychological community as it arguably precludes them from scientific study (explored further below). However, it seems that in reality, Jung was equally interested in both the empirical[2] basis and experiential aspect of the archetypes, once stating that "we can therefore speak of an unconscious only in so far as we are able to demonstrate its contents" (Jung, 1954b, p.4, *CW* 9, p.1, para 4).

A thorough study of the archetypes must therefore aim to be scientific and replicable and yet it must also capture the subjective experience. Although this has caused issues for the scientific acceptance of the archetypes (as an archetypal experience is extremely difficult to generate or externally validate, and only being able to study something through its secondary effects is far from ideal), this is not necessarily a bad thing; Knox (2004) captures the sentiment well when discussing the position of neuroscience in psychoanalysis and the importance of the subjective element, stating that "the brain does not have to be studied at the expense of the mind" (p.2). This leads us neatly into the following discussion - could the archetypes be studied scientifically? And what is it, precisely, that we mean by 'science'?

Jung and Science

In order to fully understand how the archetypes and collective unconscious were conceived within the sciences (and what their position is there) one must first understand the societal

[2] 'Empiricism' is defined here as the view that observations from the senses (especially those which are quantifiable) are the *main* way to 'know' or gain knowledge about reality (Hjørland, 2005), with the exclusion of introspection, intuition or metaphysics. 'Empirical' equally is used to mean study through observations which can be gained through the senses, with the same exclusions applied.

context that Jung was emerging from, and what bearing this had on his ideas of science. It is no secret that Jung's work has been the subject of controversy. Throughout the course of his career, he produced titles such as 'Psychology and Alchemy', routinely delved into the occult and as previously mentioned, held to the Hermetic Law of Correspondence; (Jung, 1929/1975, p.9, *CW* 15, para. 12) "… everything without is within, everything above is below. Between all things … reigns "correspondence" (*correspondentia*)…". It is no surprise, therefore, that Jung was and is frequently charged with accusations of mysticism. This reputation has only acted to cast further doubt on his theory of the collective unconscious as a scientific concept. As Jones (2014) highlights, there are also suggestions that Jung's work was not scientific and cannot be scientific because of its basis in analytical psychology. To assess this claim, it is initially important to discuss, in a broad sense, what science is and what it will be used to mean throughout this paper, beginning with its etymological root.

The word 'science' stems from the Latin 'scientia' which simply means knowledge, or what is known. It is therefore impossible to go further without recognising that this is a discussion of epistemology, since this is inherently a conversation about what it is to know, and by which means we are capable of knowing. The writer has no intention of attempting to resolve this question, but it is sufficient to know that there are at least two derivatives of 'science' or 'ways to know' which will be relevant for the following discussion. Having an awareness of both strands will not only help to understand the apparently conflicting ontology of the archetypes but will also shed light on the narrower scope of 'science' as we know it now.

The reasoning for the existence of these two derivatives of science, as outlined by Cambray (2014), is not wholly academic, but also historical and socio-political. By the end of the Enlightenment era, both the science of England and France had largely been brought under the wing of their respective national agendas (Cambray, 2014). Anglo-French science dominated the field, prizing rationality and objectivity at the expense of subjectivity and the arts. In response, Germany had developed a science that was more independent of state-interest and which retained its holistic viewpoint. It also made less differentiation between the arts and science, and is best known as Wissenschaft. Cambray (2014) suggests that this was, in part, a result of central Europe being embroiled in the 30 year war in 1618-1648. The human capital involved in this (20% of the German population, and eight million fatalities in total (Clodfelter, 2017)), combined with the time period of the war, meant that Germany was

largely cut off from other centres of scientific thought during the dawn of the Enlightenment period, and thus did not share the influences of the science of England and France.

Anglo-French Science

The science of England and France that developed during the Enlightenment era was characterised by knowledge discovered and collected via empirical methods. Objectivity was prized, along with the ability to investigate through observation and replicate results through the scientific method. Mathematical grounding was and is also important for the sciences which belong to this derivative, particularly as this form of systematisation theoretically increases objectivity; relations between elements can be described by equations without need for subjective interpretation. The empirical method brought great gains for the quality of life of mankind, and had a direct impact on national interest; Lawrence (1999) suggests that the Enlightenment period brought with it a view of human perfection in the future, and placed itself as the hallmark of progression. However, implication of this was that cultures not part of this trend were considered 'behind' and in need of education, thus creating the necessary ethical grounds for imperialism (Lawrence, 1999). This new socio-political shift thus resituated science, both as the basis of ethical justification for proponents to consider themselves teleologically 'forwards' (Shank 2008), and as a key means production of weaponry and medicine that could be used for military purposes. The furtherance of science had thus become politically and socially relevant, leading to an increase in state-funded expeditions. These provided a profitable and largely successful form of knowledge collection for scientists of the time (Cambray, 2014). As a result of this, the trend of knowledge collection shifted towards utilising the scientific method for ends which would have economic, political or military use, as these were more likely to garner funding and high societal status. This therefore became socially lauded as the hallmark of good scientific practise and was purportedly the foundation of scientists such as Banks (prolific botanist), Newton (mathematical physics) and Boyle (one of the founders of chemistry) (examples from Cambray, 2014).

However, Cambray (2014) highlights that although Boyle and Newton are famed today for their empirical findings, there is often little mention that they also studied alchemy, which Newton wrote about more than physics, and which Boyle held a life-long interest in (alongside transmutation). These interests were beyond the quantifiable, and yet increasingly

became edited out of their history. As Shank (2008) suggests, the narrative that this creates is that these key findings were a natural result of scientific progression and the Enlightenment era, rather than arising from a complex social and philosophical backdrop, over-emphasising the role of the former and minimising the role of the latter. It may be argued that these interests were no longer mentioned due to their decreased relevance to society, but the result is that the breadth of their views and interests have been largely minimised in the retelling (Cambray, 2014), and the rich philosophical foundations that underpinned some of the greatest scientific findings in European history are thus diminished if not publicly erased. Thus, by the end of the Enlightenment era, the holistic backdrop for progressive thought had been replaced with an empirical gold standard, based purely in rationality, reason and strict scientific method. The subjective was framed as lesser and deemed largely irrelevant to the field of science. In this sense, the situation has not much changed today, and we continue to emulate a great scientific past divorced from philosophy and the subjective - a past which simply did not exist.

To illustrate this; within modern Psychology the gathering of knowledge outside of state purposes is often still considered frivolous, and it is commonplace that a rationale will include notes on potential practical applications to justify why a study is 'worthwhile', as if the pursuit of knowledge in itself is insufficient. The pressure to create socially significant and economically beneficial research still weighs heavy on UK universities, as funding bids remain crucial to a researcher's position, prestige and livelihood. A glance towards these should illustrate well for the modern audience the very real pressures and biases that have been faced by science since the 1700s. These socio-political factors have in turn shaped the values and attitudes within the field of science and should be kept in mind when we consider the attributed importance or reputation of any given theory in the field.

Wissenschaft

The second form of science, which we will refer to simply as Wissenschaft[3]. Saban (2014) quotes one of Jung's letters, in which he states that '[On the continent] any kind of adequate

[3] There are multiple forms of Wissenschaft, such as schöne Wissenschaften ("fine sciences") and höhere Wissenschaften ("high sciences") but for the sake of the argument here, 'Wissenschaft' itself will be sufficient.

logical and systematic approach is called "scientific"' or ‘Wissenschaft’. As a result of this, History, Mythology and Zoology are all considered types of sciences, falling under the broader and earlier definition of science as 'knowledge collection' (Saban, 2014). This is far more inclusive of a wide range of methodology and can therefore also be applied to those fields which also contain elements which are harder to mathematically describe such as History or Anthropology. Since this approach retained its independence from state, it is also more open to information gathering that is not centred around social purposes and application. For example, poetry, before Kant, was once considered a form of Wissenschaft, since it is comprised of the systematic rules that underpinned the construction of a poem, and the capacity produce it (Bommel, 2015). Beauty was likewise studied through repeated examples and rules (Bommel, 2015). Within Wissenschaft, hermeneutics was therefore not cut off from systematised knowledge collection in the way that it came to be within Anglo-French science. In fact, it was central to the concept of ‘the true, the good and the beautiful’ propagated by the influential German philosopher, Herder (Bommel, 2015), and the Sturm and Drang movement he backed between the 1760s and 1780s. The Sturm and Drang movement prized the expression of emotion and ran counter to the Enlightenment movement, opposing the reductive elements of rationalism, which was perceived as largely cutting the subjective element from human study.

Both Cambray (2014) and Saban (2014) highlight Goethe as a prime example of Wissenschaft in their essays on Jung and science. Goethe (1749-1832) suggests that the nature of a phenomenon is both material and immaterial – art and science. For Goethe, science was inherently hermeneutic. In fact, this view was originally more common as facts were not considered something that was ‘out there’ but something that was made between the observer and the phenomenon they were observing (Saban, 2014), hence the root of the word ‘fact’ being the Latin ‘facere’ which means ‘to do’ or ‘to make’. Since things are experienced both objectively and subjectively, Goethe believed that it was necessary to describe both in order to capture the true essence of a thing. This trend was not exclusive to Germanic science and can also be seen in Banks’ notes from the James Cook expeditions, which also included rich experiential descriptions of phenomena which largely influenced the emergence of Romanticism, and writers such as Shelley and Keats (Cambray, 2014).

Social Implications and the Jungian Community

It is important to acknowledge that neither of these derivatives should be considered dominant or subordinate as neither can, without question, be said to be better or worse than the other. Each has a significant bearing on society and an impact on the knowledge that emerges. However, being aware of these two derivatives is significant to understand the scientific paradigm we currently exist in; if science is simply the collection of knowledge through logical and systematic collection, this is open to all peoples. If science requires accuracy, replicability and rigorous testing, the pool of people able to conduct it is far more limited. For example, in the psychological sciences, it is often the case that laboratory studies are considered the 'gold standard' of empirical testing. These require a venue to control for confounding variables, expensive equipment for accuracy, and access to an incentivised population, as well access to journal articles from reliable sources which must often be paid for to access. The easiest way to gain access to all these things is through university (which by proxy involves being educated under the current paradigm) or through funding bids. Even those employed at universities continue to be bound to national agenda, as there is a pressure for researchers to produce socially significant or financially significant research in order to maintain their own position and value in the university.

While this is not necessarily negative in essence, this system narrows the scope of knowledge collection undertaken, with research focused primarily around practical outcomes. This situation has only deepened since the educational policy in the UK has shifted to establish Universities as business institutions rather than free-access educational establishments. Another alternate route for acquiring funding for knowledge collection is research grants, which also approach science from the same angle. Proposals to the ESRC (Economic and Social Research Council) in the UK, for example, must also include a section where any costings are justified and a quick glance at the opportunities for funding (diabetes, healthy aging, societal transformation and climate change) still show a distinctly practical slant. Therefore, this variation of science quickly ends up in a situation where it is almost inherently bound to state-interest, and it becomes the case where this is the only form of knowledge collection that is considered viable or worthwhile for society.

This is not to advocate the abandonment of current scientific practises in favour of Wissenschaft. However, it is a warning against a society and scientific community building its ideas of prestige and truth purely on the back of empiricism and what is practically beneficial. A full adoption of 'Wissenschaft' with the same status as current 'science' would not be without its issues either, as the current dominant paradigm limits those who collect and produce statements of knowledge to those with at least a university level education and reduces ‘noise’. The scientific community is therefore safeguarded from simply anything being classed as 'knowledge' and anyone being classed as a 'scientist'. The means to produce and collect knowledge remain restricted but the cost, particularly in terms of the breadth of knowledge explored, remains a significant one.

The relevance of this in terms of the current paper is as follows; the reconsideration of Jung, this paper, and the current scientific community, do not exist in a void. If Jung’s ideas and framing are considered wholly unscientific instead of simply emerging from a different branch of science, then there would be no use in pursuing these ideas in modern psychology nor continuing to explore them within this work. These dual definitions have led to a split in the Jungian community, with a considerable amount of practitioners who maintain that science as it is, is irrelevant for Jungian practise (Jones, 2014). This is an understandable view from those who subscribe to a more 'Wissenschaft' angled perspective, versus 'science' as the only way 'to know'. There are many theoretical aspects of the collective unconscious which could not yet be adequately captured by empiricism and quantification and given much of the foundation being based in the experiential, there is reasonable concern for the theory not being done justice by modern empirical psychology. Equally, there exists a counter community who appear to search for anything Jungian in the empirical sciences and laud it as evidence that Jung is indeed 'scientific' in order to gain validation for the theory (Saban, 2014). However, like Saban (2014) it is the belief of the writer that this divide must be healed and the Jungian community itself must individuate and come to accept its contraries. The empiricists must accept the abstract, and the abstract the elements of the empirical, as opposites and equals. Much like in Goethe's concept of science, both must be used in order to capture the true nature of the thing, as neither alone will do the theory (or more importantly, the phenomenon) full justice.

Crucially, the reality is that the terms and standards that will be used (and already have been

used) to critique Jung are largely borne from the science of the Enlightenment era, which is largely positivistic and physical (i.e. emergent from a science which relies on sensory input as the only way 'to know', at the exclusion of the subjective element). This paradigm continues to be useful, but primarily to study the biological and empirical side of the collective unconscious, which Jung (1948/1977) believed to be the proper concern of "scientific psychology" (p.518, *CW* 18. para 1228), by which I believe he meant empirical psychology. Additionally, since this is the dominant paradigm, whether the collective unconscious can be reframed as a scientific concept (i.e. amenable to empirical study) without losing the experiential aspect of its roots remains critically important to its continued relevance in modern Psychology. Whether this is a reflection of it being 'truthful' or 'meaningful' is largely irrelevant; we cannot extricate the theory of the collective unconscious from its socio-political background and we cannot blindly and stubbornly ignore what bearing this will have on the theory's future; the impact of history is very much real and we see it traced clearly through the last 400 years of scientific progress. If the theory cannot be reframed and understood on modern terms, with empirical backing, then it will fall by the wayside and remain a quirky side-note in the history of Psychology rather than being recognised as a first attempt at exploring a complex and highly relevant phenomenon.

Addressing Unresolved Methodological Issues

That a phenomenon can be both subjective and objective is a position which makes clearer Jung's perspective on the ontology of the archetypes. The dual definitions outlined above ease the apparent contentions in Jung's work and highlight two potential methods of study for the collective unconscious – an approach which utilises the scientific method, and an approach which is systematic and logical, but not necessarily centered around the observable and empiricism. The latter is becoming increasingly common in the qualitative methods in Psychology, and also includes space for reflection on the interaction between the self and the interpretations provided. This fits very neatly with the underlying ethos of the psychoanalytic approach, as it again returns to ideas of facere, wherein researchers acknowledge and reflect on their positions as the interpreters and 'makers' of facts, and the recognises that the conscious and unconscious interplay between observer and phenomenon is a dynamic within itself, thereby bringing the humanistic element back into scientific enquiry. Thus, in combination (as in a mixed-methods approach), it seems to provide a potential pathway to establishing a more active and harmonious field for the study of the

archetypes.

However, even knowing about the two scientific derivatives (Wissenchaft and modern science) above, there remain some methodological issues in terms of whether the archetypes should be treated as absolute or relative phenomena. Jung's (1948/1977) position on this appears to be mixed, for example, he suggested that only the "biological [aspect], is the proper concern of scientific psychology" (p.518, *CW* 18. para 1228) but also clearly draws on German Romantic ideals of science himself, and stated that "purely intellectualistic, analytical, atomistic and mechanistic thinking has, in my opinion, landed us in a cul-de-sac, since analysis also requires synthesis and intuition" (1928/1993, pg. 39) The former statement suggests a call for a focus on a more enduring truth which is unchanging over time (i.e. an absolute), while the latter indicates the importance of interpretation on the part of the researcher, which may indicate an orientation towards a more flexible and relative truth.

This is not necessarily a flaw of character or indicative of the theory being irreconcilable with modern science, but at least partially reflective of what was, at the time, psychology as a science with a largely divided identity - an issue which it still suffers from today. Different branches of psychology continue to hold varying (and frequently opposing or contradictory) epistemological and ontological beliefs, as well as philosophical foundations, e.g. the discursive psychologists see reality as something that is made through speech, while strict empiricists might see reality as something that is 'out there', fixed and unchangeable by subjective means. As a note on this thought: the words "reality" and "truth" are often used interchangeably, yet we accept that reality is something that is changeable based on historical or emotional events. Meanwhile, truth is considered fixed and unmoving. This creates a peculiar divide wherein reality is not truth, and truth is not reality, although we mean both to be something which is 'actual'; that is to say, it may be more apt to suggest that there are relative and absolute phenomena, which each need their own distinct approach. Psychology (and the theory of the archetypes in particular) sits at the crux of this tension between relativist and positivist truth and continues to struggle in finding its place among the sciences methodologically as a result, being a curious blend of the methods of the natural and social sciences. Jung lived while this tension was still very much emergent, and it is this conflict which shapes the apparent contradictions surrounding the ontology of the archetypes and collective unconscious.

Accordingly, in the writer's opinion, the archetypes are neither biological nor experiential, nor can they strictly be classed a relativist or absolute phenomenon. They are both; with the archetypal patterns appearing invariably as an absolute phenomenon (unchanged by time or circumstance), and the transformations they manifest as appear relativist (influenced by the historical epoch, perception, culture, aesthetic values etc). Crucially, they are thus much like the flowers that botanist Humboldt (1793, cited in Cambray 2014) observed during his expeditions; "[they] should not be studied in isolation, but as an integral part of the environment in which they are found" (p.14). This requires us to study the context they arise from (supra-individual; experiential) and neurobiology (individual; material) and how these come together to produce the emergent subjective experience, or 'qualia' (individual; experiential). A full exploration of the archetype phenomenon must account for all three in order to capture the 'true nature' of the archetypes and their transformations and requires relativistic *and* absolute approaches to truth.

The Modern Relevance of the Archetypes

If it can be supported with empirical evidence to complement what already exists in experiential and cultural explorations, the very act of appropriately revising Jung's theory of the archetypes would offer a valuable model for how the empirical and experiential approaches within psychology might be reconciled to approach a single theory.

The methodology and philosophy underlying neurophenomenology may be well-suited to this task, as it opens an avenue to study both the qualia and the biology in tandem, and to study the empirical aspects of experience while not reducing them to biology alone. This method was developed by Varela in the mid-1990s (Gordon, 2015) as an attempt to tackle the dualist paradigm that had emerged as a result of the mind-body split (Guendelman, 2017), and as an attempt to tackle the hard problem of consciousness, i.e. how can neuronal activity produce subjective experience? (Chalmers, 1995; 2007). Varela (1996) suggests that the only way to overcome this is by studying the structure of human experience itself. The approach places great importance on the irreducibility of conscious experience, and seeks to find structure in phenomenological accounts, and to link these to their cognitive counterparts (Varela, 1996). This methodology has produced fruitful results, as in the case of the research

by Petitmengen, Navaro and Levanquyen (2007) which combined studies of the conscious experience of epileptic patients with neuroscientific methods. The experiential studies conducted acted to guide the EEG procedures and led to the finding that cerebral activity had undergone subtle change hours before the seizure. As well as the advancements in understanding epilepsy, the results of this study have opened the doors for possible therapeutic interventions for sufferers prior to seizure. This is an extremely a promising result for a relatively new methodology, and indicative of the potential of neurophenomenology.

However, as it is, neurophenomenology does not allow for the harmonious convergence of context alongside qualia and biology, which may well be better approached through anthropological methods such as ethnology. For example, if one were to explore the symbol for death in a specific culture, cultural studies would may need to be conducted to explore how this may influence the archetypal expression, as well as work with the individual's experience and underlying neurology. Additionally, neurophenomenology has yet to be popularised in modern psychology; there is still more work to be done to make this methodology clear and accessible to the typical researcher, making it hard to ensure scientific rigor and reliability at present. In the case of the archetypes, since any neural research would still be exploratory, it is also likely that (issues with identifying archetypal experience aside) functional magnetic resonance imaging (fMRI) research would be the most promising place to begin this kind of research in order to take advantage of their spatial accuracy. By contrast, the electroencephalogram (EEG) provides exceptional temporal accuracy (Cohen, 2011) but is not the best approach for precise location or investigating deep brain areas (Cohen, 2014), and without much prior research to identify the areas involved in archetypal experience, we cannot guarantee that we would not expect to see deep brain areas involved. It would also be incredibly difficult to find event-related-potentials (ERPs), and even for time series data, there is no current research to give us an idea of what the time course of an archetypal experience would be. Without a clear idea of this, the temporal precision actually acts as a disadvantage, as an uncertain time course renders the analyses unreliable (Cohen, 2014). This essentially makes EEG largely suboptimal for this investigation. The issue with the alternate, fMRI, is that it is currently extremely expensive and thus not accessible to all researchers, and so neurophenomenology's use for more exploratory research (as would be the case with the archetypes) is somewhat limited by these practical factors.

However, practically speaking, if an appropriate empirical approach to Jung's work could be found, this would also provide a useful tool for self-development, clinical work, and even for understanding the underlying tensions in religious conflicts. The last of these may seem a rather sudden departure from the former two which are arguably self-evident, but the theorised presence of archetypes in religion, and their relevance to socio-political conflict is not to be underestimated. This is particularly relevant today, as Huntington's (1993) Clash of Civilisations hypothesis (which suggested that post-cold-war conflict would centre around religious and cultural identity), appears to have come to fruition over the last twenty years with religion emerging again, above ideology, as a major source of intergroup conflict (Kashima 2016, cited in Saroglou, 2016). As Jung suggests that the primary form of archetypal expression is religious belief systems (1964), it is plausible that a broader adoption of Jungian ideas would provide the psychological community with new insights into religious conflict and potentially new methodologies to counteract them. This is a global issue, and certainly the business of social psychologists. It should also be noted, for example, that religiosity has been suggested as a lesser determinant of religious prejudice than traits such as authoritarianism (Saruglou, 2016), and while this intersection may be beyond the ken of contemporary theology or political studies, it is not beyond Psychology. A theory like that of the collective unconscious and the archetypes, which can grapple with intergroup conflict, religion and individual personality tensions, would thus have much to offer the modern world - but only if it can be shown to have foundations in material reality.

The current state of affairs is that the theory (despite lack of empirical evidence), is still used to understand individual cases, and there exist associations which offer therapy and training in Jungian practises. For example, the British Jungian Analytic Association offers four year courses in Jungian psychotherapy, an MSc and various training sessions. A glance at recent publications shows that Jungian theory is used to understand everything from organisational practises (Moxnes & Moxnes, 2016) to Indian music and alchemy (Singh, 2013). This may raise the question of why it is necessary to begin testing the very foundations of the theory if it is already in popular use in certain circles.

However, as already discussed, Jungian theory is still largely divorced from modern psychology and very few psychologists take the theory seriously – and with good reason. The theory lacks clear scientific foundations and efforts to unite it and modify it in accordance

with modern research have been few. The theory is often written off as needlessly esoteric (thereby making itself deliberately obtuse and hard to debate) or pseudoscience. This is not a promising progression for a supposedly scientific theory, which ought to be clear, accessible and falsifiable.

There is an additional importance for continued study in this field. If the theory is found to lack solid empirical evidence, by the current study and by future studies, then it raises the question of ethical implications for the practises outlined above. After all, scientific qualifications should not be offered in a subject which is not scientific and it is, arguably, unethical to provide psychological advice based on a theory which cannot be scientifically validated.

Finally, on this note, while the primary purpose of this thesis is to discuss whether the collective unconscious can be reformulated to be amenable to scientific practises, I would like to highlight that this is not a value judgement; as Popper (1963/1998) himself states when discussing his own demarcation criteria for scientific practise, this is not an assertion regarding meaningfulness, significance, truth or acceptability. It is simply a determination of whether this theory and the research stemming from it is scientific. The following chapter will therefore further discuss mainstream scientific practises, what we have come to regard as scientific honesty, and how the collective unconscious can (and cannot) fit within the current scientific paradigm.

Breaking Down the Collective Unconscious.

Before beginning to assess whether the theory of the archetypes is amenable to empirical investigation, it is first necessary to be able to understand the theory plainly in order to make it accessible; one of the first obstacles in understanding Jung's theory of the collective unconscious is the sheer number of sub-ideas that seem to exist within the core concept. Thus, in the interest of clarity, the core idea along with the most relevant intricacies are listed below.

1st [Core]: All individuals are capable of engaging in recurrent patterns of thought and behaviour, identifiable in other times or cultures.

2nd : That the recurrent patterns are due to a substratum of the unconscious, known as the collective unconscious. Unlike the personal unconscious, the collective unconscious is an inherited structure bearing much resemblance to the biological instincts.

3rd : That these recurrent patterns of thought and behaviour (often manifesting as spontaneous images, known as symbols) can be broken down into distinct archetypes, recognisable by the function they serve rather than by the image they appear as.

4th : That the type of thought forms and symbols that arise from this structure have some relationship to the psychological state of the individual and are functionally significant.

5th : That the archetypal symbols which arise are indicative of the hidden or repressed personality traits of the individual.

6th : That there is a part of the mind which seeks to strike a balance between the conscious and unconscious elements, referred to as 'The Self'.

7th : Failure to strike a balance between the archetypes or having poor or over-zealous relationships with the archetypes, leads to the development of complexes.

The core idea will be analysed in order to establish the fundamental issues with reconciling the theory of the archetypes with modern science, with occasional reference to sub-ideas 2 and 3; if this core foundation is not amenable to scientific practise, it is also unlikely that it would be possible to falsify those ideas which are contingent upon it. The bulk of the idea that will be discussed throughout the rest of the paper is summarised below, with bolded statements representing particular areas of interest.

'That there are **recurrent patterns of thought and behaviour** which arise from a source **outside** of the individual's own life experiences. The recurrent patterns are due to a substratum of the unconscious, known as the collective unconscious. Unlike the personal unconscious, the collective unconscious is an **inherited** structure bearing much resemblance to the biological instincts. The symbols that arise from this subconscious structure

(transformations) are recognisable by the **function** they serve rather than by the image they appear as, i.e. the Hero could be anything from Hercules to Harry Potter to an admirable friend.'

Empirical Science and its Demarcation Criteria

My aim is to use the statement above and to test it against the standards of modern scientific practise. This is in order to establish (1) where the original theory failed in meeting these standards, (2) where it can meet these standards at present with the aid of modern research, and (3) which elements of the theory continue to be outside the scope of empiricism today. To do this effectively, it is important to be absolutely clear on what the standards are for modern scientific practise. Pinning down what it is that makes a theory or practise scientific or pseudoscientific is no easy task, but it is a necessary one in order to attempt to discern which elements of the collective unconscious are amenable to scientific study in the current Anglo-American tradition (i.e. following the Anglo-French tradition earlier outlined, but which, in psychology, is now also largely championed by American Psychological Association).

Previous notable attempts at demarcation between good and bad scientific practise can be found in the work of Popper (1963/1998), Kuhn (1970/1998) and Lakatos (1973/1998). Throughout these works, the most typical example used for the sake of contrast to 'true science' is Astrology. This is because, although now widely recognised as a pseudoscientific practise, Astrology too once enjoyed some scientific credibility. Astrology is therefore set up as an example to illustrate what is seen as a mistake that was once made in the demarcation between science and pseudoscience, and which therefore provides a relatively uncontentious marker along the borderline between the two poles. The work of Popper (1963/1998), Kuhn (1970/1998) and Lakatos (1973/1998) all endeavour to explore this borderline and establish exact limits to avoid this mistake (of false knowledge being presented as science) being made in future scientific endeavours. In doing so, they attempt to set the standards for good scientific practise and elaborate on the expectations of the Anglo-American tradition, which we would hope would provide stable foundations for seeing where Jung's original theory did not meet hard scientific standards. The criteria posited by these scholars will now be

discussed with occasional reference to the Astrological comparisons made in the original texts.

Falsification

Popper (1963/1998) states that science and pseudoscience are differentiated by their use of the empirical method, and that true science arises from 'observation or experiment'. He criticises psychoanalysis on the basis of it having apparent explanatory power, but only seeming to accept verifications rather than refutations of the theory. He sets this up against Einstein's gravitational theory, and explores how in this case, as in many cases in the physical sciences, the theory is 'incompatible with certain results of observation'. This is known as falsification.

One of the adopted cornerstones of empirical science has long been the idea of falsification, which was Popper's (1963) suggested primary means for differentiating scientific from non-scientific practises. Falsification is the idea that a theory must have a set of conditions against which it can be tested and contradicted by a basic observational statement. This was suggested on the basis that we cannot confirm theories through observation, as a discrepant observation may always potentially occur. The most typical illustration of this stems from the 'all swans are white' example. The concept underlying this is that the statement could never be proven true as the evidence would have to be infinite (all swans in the world, from past, present and future) and attempting to gather this information would be unfeasible and leads to the problem of induction i.e., we cannot achieve firm knowledge when there is always the potential for contradictory evidence (Henderson, 2019). However, the all-statement can be rejected with one singular observation, e.g. of a single black swan. The observation of a single black swan shows that the initial statement is false as it is not a universal law, allowing it to be dismissed.

However, Kuhn (1970/1998) and Lakatos (1973/1998) criticise the idea that a theory should be dismissed and overthrown in the presence of contradictory evidence, indicating that many notable scientific theories such as Einstein's Relativity theory and Newtonian theory were resistant to refutation and dismissal, instead generating saving hypotheses or altering auxiliary hypotheses. It is also suggested that this is not a bad thing; in a theory's initial conceptualisation, inaccuracies and refining adjustments are an inevitability. To elaborate on the textbook example given above, the true answer is not that there are no white swans, for

example; it is far more nuanced than that. The answer is that black swans are far less common than white swans, that they are native to Australia and New Zealand (RSPB, 2007), and that there exist approximately 37 breeding pairs in the UK (McCarthy, 2012). Why is this relevant? Because the theory of 'all swans are white' had to be retained but moderated and tested in greater detail, including in a study of over 20 million birds (British Trust for Ornithology 2012, cited in McCarthy 2012), and has also been found to be dependent on variables such as geographical location. It is therefore not sufficient to say a single falsification alone is enough to dismiss an entire theory; 'all swans are white', when further tested, simply becomes 'most swans are white – with a number of conditions.' Therefore, while the initial statement is therefore false, it would not be correct to dismiss the theory entirely either, as its initial suggestion is still close to the truth and the observation was based in a reality - that most swans *are* white. And the same is often true and the mentality of 'adjustment, not overthrow' applied within psychological theory, as further explored by Kuhn (1970/1998).

Kuhn's (1970/1998) criticism of the emphasis on falsification arises from the fact that instances of theory overthrow (following testing) only occur when a field is already in crisis, and when there is a better alternative explanation for the exponents of the past theory to shift to. In support of this, Kuhn (1970/1998) highlights the way that commonly accepted theories may be tested, but if the evidence is unfavourable, it is often the practitioner that is blamed for failure to support it. Only after repeated failure does opinion begin to shift. These arguments, if accepted, indicate that falsification alone is not sufficient to distinguish a science from a non-science.

This said, generating theories with clear support and refutation conditions is a gesture of willingness to be disproved and a mark of scientific integrity; a hypothesis ought not to be a statement which one must dogmatically believe or disbelieve. Thus, while falsification alone may not be enough to singularly mark out a science from pseudoscience, it is surely an important factor to ensure good scientific practise.

It is therefore important to now ask the question 'to what degree is the theory of the collective unconscious and the archetypes falsifiable?'.

Falsifiability of the Archetypes

> 1st [Core]: All individuals are capable of engaging in recurrent patterns of thought and behaviour, identifiable in other times or cultures.

On the surface, the core idea appears to pass this check. Firstly, with regards to whether or not the 'true' condition can be observed, as already explored in the introduction, there are indeed observable patterns which can be identified across different times and cultures, such as the idea of the Mother (expressed as Earth Mother, Demeter, Virgin Mary, etc), the idea of the Shadow (expressed as Demons, Jinn, Asura/Rakshasa, etc), the idea of the Hero (expressed as Hercules, Joan of Arc, Ivan Tsarevich etc). As for establishing a false condition, we are also working with an 'all' condition which initially looks promising. Looking at old collective data (i.e. mythology) is little different to archival research. Accordingly, if we were to examine a society with largely the same living conditions as the others, but which did not produce archetypal ideas of any description (e.g. Jester, Shadow, Anima etc) then this would be counter-evidence for the theory of the archetypes, and the theory would indeed either need to be modified or discarded. If there are no identifiable patterns in the same manner as those that exist in other cultures, or if there exists a person who does not engage in any of these patterns, then the theory can be falsified as this provides a theoretical 'false' condition. In this sense, the theory is indeed falsifiable - provided 'true' and 'false' conditions can actually be observed in practise.

It is arguable that there are some issues with these conditions. For example, that we are projecting these themes on ad hoc, and may be subject to confirmation bias. That said, the fact that it is possible to group many of these concepts together is indicative that they share enough abstract properties and functions (e.g. to save, to trick, to nurture, to be above, to be below, light, dark etc) to group them and see them as one 'class' of objects. Similarly, we may see that an animal has four legs, barks, wags its tail, and is typically between our ankle and hip height and call it a 'dog'. Although the categorisation is ad hoc, we would also be correct to assume that all these animals had a similar 'root' or ancestry. Thus, while important to take into account, this in itself is not enough to rule out the validity of the observable similarity we can see in patterns of thought across culture, and there does seem to be an observable 'true' condition for the theory.

A much more critical issue is the testability of the theory; in order to properly falsify a theory, you must be able to test it, and clearly identify results which run counter to the hypothesis. For the archetypes, this is where the issues arise. Although we can *theoretically* say that 'all people can produce repeated patterns, and this can be evidenced', the original concept lacks conditions that are specified clearly enough to allow testing for counter evidence. To be able to actually observe a 'false' condition, an experimenter must first have some kind of stimuli to elicit the desired response. This would require a list or description of necessary stimuli which have elicited the behaviour in previous observations. Secondly, the experimenter would need a hard definition of observable outcomes so that they can differentiate an archetypal response from a non-archetypal response. There are multiple ways that this could be achieved. For example, a full list of archetypes and their potential transformations as they would appear in verbal behaviour could be provided (if the researcher is taking a more quantitative approach), or if they are taking a more materialistic approach, it would be important to have a list of brain areas or physiological responses that would be indicative of an archetypal response. This would allow the researcher to present the participant with a discrete stimulus and observe an outcome which is clearly either archetypal or non-archetypal. This information would allow researchers to make a clear all-statement which could then be clearly disproven by a future negative statement, e.g. 'all people will repeat ideas of the Mother, Shadow or Heroes in the presence of x stimuli', or even 'all people will repeat ideas of the Mother, Shadow or Heroes in the presence of x stimuli, if there is significant emotional impact containing these images in early development)', or 'all people will repeat ideas of the Mother, Shadow or Heroes in the presence of x stimuli, *regardless* of their early development'. The clear delimiting of which patterns we ought to see would allow us to look retrospectively through history and to mark a dividing line between ideas that will be repeated and prospectively to understand those that will not. This specificity would be essential in order for us to make a prediction and posit a hypothesis which can be supported (or refuted) by alternate observations, i.e. falsified.

This specificity is not provided in the original theory; what Jung (1936/37/1975) instead suggests is that "there are as many archetypes as there are typical situations in life" (p. 48, *CW* 9 pt. 1, para 99), and that they cannot be studied directly but only through the secondary phenomenon, which themselves are of "infinite variety" (Jung, 1954a/1975, p.81, *CW* 9 pt.1, para 156). This makes the concept difficult if not impossible to define in practise; if a church

can be a manifestation of the mother archetype, who is to say a pen, being vaguely phallic, is not a product of the Animus? That a cup, being a container, is not somehow a product of the Mother archetype, or the Anima? What a researcher is instead left with is a sliding spectrum of things that are 'more or less' archetypal rather than things that strictly are not. It is possible to make some inferences from the qualitative accounts, e.g. experiences of the archetype are associated with profound psychological impact and feelings of numinosity (Jung, 1956/1977), so we may presume that a Mother archetype will be *more* present in the Virgin Mary than in a cup. However, there is no reason to be found in Jung's work that a cup, too, could not represent the Mother archetype under the right circumstances. Particularly if that cup is warm, and if the edges of the cup are rounded, if the individual drinks from the cup and fills their mouth with liquid that will sustain them for days to come or protect them from harm. The form itself, the physical manifestation or observable phenomenon, is treated as irrelevant provided the core function and set of underlying attributes is met, as outlined in the 3rd core idea, i.e.

> 3rd : That these recurrent patterns of thought and behaviour (often manifesting as spontaneous images, known as symbols) can be broken down into distinct archetypes, recognisable by the function they serve rather than by the image they appear as.

Therefore, any observable outcome could be observed and interpreted as providing support for the theory, which would clearly be pseudoscientific if this testing approach were to be adopted. It is difficult to build a nomothetic law (or even establish a control group) upon the back of what appears to be personal associations, as the basic statement for contradiction becomes contingent on individual differences. Thus, Jung's unwillingness (or inherent inability due to the phenomena itself) to provide clear boundaries on what an archetype is, makes them extremely difficult to test and falsify.

This causes clear issues for testability; not only does the stimuli have no precise definition, but the outcome lacks definition as well. With no clear delimiting (because the transformations are dependent on individual differences and thus more of a relative truth than an absolute one), the theory is essentially reducible to 'unknown stimuli/infinitely variable stimuli causes unknown outcome/infinitely variable outcome'. This is an entirely unworkable state for any scientific theory and has ramifications for the testability of every other sub-

theory that branches off it.

To illustrate this, to test this, we will discuss the potential testing of the 2nd core idea.

> 2nd : That the recurrent patterns are due to a substratum of the unconscious, known as the collective unconscious. Unlike the personal unconscious, the collective unconscious is an inherited structure bearing much resemblance to the biological instincts.

This statement, outlined above, could be tested by observing whether or not a child who grows up without a mother would still develop ideas of the Mother archetype. If found to be true, this would perhaps indicate some kind of inherited underpinning as Jung initially suggested and provide support for the idea that the archetypes may work akin to biological instincts. In reality, this would be impossible to test because there is no clear list of observable outcomes which would indicate the presence of the mother archetype and no clear stimuli to test it. In addition, there are also 'saves' built into the theory even if a false condition was observable; e.g. a Jungian analyst might determine that there is no obvious material for the Mother archetype because it has been pushed into the unconscious. Despite the Jungian idea that this repression would cause an equal and opposite reaction, there is no irrefutable observation of what this would look like either, except a set of behaviours which may (or may not) be reflective of a mental preoccupation. What would be more useful here for clinicians or experimenters is a set of definitions for verbal behaviour e.g. number of mentions of a particular subject, or number of times topic is changed when specific subject is mentioned.

There are some areas where the prospects for falsification are more promising if the issues with the core statement could be overcome. For example, for statement 2 to be falsifiable, all instances of the archetype must stem from unconscious, rather than conscious processing or individual experience. To test this, instances of archetypal experience would need to be recorded, and the researcher would need to be able to determine that the basis of these patterns is not due to conscious or personal experiences. The basic counter-statement for this hypothesis would be that if these thoughts *are* conscious, or that if all 'archetypal experiences' are derived from personal experience, then there is no such thing as the

collective unconscious (because there would be no need for it, and its only function would be explained by already known structures). In this regard, this aspect is in line with Popper's criteria for falsifiability, as there is a significant element of risk involved, and practically determining that processes are unconscious rather than conscious is arguably possible with modern technology. In essence, brain activity in the absence of observable behaviour or self-report of experience to indicate consciousness, may suggest unconscious processes.

For example, a recent study into unconscious access to conceptual action and sound information explored unconscious priming effects (Trumpp, Traub and Kiefer, 2013). This study investigated the unconscious processes through topography and ERPs (event-related-potentials) and suggested different underlying neural generators for conscious and unconscious responses to stimuli. It also demonstrated that activity in sensory-motor brain systems could occur unconsciously, without imagery (in Jungian terms, this might be considered akin to the archetype without transformation). This suggests that it may be possible to tell conscious from unconscious processes activation using ERPs and topography. Therefore, at least in theory, a person who presented with a transformation of an archetype should be unconsciously primed to further archetypal content, which could be observed to be unconscious or conscious by the differing electrical potentials that occur in the brain. There is plenty of other research on differentiation conscious and unconscious content through neural-correlates, including tests of bistable visual phenomenon (where the stimuli is constant, but conscious perception fluctuates) such as the work on binocular rivalry by Clifford (2006) or the Necker cube, work by Dieter, Melnick & Tadin (2015).

In reference to the second core idea, determining whether the source of a pattern of behaviour or thought stems from an individual's personal experience (as opposed to the collective unconscious) is again more difficult to achieve reliably. Self-report measures are perhaps the easiest way to achieve, this, but if the experience has been acquired subconsciously, then it would not appear on this measure. Another method to help limit the potential knowledge the participant already has, could be by working with younger children. Jung even suggests this as a way to examine the archetypes, particularly in children 3-5 years old (1936/37/1975). However, in practise this approach could face difficulties with the child's articulation or focus, and even then, many ideas may be subconsciously borrowed from the child's environment instead, e.g. movies or books.

Additionally, even if the archetypes do stem from an inherited structure, this could be difficult to demonstrate in children. For example, if the archetypes do exist, and if they worked akin to Chomsky's theory of language (2013) (i.e. an innate, biological structure that develops reliably, like a limb), then there may be a structure that is inherent (i.e. the collective unconscious), but which is not observable until a certain age, e.g. when the child first begins to speak/when the child first produces symbolic content. For this process, there first needs to be sufficient environmental input to give the child the tools to express, and although this may theoretically be linked to an inherited structure, at this point it becomes impossible to discern the inherited aspect, as all that can be reproduced is what has been experienced, i.e. the transformation of the archetype. However, in Jung's original work, it should be noted he often refers to the mind like a "museum" (1964, p.57), or as a form of evolutionary structure with traces left behind. This is the core basis for his idea of the archetypes, i.e. that archetypes are can arise through images which are not consciously acquired, and which are expressed particularly in dreams or fantasies. Thus, theoretically, even a child who does not watch movies or have access to stories (which may indeed be an accessible population at pre-school age, particularly in areas of poverty or simply in cultures which are less media driven) should be able to dream or fantasise with archetypal content. Since verbal behaviour is observable, this could be one potential way to falsify the theory that the archetypes stem from a collective as opposed to personal unconscious, although the core issue of what is and what is not example of an archetype arises again here, so in practise it would still be near impossible to discern a 'true' or 'false' condition.

However, even with this already tenuous basis there are issues of clarity. It is not defined whether it is *only* the origin of the archetypes at birth that stems from the collective unconscious or if each experience of the archetypes must also stem from that same unconscious, collective structure. For example, if a person is conscious of a hero archetype in a piece of literature or film, can they then truly experience the hero archetype or are they precluded from this purely because the process is now conscious? If an individual specifically sets out to create a piece of art or a myth relating to an archetype, would this process too only be a pale conscious imitation, and not elicit the same neurological response as an archetypal experience? After all, Jung claims:

> No one can take a more or less rational thought, reached the logical conclusion by deliberate intent, and then give it "symbolic" form. No matter what fantastic trappings one may put upon an idea of this kind, it will still remain a sign, linked to the conscious thought behind it, nor similar hints at something yet unknown. (Jung, 1964, p.41).

However, in the case that it is only the original basis which is intended to be unconscious (as opposed to the cause for each experience of the archetype), then the entire method of testing could be different entirely. Jung's theory remains vague and exploratory in this aspect, without committing fully to one or the other, leaving the route to testing unclear.

In its original form, the theory is not specific enough to be falsifiable. Although there often seem to be potential or theoretical routes for falsification (e.g. for possibly identifying unconscious elements), all roads to testability are effectively blocked by the issues outlined with the very core of the idea due to striking problems with clarity and an unwillingness or inability to define terms, as well as issues with falsifying the basis of the inherited and 'collective' claims originally set out. Of course, this is not to entirely discredit the life's work of Jung. There is a phenomenon that is observable, and which lacks explanation - a collection of repeated patterns that seem to transcend time and culture. The theory of the collective unconscious and the archetypes begins to explore this and has some valuable things to offer, but despite Jung's intentions, the theory, in its original form, does not appear to be falsifiable and thus not amenable to scientific enquiry.

Why Revise the Theory? : Alternate Demarcation Criteria

If the original theory is not falsifiable, why then is it worth re-examining? It is impossible to ignore the immense importance of falsifiability for testing, and for establishing a flourishing research programme. However, as was previously mentioned, falsification alone is not enough to mark out whether a theory is science or pseudoscience; there are other approaches to demarcation which may be useful for understanding the research culture around the phenomenon Jung originally set out to explain. By exploring these alternate criteria in a revision of the theory, we enable ourselves to expand within the realms of good scientific practise and, where possible, develop the theory in line with this criterion. This in

turn may produce a result which is more amenable to scientific enquiry and, as will be suggested, may actually bring us a step closer to falsifiability itself.

By contrast to Popper's (1963/1998) interest in revolutionary theory overthrow, Kuhn (1970/1998) focuses on what he refers to as 'normal science' and problem solving, which is essentially the work done within a theory, taking the general premise for granted. By engaging with the puzzles which challenge the theory, the scientists in question slowly expand the theory's explanatory power. Kuhn (1970/1998) argues that the issue that emerges in non-sciences, like Astrology, is that the puzzles remain unsolved for an extended period and are often placed outside of the researcher's control and resultantly no attempt is made to refine tools or theory in order to increase its predictive power. The link between predictive power and scientific status of a theory is a simple one, which Ruse (1982/1998) underlines quite clearly; science is about establishing universal laws. Thus, the presence of unsolved puzzles for extended periods, or puzzles not being solved for reasons that are placed outside of the researcher's control, are often indicative of pseudoscientific study. Therefore, concerted effort to resolve puzzles, and to bring factors set beyond researcher control back into the realm of the testable appears to be a reasonable mark of a progressive scientific programme.

Lakatos' (1973/1998) criteria is similar to Kuhn's, in some respects, as both centre around the idea of progress. However, in Lakatos' case, the hallmark of science is that the theory progresses enough to predict novel facts and provide an explanation for phenomena which other theoretical approaches cannot. It is also stressed that refutation is not the mark of total failure; theories fail instead when they lag behind the facts. To illustrate this example, Lakatos (1973/1998) highlights the persistent failure of Marxism to predict future events predictions (i.e. that socialist societies would have no revolutions, or that the first socialist uprising would occur in the most industrially developed society), instead generating ad hoc explanations for political outcomes. However, no substantial alterations are made to the theory to account for these failures and the innumerable explanations for the failures make it too difficult to pin down exactly which part of the theory needs alteration in the first instance. Therefore, regardless of the significance or usefulness of the theory, the number of anomalies without reasonable adjustment to the theory make its persisted use untenable. As Ruse (1982/1998) highlighted, since science is focused on establishing universal laws, it does not

allow for breaks in these laws, thus, if the law were true, the theory could account for any outcome, anomalies included. Thus it is clearly important that any scientific theory should be able to make accurate predictions about reality; it is not the presence of refutations that indicate empirical failure, but the absence of predictive successes (Lakatos, 1973/1998).

From these three approaches (that of Popper, Kuhn and Lakatos), there are a number of general principles which seem to define the sciences from the non-sciences, and suggest what allowances ought to be made for those theories that are in their early development. For example, a theory must produce a testable hypothesis. Regardless of whether or not it is refuted, the very act of doing so is a mark of scientific integrity and this act is the beginning of a process towards establishing a flourishing research program. If a theory is refuted, it need not be instantly abandoned. Instead, modifications can (and sometimes should) be made to the theory; these are research puzzles, and in this time, the scientist may take the general premise as true. There is the clarification point that researchers cannot be selective about the evidence which emerges. Finally, after a period time that is unspecified here, this theory must be able to produce novel predictions and provide explanatory power that other competing theories cannot. If it cannot do so and the puzzles remain for a long time, it may be noted that the research program is stagnating. At this point, changes are ordinarily made or else the theory begins to go into decline. At this point, it is still not unscientific to remain with the theory, as many scientific theories still face unsolved problems and the reformation of the research program is still an option. The point at which it becomes unscientific is when an alternate theory emerges which is more successful at explaining and predicting the phenomena in question, but the less successful theory is still retained by the community in favour of the one that is objectively more successful. The most profitable course of action at this point, if the researchers involved consider themselves to be scientists, is to shift into testing the new theory. It is this point in particular which Astrology failed upon, as the emergence of early psychology and personality theories and psychometrics (such as Jung's MBTI test) provided fuller and more plausible explanations for personality development.

One notable (and successful) example following these research guidelines is Newton's theory of Universal Gravitation; it was first published in 1687 and struggled to fully and satisfactorily explain action at a distance, which was not resolved until the addition of Einstein's theory of General Relativity, published in 1920. This demonstrates the length of

time that a scientific theory can take to resolve internal puzzles, and to make predictions which provide a fully accurate model of reality. Despite there being certain research puzzles unsolved, continued effort was made within the research community to resolve it, with steps clearly taken to understand anomalies. The theory of the collective unconscious and the archetypes by contrast is therefore very much in its infancy, and despite current issues, there is no reason at present why there might not still be a successful research program developed if the issues with falsification can be resolved. As Lakatos (1973/1998) highlights, even in the presence of so many refutations, it is not unscientific to stick to a degenerative programme if steps are being made to try and make it empirically progressive. That is to say, even with the current issues of falsification, a programme of study around the archetypes can still be scientific – but only if researchers continue to try and resolve these problems. It is this category which the current thesis aims to be part of.

From this aggregate of the above approaches it would seem that the difference between science and non-science is not defined by the theory alone, or whether or not it is falsifiable, but the entirety of the research process surrounding it. It is also notable that the differentiation is context sensitive; a theory can be scientific at one time and non-scientific at another, which was the case for the once credible Astrology.

However, there is another particularly interesting criterion from Thagard (1978/1998) which incorporates some elements of the previously mentioned as well as adding its own contribution. It is a little more complex than those listed above, a tripartite matrix, which I feel necessitated the summary before this addition.

The three elements posited for examination are theory, community and historical context (Thagard, 1978/1998). The three elements will be divided below and the relevant areas of interest suggested by Thagard (1978/1998) will be outlined.

Theory: According to Thagard, a theory is scientific if it is structured so that it may be falsified, if it can make predictions, if it has significant explanatory power and if there is an element of problem-solving involved in its development (see Kuhn's 'normal science'). Ideally, there should also be a clear, plausible mechanism by which the posited relationships are possible. However, this is not essential; as Thagard (1978/1998) highlighted, the links

between smoking and cancer are well-accepted although the physical means are not fully understood.

Community: Scientific versus non-scientific communities are marked out by their agreement on the core principles of the theory and having a cohesive narrative about how to solve the problems the theory faces. The research community should also be involved in the explanation of anomalies and demonstrate an awareness of other emergent theories and lastly, the community ought to be actively involved in confirmation and refutation of their theory.

Historical Context: Here, Thagard (1978/1998) draws on the work of Kuhn and sets up an additional two considerations: the theory's ability to explain new facts and capacity to deal with anomalies, and whether or not alternative (and more successful) theories are available; contrary to Lakatos' demarcation, the presence of long-standing unsolved problems are only an issue when there is a more viable alternative available.

This is a somewhat oversimplified overview, as there are arguments to suggest that it is not possible to determine pseudo-science from science. For example, the concept of epistemological anarchy posited by Feyerabend (1975) suggests that scientific process is not guided by rational thought but is instead a more chaotic process which ought not to be dogmatic and rigid. However, assuming all theory adoption and abandonment is a matter of mere chance and that progression of any kind is illusory is unhelpful in discerning what works and what does not in the ongoing attempts to attain truth and control in an otherwise admittedly chaotic world. It is better to work with an imperfect model, in this respect, than to proceed without one at all. Without some clear shift towards falsifiability, other routes to scientific enquiry are also barred; we cannot problem-solve as recommended by Kuhn because the problems cannot be put into clear, testable terms; we cannot make predictions as recommended by Lakatos because what we are predicting is unclear, and false outcomes are impossible to discern; and there can be no active research community with agreed upon principles as suggested by Thagard (1978/1998), because, again, the route to testing is barred. It is thus my opinion that in its original form, the theory is not amenable to scientific enquiry. In light of this, laid out below is the demarcation criteria which I intend to adhere to and which I believe should underpin further research of the archetypes. By following this

criterion, it should be possible to develop a variation of the theory which is amenable to modern scientific practises; without doing so, the field risks further degeneration.

1) The theory must have some clear hypothesis/hypotheses and a null for each.

2) The hypotheses must be clearly testable in a way that can be repeated and verified by other practitioners. If the hypotheses cannot be tested with current means, there ought to be a clear indication of how it could be tested, e.g. with the right technology.

3) Researchers ought to be actively involved with testing to support or refute the theory.

4) A theory must be falsifiable with observable 'true' and 'false' conditions with clearly defined terms. All evidence must be retained regardless of if it supports or refutes the theory. However, refutation is not necessarily failure; auxiliary hypotheses may be amended to account for the changes in an attempt at refinement of the theory.

5) There ought to be an indication for the physical means for the link between the cause and effect relationship posited. While this is not crucial, the presence of a physical link will help to ground the theory in physical reality, which is much in line with the current scientific paradigm.

6) The core principles of the theory should be largely agreed upon by the researchers involved in the research program.

7) If there is no attempt to change the theory in light of counter-evidence, then the community ought to at least be engaged with explaining the anomalies. This may be in line with 'normal science', which does not necessarily have to dramatically challenge the given theory or paradigm.

8) There ought to be evidence of the community examining other theories and explanations for the phenomenon which they are attempting to explain, and perhaps reconsidering their own theory in response to this.

9) After a grace period for refinement, a theory ought to be able to predict novel facts. If a more successful theory emerges which provides more explanatory and predictive power, and the current theory also faces unsolved problems for long periods with little progress, then it is unscientific to maintain this theory.

From Then to Now: Our Criteria and the Theory of the Archetypes.

Of course, much has changed in Psychology even since the time of Jung, and there are issues the original theory faced which are possible to overcome today. Before turning to modern research to begin revising the theory, it is of interest to compare the old to the new, and see to what extent the constraints of the past are still relevant to the archetypes.

It is notable that Jung himself acknowledges the dominant paradigm of his time and makes a certain critique of it that is relevant to the current discussion. Jung states that:

> the thesis runs as follows: we accept as valid anything that comes from outside and can be verified. The ideal instance is verification by experiment. The antithesis is: we accept as valid anything that comes from inside and cannot be verified. (1954a/1975, p.76, *CW* 9 pt.1, para. 149).

Jung goes on to call this position as "hopeless" and laments the triumph of Aristotelian reasoning (i.e. mathematical logic) and Greek empiricism over Plato (whose theory of Forms in Phaedo suggested that there is the ever-changing material world, grasped by the senses, and the inner world of concepts such as justice and beauty, which are real in themselves and reachable through thought alone). The quote appears to lament the exclusion of the introspective and metaphysical from scientific study. Indeed, this narrative may be considered quite damning for the archetypes; the outlined paradigm almost prohibits the study of intra-subjective material, which a study of the human psyche (and certainly that of the archetypes) is arguably incomplete without.

In terms of direct reports of the subjective experience, it may now be argued that we could produce interviews and deploy self-report measures, and correspondence between these could

be taken as verbal behaviour in *response* to an experience. However, it remains the case that that 'from within' is taken merely as a response rather than reflection of reality, as it is difficult to objectively verify that these accounts represent the experience itself (i.e. a verbal account of motherhood is not considered a representation of motherhood, but an account of the subjective experience of motherhood). As a result, 'that from within' is still not taken to be an accurate representation of reality as may have been Jung's wish judging by his fondness of Platonism (Mills, 2012).

This said, the inaccessibility of the inner world has vastly changed since Jung's time and consciousness is no longer in the 'black box' that it used to be; accordingly, the narrative has shifted somewhat, and 'that which comes from inside' is no longer considered wholly unverifiable from a biological standpoint. There are now a myriad of techniques available to study the brain and potential neural-correlates of consciousness, which range from positron emission tomography (PET scans) and computed axial tomography (CAT scans), to electroencephalograms (EEG), functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI). Studying the archetypes scientifically is now more plausible than before (provided the researcher accepts the premise that the brain and mind are related), and these possibilities will only continue to grow as technology advances. Additionally, as behaviourism has declined and humans are no longer seen as purely rational actors, systematic qualitative approaches have increased in popularity (such as discourse analysis, ethnographic research, focus groups, etc.), and are more often seen running in tandem with empirical methods; it is easier now than before to begin to connect qualitative experiences with their potential neurological underpinnings.

It is important to note here that some of the failings in meeting scientific standards belong not only to the archetypes but to the entire psychoanalytic practise of the past. These are highlighted by Bornstein (2001;2005), who remarks on the disconnect between psychoanalytic psychology and other modern research branches. In an earlier paper, Bornstein (2001) attributes the failures of the research programme to the following 'seven deadly sins': insularity, inaccuracy, indifference, irrelevance, inefficiency, indeterminacy, and insolence. The question thus arises of whether it is possible to overcome this issue in the modern research programme, as the rooting of the archetype theory in this field has therefore led to similar issues.

However, as is highlighted by Popper (1963) and Thagard (1978), many theories have roots in the mythology and the occult, such as the role of alchemy in the development of chemistry; that does not prevent them being scientific today. This is not to say that psychoanalytic practise is mythic or occult, but simply to highlight that origins of a theory are irrelevant to the way they may be later transformed. Post-Jungians, for example, now have a mixture of biological, developmental and cultural adaptations of Jung's original theory. At a glance, this seems to violate the need for a community to be cohesive and unified in their basic principles. However, the diversity present in post-Jungian explanations is not necessarily an indication of discord in the community; each of these should be treated as a competing and occasionally complementary research programs. Additionally, many ideas from the psychodynamic school of thought have been prematurely abandoned, discredited, or plagiarised (Bornstein, 2005). For example, the traditionally Freudian 'ego' has been rebranded as Baddeley's (1992) 'central executive'. The re-emergence of these concepts, albeit under different names and in different branches of psychology, indicates their continued scientific relevance. The development of competing research programmes (despite testing issues) and the persistence of similar ideas from the psychodynamic school of thought indicates that it is possible to revise the theory in such a way that it may be possible to re-enter mainstream psychology, and to establish a more active research programme once more, despite its roots in a typically more insular field.

In its initial conception, Jung did not manage to present us with a theory that allows us to empirically examine the archetypes (even with the more modern qualitative approaches available to Psychology), there are important exploratory steps made and elements of plausibility in the explanation Jung describes. As in the all-white-swans example suggested earlier, there may be at least a grain of truth in the original theory, and parts of the puzzle which might be accounted for by Jung's suggestions. The issue for the theory of the collective unconscious at present is that it fails the very first hurdle of being amenable to scientific enquiry – it does not appear falsifiable. Falsification may not be the sole divider, but it is necessary to enable the other pursuits which divide science from pseudoscience, which have been outlined above. We must be able to *test* the theory – and currently, this is what the theory of the collective unconscious and the archetypes is missing. There is an active community who are interested in engaging in normal science with Jungian theory, but

the path to testing is unclear, leading to disunited practitioners and theorists with conflicting narratives about what the archetypes are and how they can be tested. Even in modern papers, there are very few attempts to empirically explore the archetypes, which arguably is because there is no clear route to do so, and due to the disputes regarding ontology which have been explored above.

However, there are a few papers which have suggested revisions of the theory based on other modern findings, which may provide a path for falsifiability, but these provide no actual explanation of how this might work beyond a reference to the theories they draw on. The remainder of this paper will be dedicated to doing just this – outlining and exploring modern interpretations of the theory of the collective unconscious and the archetypes, and then using the ideas suggested by the modern Jungian community and alternate research communities to create a model of how the core phenomenon might occur. This will ideally account for culture, experience and biology. It is my hope that this would enable the Jungian community to engage in normal science, with a falsifiable theory which remains true to the core elements of its predecessor, and which does not reduce the archetypes to pure experience or pure biology, in accordance with my reading of Jung's original wishes.

Modern Conceptualisations of the Collective Unconscious and the Archetypes

To begin to approach this idea afresh, I am first going to strip the idea back to its very basics, and examine how, in light of modern psychology, this phenomenon could come to be. The phenomena we are looking at is, as has now been said many times, recurrent patterns of thought and behaviour typified in mythology, dreams and fantasy. In its original form, the theory of the archetypes was not able to address this issue scientifically. My aim is to utilise theory of post-Jungians such as Knox (2004) and Merchant (2009; 2016), combined with works in neuroaesthetics and the recent work of Lakoff (2015) to provide a testable alternative to the original theory, while remaining true to its basic foundation. The parts of the original theory that simply are not needed to explain the phenomenon will be discarded, as it does not make logical sense to retain the parts of the theory which are not somehow required to explain the behaviour that support the phenomenon, i.e. if variable A accounts for all the variance and outcomes, there is no need for a variable B.

The most useful and logical place to start this process is by looking at the major adaptations

of the theory since Jung's original conception of the archetypes; they are no longer considered undoubtedly innate, even by Jungians. This is not to say it is an abandoned position, as evolutionary psychologists such as Stevens (2006) and Goodwyn (2012) uphold that there may yet be an evolutionary basis for the archetypes. However, researchers such as Knox (2004), Sotirova-Kohli, Rosen, Smith, Henderson and Reece (2011), and Merchant (2009;2016) contest this and suggest that the archetypes are developmental. In revising the theory, an initial position on this must be taken and supported by current scientific theory.

In the case of innate archetypes, we must entertain the possibility that there are mental instincts in the same way that there are physical, and that humans possess some core similarities in emotional displays. This may be considered something akin to a human condition which is unchanging over time and culture. Goodwyn (2012) suggests that this is plausible and that a core psyche may be due to a 'genetic bottleneck' 200,000 years ago which makes humans more similar to each other than is the case in other species. Contrariwise, in the case of developed archetypes, the possibility we need to entertain is that regardless of temporal and cultural changes, all humans have a shared experience which causes them to generate images (transformations) which stem from the same core archetypes or function.

The most salient theory which has emerged (and which both camps appear to entertain) in post-Jungian research is that archetypes are a kind of image schema (i.e. abstract mental representations of embodied experiences, arising from our interactions with the environment). Knox (2004) suggests that the archetypes are emergent cognitive structures, resulting from the individual's interaction with the environment. The theory draws on Lakoff and Johnson's (1999) work on metaphor, which suggests that our physical interactions with the world are the primary determinant in the way that we come to visualise situations or concepts which are difficult to articulate. These can be visuospatial in nature, for example that happy is up and sad is down, or somatic, such as affection as warmth (Goodwyn, 2012). Echoes of this can be seen in tales of katabasis, a descent into the underworld to retrieve something lost. For example, in the Sumerian Epic of Gilgamesh, the King suffers the death of his companion, Enkidu, and embarks on the journey down to find the secrets of eternal life, or in the Greek Orpheus wherein the hero attempts to steal back his wife Eurydice from Hades. It is grief and sadness of some kind that drives the protagonists down in each case, which appears to

support the idea that there is a shared association between spatial positions and emotions (sad is down). Lakoff (1997) highlights the usefulness of this; our metaphors become a vehicle for finding meaning in our own lives, a desire which is also relevant for understanding our biological evolution as thinking, remembering and dreaming organisms (Stevens, 2006).

Since these image schemas provide a reliable sensorimotor scaffolding for the world, as all individuals experience various environmental interactions with concepts such as 'up', 'down', 'out' or 'path' through their own spatial navigation, Knox (2004) argues that this may be the basis for the archetypes. What we are seeing in the archetypes, therefore, is not necessarily a spiritual, archaic experience in the sense Jung originally intended, but a universal metaphor pattern produced by early psychological self-organisation processes, interplaying with physical laws. Further to this, Merchant (2009;2016) states that this basic template of understanding is produced when a person is pre-verbal, a time of extreme affectivity, and thus the experiences are encoded primarily with visual and affective associations (Sotirova-Kohli et al., 2011), which may in part explain the numinous experience that Jung describes as being associated with the archetypes (Jung, 1964) and their visual manifestations.

The divide between the developmental and evolutionary position is not a clear-cut dichotomy. It is more a case of arguing whether this developmental experience merely shapes the means of expression for the archetypes (the evolutionary position), or whether it *is* the entire archetypal basis. In the case of the former, the experience of the archetypes is due to evolutionary history and is still beyond the realm of individual experience (Goodwyn, 2012), while in the latter the archetypes may be reduced to the re-experiencing of pre-verbal memories and thus entirely belong to the personal unconscious (Merchant, 2016), which indicates that concepts of innatism should be abandoned (Knox 2004; Merchant 2009). This may be considered something like a hard or soft influence debate around the influence of early sensorimotor experience on archetypal phenomena.

If it is the case that archetypal experience is simply a universal metaphor, the question arises of why some metaphors tend to be expressed using the universal basis while others are not – for example, the idea of setting a baby down a river, appears in the Kanji for '*nagareru*', the Japanese tale of Izanagi and Izanami and the Christian tale of Moses being set

down the river (Sotirova-Kohli et al., 2011). The connotations of abandonment are prevalent in each. By contrast, many metaphors use specific cultural artifacts as their basis and are not translatable to other cultures, such as references to a ‘one way ticket to x’ or ‘pancaking’ to describe a flattening effect (Merakchi & Rogers, 2013), which indicates that not all metaphoric expression utilises a universal framework. Further work on this subject has been conducted by Kövecses (2005), who suggests that metaphors can vary based on region, style, subculture and individual. Where universality does exist, the explanatory power of the sensorimotor explanation is frequently limited to small visuospatial suggestions, such as ‘sadness is down’ and while this explains elements of mythology such as the journey down in katabasis, it does not explain the reoccurrence of the overall tale or its interplay with the idea of the hero.

It should also be noted that this is only a partial explanation of Jung’s original theory; there is no mention of enantiodromia and no clear attempts to identify which archetypes are present in the case studies of the most recent developmental literature. The interplay between the different archetypes within the self is of critical importance in Jung’s work and since it is the attached emotionality and connection with affective states which give the archetypes their psychological significance, it is arguable that at present this theory lacks the explanatory power of its predecessor.

Despite these drawbacks, the theory that archetypes are rooted in metaphorical thought, (which in turn are based upon developmentally based image schemas) is a significant contribution to Jungian psychology, providing researchers with means to establish expectations of what we might see when observing the archetypes and may provide the groundwork for understanding the archetypes through neurobiology. Considering the fantasies, fairy-tales and myths that the recurrent ideas are visible in, it does seem plausible that this could be a metaphorical mapping of emotional experience; the primary questions that arise are ‘why are those particular metaphors universal?’ especially because many seem far more complex than usual linguistic metaphors, and also ‘how could that work in a way that fully accounts for the observations made in the original theory?’.

If the outlined theory is correct, then this could go some ways towards resolving the issue of defined outcomes which was discussed earlier, i.e. the brain areas activated during archetypal

experience ought to be largely akin to metaphor processing; as is pointed out by Goodwyn (2012), the archetypes (or rather, their transformations) act as metaphorical representations of lived experiences in the world, which is highly akin to image-schemas. Additionally, in Jung's explorations of archetypes in dreams, he often explores how the transformations are used to express personal issues that are below the level of conscious awareness. For example, Jung (1964) suggested that those who attempt to live beyond or above their capacities often have dreams of flying or falling. This is non-literal, and Jung interprets this as a figurative representation of their situation, which is meant to act as a warning for the dreamer. Jung refers to this as the compensatory function of the archetypes. This is highly akin to metaphor processes, where the properties of the figurative situation are used to provide a fuller explanation of a literal scenario. As a result, we would expect to see similar brain processes involved as in that of a standard metaphor.

If they are not similar to these brain areas, then the archetype is not a type of metaphor and we would need to revise the underlying mechanisms of this phenomenon. If archetypes *are* reducible to universal metaphor alone, then this should be the only activity we see. However, if there is a difference between standard metaphor processing and the archetypes, then we may expect to see additional arousal or other brain areas involved with processing the archetypes in comparison to the neural correlates associated with metaphor (provided stimuli can be found to trigger archetypal experience). To test this, it is first necessary to explore to what extent theories around metaphor processing could explain the archetypes.

Metaphor Processing

Although primarily studied as a linguistic device, metaphor has been increasingly studied by psychologists as a conceptual tool to represent situations we are yet to understand or find difficult to articulate (Ma & Liu, 2008). When utilising conceptual metaphor, we are figuratively comparing two different objects, using the connotations of one to explain the qualities of the other. Through this process, we are able to elucidate sensory and affective details of abstract concepts with ease, for example in a phrase like 'the warmth of her love', the target 'love' is linked with the source 'warmth', setting the narrative of 'love as temperature'. This can be seen in other 'love as temperature' metaphors, such as 'she was acting cold' to describe a situation where love or care seems absent. This source-target domain mapping thus expands our capabilities to reason about abstract targets (Shutova, Teufel, Korhonen, 2013). It could be argued that this is merely a comparative likening of one

object to another, and there is some in-field debate over whether a metaphor is simply an abbreviated form for a simile, where again one object is likened to another. However, Searle (1979) highlights that metaphors made into similes are frequently less informative than the former. For example, one might say of a lover, 'that person is a home to me'. Here, the meaning is quite clear; the person in question serves the function of a home, they may keep the speaker safe or bring them a sense of comfort. By contrast, saying 'this person is like a home to me' begs the question of 'in what way?'. This suggests that the conceptual metaphor, while similar, is separate from the simile and serves a distinct psychological purpose of its own; recent psycholinguistic models suggest that beyond similarity, metaphor is a form of class-inclusion and is better understood as a categorisation process than superficial comparison (Shibata, Toyomura, Motoyama, Itoh, Kawabata and Abe, 2012).

Although the language domain is associated with the left-hemisphere, it is the right hemisphere which has been considered to have particular significance during metaphor processing (Schmidt and Seger, 2009; Zeev-wolf, Faust, Levkovitz, Harpaz and Goldstein, 2015). The reason for this is thought to be because the left hemisphere produces a narrow semantic field of associations to support further processing (fine semantic processing) while the right hemisphere generates a larger (but weaker) field of associations which may only be peripherally associated with the initial stimuli (coarse semantic processing) (Beeman & Bowden, 2000). Since the language used in metaphor is figurative, broader associations need to be made in order to ascertain the meaning of the statement and as a result of this, increased right hemispheric activation is frequently seen in the processing of metaphors.

Consonant with this theory, patterns of lateralisation in language comprehension have been found to be reversed in schizophrenic patients, with higher right-hemispherical activation found during typical language processing. This has led to suggestions that schizophrenic patients are over-reliant on coarse semantic coding (right hemisphere) during initial exposure to stimuli, which can lead to loose association patterns (Zeev-wolf, Faust, Levkovitz, Harpaz and Goldstein, 2015). This has both positive and negative outcomes; schizophrenic patients are able to notice remote relations between words which neurotypical subjects may not be able to, but are also unable to meaningfully distinguish between related and non-related items due to a failure to recruit the left hemisphere for fine semantic coding.

Right-hemispheric preference has also been found in the processing of certain image-based alphabets, such as Kanji. Since Kanji is derived from mythology, this graphic language serves a symbolic purpose rather than simply phonetic; kana (hiragana and katakana) are by contrast phonetic and show left hemispheric preference during processing (Sotirova-Kohli, et al. 2011). Thus it seems to be the case that it is not the image alone but the associative process required to understand it which gives rise to right-hemispheric preference.

The role of the right-hemisphere in metaphor processing has been further expanded on in recent research which indicates that familiarity is a key mediator in right-hemispherical involvement. Schmidt, Debuse and Seger (2007) suggest that it may not be the metaphor itself which gives rise to this processing, but any novel sentence containing a distant semantic association since this is in accordance with the lateralisation suggested by Beeman and Bowden (2000). Schmidt and Seger (2009) tested this further to assess the role of figurativeness and familiarity and found that metaphors (by contrast to literal sentences) recruited the right insular, left temporal pole and right inferior gyrus. Metaphors which were familiar showed activation in the right middle frontal gyrus whereas difficult metaphors showed activation in the left inferior frontal gyrus. Since this area is associated with response inhibition (Swick, Ashley and Turken, 2008), it might be suggested that the left inferior frontal gyrus is activated to prevent undesirable learning while assessing whether or not two items are related, though this is only speculative. Schmidt and Seger's (2009) study also includes a table (see appendix) outlining previous studies on metaphor processing, including cases where no right-hemispherical activation was found. In these cases, the metaphors used are typically embedded in simple sentences and/or are conventional or familiar. This indicates that sentence complexity and familiarity are important determinants of right hemispherical activation, with novel and higher complexity metaphors being more likely to give recruit the coarse semantic coding of the right hemisphere.

The left inferior frontal gyrus also seems to play a role in metaphor processing (Rapp, Leube, Erb, Grodd and Kirscher, 2004; Kirscher, Leube, Erb, Grodd and Rapp, 2007; Stringaris, Medford, Giampietro, Brammer and David, 2007). This has been observed in schizophrenic patients, for whom metaphor processing is often faulty, leading to concrete or literal interpretations (Rapp et al., 2004)). Kirscher et al. (2007) conducted a study comparing the fMRIs of a control group to a group suffering with Schizophrenia. The groups

were processing unfamiliar, non-salient metaphoric sentences compared to literal sentences. The results suggested that the patients demonstrated increased activation in the left inferior frontal gyrus (BA 45, an area thought to be associated with semantic inference processing (Rapp et al., 2004)) which was significantly negatively correlated with concretism. Although both groups experienced the increased activation in the left inferior frontal gyrus in the metaphor condition as opposed to the literal sentence condition, schizophrenic patients exhibited this roughly 3cm dorsal to the control group. This displacement combined with the frequent loose associations in schizophrenia may tentatively be considered as support for the prior speculation that this area if used for the decision-making process in accepting relations between items in difficult metaphors. This has some support given that lesions to the left inferior frontal gyrus also lead subjects to perform significantly worse than control subjects on Go/No-go tasks which require greater inhibitory control (Swick, Ashley and Turken, 2008). In addition, the control subjects also demonstrated higher activation than patients in the right posterior temporal cortex and the right middle/superior temporal gyri (associated with syntactic processing) while reading metaphors (Kirscher, et al. 2007).

Interestingly (in conjunction with the sensorimotor theory underpinning the study) the angular gyrus is also purportedly involved in metaphor processing (Bonner, Peelle, Cook and Grossman, 2013). This is notable because the angular gyrus in particular has been associated with sensorimotor cognition when participants are asked to imagine actions or the features of objects (Bonner et al. 2013) and is the most commonly activated region found in semantic memory studies (Binder, Desai, Graves and Conant, 2009). The angular gyrus is associated with sensory integration in the heteromodal cortex and is thought to give rise to higher order multisensory perception. In relation to metaphor and the angular gyrus, there is a study which involves two shapes, originally designed by Köhler (Ramachandran & Hubbard, 2003) which is of particular interest. In the study, one of the shapes presented is jagged and sharp and the other soft and rounded. When asked which is the ‘kiki’ and which is the ‘bouba’, most people will respond that the soft and rounded shape is the bouba. Ramachandran and Hubbard (2003) suggest that this is due to the soft, rounded shape of the lips involved with the auditory input of ‘bouba’ resembles the soft roundness of the shape in contrast to the sharper ‘kiki’. It is then suggested that that the connection between these abstract properties is made due to them creating similar waveforms which meet in the angular gyrus. Ramachandran and Hubbard (2003) also point out that the ‘kiki’ and ‘bouba’ effect does not work on those who

sustain damage to the angular gyrus, indicating that this may be the area responsible for our ability to process abstract properties and thus a key area for metaphorical processing.

In the case of the archetypes, since it is frequently suggested that the experience of this phenomena feels as if it comes from outside the self (Goodwyn, 2012), we might presume that the metaphor would be classed as complex or at least unfamiliar and would thus recruit the right-hemisphere and the left inferior frontal gyrus. We would also expect to see the angular gyrus activated in the process of recalling vivid sensory experiences. In cases where the symbols appear to derive from the personal unconscious, we may expect less right hemispheric activation, or potentially none at all because the individual would be drawing on existing memories rather than drawing on creative faculties to create a novel symbol. This may also require less interpretation and cross-sensory integration than novel metaphor generation.

This theory goes some way towards solving one of the main issues of the original theory of the archetypes; if the archetypes are some kind of metaphor, this gives us a more concrete idea of the outcome we might expect to see, particularly in terms of neural correlates. However, as Hogenson (2019) points out, the theories which suggest the archetype is emergent from interaction with the environment do suffer from being somewhat divorced from the original theory of Jung because of the continued failure to set clear dividing lines up to set archetypes apart from other phenomena, and because of the absence of any account for the collective unconscious. Indeed, Knox's attention shifting to attachment theory and child development (Hogenson 2019) seems almost reductive when set against a vast cross-cultural history of myth, symbols and dreams, and their visceral impact on the psyche.

This said, the emergent theory posited by Knox (2004) is a great contribution to the study of the archetypes. For example, it goes some way towards resolving the issue of an undefined outcome for a 'true' or 'false' condition, as if the archetypes are expressed through metaphor, we would expect to see activation in the brain areas outlined above. However, as Hogenson (2019) points out, the theory does remain divorced from many of the original complexities of the phenomenon and fails to resolve many important questions about the archetypes. For example, if they are both rooted in image-schema, how is an archetype phenomenally different from a standard linguistic metaphor? How do we elicit an archetypal image as

distinct from a standard metaphorical image? Despite providing some clarity, this again leads back to the same issues for testability outlined earlier in this work; the archetypes remain undefined, and as a result, difficult to elicit and discriminate from alternate phenomena. The emergent theory also does not particularly explain the affective connection between the (subconsciously) selected image and the archetype it represents; if the Mother archetype is purely based on containment, for example, why is a church an adequate symbol, but not a skyscraper or a restaurant? Additionally, why do the archetypes remain unconscious while other forms of metaphor are conscious enough to appear in our verbal behaviour? There are some elements of the theory which also seem confusing and discordant with their supposed basis; Knox appears to entirely reject innatism, and yet there are innate elements which underly metaphor, and the universal behaviours which support early development of image-schemas in the first place. What would be more helpful is a clearer attempt to apply the image-schema theory to real cases of archetypal phenomenon, an explicit explanation for the role of image-schema (and how this occurs), an explanation of affect, and a fuller attempt to define the psychological purpose and significance of the archetypes. Instead, Knox seems to have discarded the theory of the archetypes in favour of pursuing theories of attachment in 2010 and suggested that further study of the archetypes would not be productive (Hogenson, 2019).

As outlined earlier, ‘usefulness’ should not be the sole criteria of whether or not knowledge is pursued. The phenomenon of the archetypes remains without full explanation, and while Knox made great steps towards solving this puzzle, there have since been advances in numerous fields (including in the work of Lakoff, whose work Knox originally drew upon) which could help inform the theory. The section below will therefore largely seek to elaborate on Knox’s original theory, with additions and challenges based on modern research.

A Revised Theory of the Archetypes

In terms of Jungian theory itself, I earlier stated how the theory posited by Knox (2004) based on the work of Lakoff and Johnson (1999), to me, failed to fully explain concepts such as the Hero archetype as this is not built purely on spatial relations. The theory also failed to explain the behaviour of the archetypes that was posited by Jung (e.g. why they occur in the places they do, what their function might be, how cultural deviations might occur, why the

images often feel that they are from an Other rather than within self), nor drew a clear dividing line between this phenomenon and other metaphor. This again brings us back to the issue of falsification outlined in earlier chapters.

One of the crucial points to be taken account of in the study of the archetypes (and transformations) is that while the transformations of the archetypes are a relative truth (arguably influenced by culture, early-development and personal experiences which shape the symbol), archetypes (defined as a reliable process rather than a thing in and of itself) are an absolute. That is not to claim that archetypes are a fixed object; on the contrary, it is the belief of the writer that this would be an erroneous reification. It is my belief that archetypes are a process, present in all people wherein embodied experiences are used to represent psychodrama through symbols (which repeat not due to any innate image store, but due to unchanging parts of the human experience, such as common physiology and experiences of spatial relations and force dynamics). Although the transformations may be 'infinitely variable', the underlying process is fixed and unchanging. Therefore, if we can fully understand the process that leads to the generation of a particular symbol (i.e. transformation), then we have understood the archetype and should be able to make precise predictions about future outcomes. It should therefore be the highest priority of Jungian researchers to examine the conditions that elicit similar transformations in order to find commonalities which may provide a glimpse at the underlying process (i.e. the archetypes).

Accordingly, what I plan to do in the following section is provide the reader with a breakdown of how the archetypes could work, using existing theories and approaches in contemporary psychology (particularly neuroaesthetics), and tying them to the original theory of Jung. This will include the elements of innateness which could plausibly exist, as well as developmental and environmental factors which shape the way an archetype manifests, as well as a brief suggestion of what function archetypes serve. This theory will draw on, but not be exclusive to, the theory posited by Knox (2004). I aim to present a comprehensive idea of what the archetypes could be, but most crucially, attempt to grasp at why particular transformations are generated.

It should be noted that, in the writer's opinion, there is some innatism involved in shaping the transformations of the archetypes. This could raise issues of falsification, particularly due

to the difficulties of determining nature versus nurture influence, but the only innate elements I am suggesting are those of physiology and basic emotions (which exist in or are dependent on subcortical areas and are thus resistant to environmental change). Basic emotions have also already been empirically studied and supported using selective electrical stimulations of neural circuitry in non-human animals.

My aim is to demonstrate how the archetypes could potentially be explained using modern research, avoiding mysticism and unfalsifiable concepts where possible, without reducing archetypes to simple biology divorced from the accompanying qualia. To follow my earlier criteria for good scientific research, I will aim to do following: provide alternate and null hypotheses; suggest a clear route for testing; define my terms as clearly as possible to allow for refutation and clear true/false conditions; and attempt to establish physical means by which suggested causal relationships could be possible.

To be explicit in the definition of my terms to aid falsifiability; I am defining the archetypes as: 'an innate process of environmental internalisation and metaphorical expression, which leads to the human tendency to repeat patterns in thought, observable in behaviour (including verbal behaviour) across time and culture'. There will be additional focus on transformations, which will be defined as: 'spontaneous images (as in visions, dreams and myths) which arise as a metaphorical representation of the core underlying pattern'. These patterns, e.g. that of the Hero, Mother, or Shadow are not genetically pre-determined, but arise due to common ecological and developmental experiences.

My intention is to write this elaboration up 'chronologically', that is, in the order of development I feel is plausible, beginning with that which could be innate at birth, moving onto the internalisation and elaboration of image schemas, and how these might be moderated by cultural factors. I believe it may well be possible to create a formulaic map of situations and link them to archetypal outcomes based on this theory, but that is beyond the scope of this paper. Throughout, I will also attempt to highlight which brain areas may be involved (in addition to those from metaphor processing which were outlined earlier) to help pave the way for future neurological research in this area, and also as an attempt to create a more falsifiable basis for the theory.

Innateness

I am beginning by considering the potential for innate elements of human psychology, which may support some of the claims for innateness in Jung's original theory. Knox (2010) intensely critiques Goodwyn's claim (2010) that Panksepp's (1998; 2011) concept of the 7 core emotions, can be used to explain similarities in symbolic content. However, it is exactly this work that I intend to draw on to counter this claim, using the very information that Knox outlines as their critique.

Panksepp's work belongs to a field of research investigating basic emotions. Basic emotions are those which are considered psychologically primitive, genetically determined, and are generated by subcortical neuronal circuits. Even if we do not subscribe to Panksepp's model, this much is agreed on across the four major emotion models put forward by Ekman and Codaro (2011); Izard (2011); Levenson (2011); and Panskeep and Watt (2011), (Tracy & Randles, 2011). It is also agreed upon that although environment can have a moderating effect on basic emotions (such as intensity and the conditions under which they arise), no environmental influence can create a basic emotion. This demonstrates the limited effect of external influences (and even higher order emotional systems) on these basic emotions. To give an idea of their universality, basic emotions are only considered so if they are also present in non-human animals (Tracy & Randles, 2011). This is highly reminiscent of Jung's suggestion that there are innate, primordial elements of the archetypes.

Although Panksepp's claim is one of a few, a brief exploration of this theory will be sufficient to demonstrate the point about basic emotions. Panksepp (1998; 2011) suggests that there are 7 core states in the mammalian brain; SEEKING, RAGE, FEAR, LUST, CARE, PANIC and PLAY. As noted by Tracy and Randles (2011), these basic emotions are also largely agreed upon across the four major models, with the only difference tending to be slight terminology changes, e.g. SEEKING as 'interest' in Ekman and Codaro (2011), or PLAY as 'enjoyment' in Levenson (2011). These have all previously been empirically supported using causal, cross-species studies, using electrical stimulation of discrete neurons (Tracy & Randles, 2011). These basic emotions moderate behaviour and are conducive to the continued survival of the animal, factoring largely into behaviour and motivation, particularly in early development. For these to be genetically passed, Ekman and Codaro (2011) argue that these emotions must have some specific behavioural output that aids the adaptation of the

species and enables them to better cope with ecological challenges. As Knox (2010) highlights, the mid-brain areas lose plasticity very early in development, and thus these primitive emotions become fixed. What Knox (2010) claims is that "these circuits are those that control behavioural and physiological patterns of action and reaction, not subjective mental experience." (Knox, 2010, p. 524). However, behaviour and physiology are arguably not separable from subjective experience. These emotions not only colour the nature of our early embodied experiences (the same embodied experiences which Knox (2004) seemed to be suggesting as underlying the archetypes) but dictate our willingness and likelihood to engage in them in the first place (Panksepp & Watt, 2011). Additionally, since these states are accompanied by physiological patterns and experiences, they also play a role in how we envision abstract concepts, as suggested by the work of Kövecses (2005) which will be further elaborated on below. This seems to suggest that there could indeed be basic emotions (i.e. innate subcortical circuits) which colour our conceptual representation and subjective experience of events. These subcortical systems are genetically determined, and work in tandem with higher order neocortical systems to produce emotional output, requiring very little environmental input (i.e. they are largely independent of time and culture). Notably, it is also highlighted that affective shift can occur in these deep subcortical systems without their being cognitive consciousness of the stimuli which causes the change (Tamietto and de Gelder, 2010, cited in Panksepp & Watt, 2011). This would suggest that these mammalian core states can be influenced unconsciously (which would be wholly necessary if they were to play a part in the archetypes) and suggests there may be a pre-determined element in the development of the image schemas, counter to Knox's (2004) claim.

It should be noted that it is superfluous here to suggest which of these basic emotion theories we use to explain early developmental behaviour, i.e. why a child engages in the early SEEKING behaviour that helps shape image schemas and later cognitive development. It is just enough to know that they *do* engage in this behaviour, and that whatever basic emotion underlies this behaviour is likely to be pre-existent, because it is something that emerges reliably in every typically functioning child. The only reason it is necessary to acknowledge these basic emotions is to indicate that there may be some innate disposition underpinning image schema development (and thus archetypes and their transformations) which is not environmentally dependent. This would account for the emergence of common feeling-tones (i.e. emotional qualities underlying transformations) independent of time and

culture.

This provides a plausible basis, at birth, for some of the innateness Jung refers to in his original theory, since these affective states are common to all humans. The basic emotions thus provide a basis for the universality of certain behaviours (arising as a response to basic emotions) and emotional states throughout life, particularly in times of crisis in early development (Tracy & Randles, 2011); in adulthood their influence is lessened by the pre-eminence of higher order emotions. Accordingly, if these states are involved in the affective similarities underlying archetypal experience, then mid-brain activity would be expected in archetypal experience.

Image Schemata

If it is accepted that the basic emotions may underlie similar affective states (which also leads to some predictable and reliable ways of interacting with the environment), it is important to consider the means through which the environment may be internalised and represented, i.e. by what means does external phenomena come to be internally represented in the form of image schemas?

When Knox (2004) discussed image schemas, she referred to UP-DOWN, CONTAINMENT, PART-WHOLE as examples, drawing on the metaphor work of Lakoff and Johnson (1999). There is a lot of discussion in Knox's (2004) paper of the image schemas and early conceptualisation and developmental changes but little discussion about the mechanisms underlying these conceptualisations, how precisely these support the archetypes, or what the limits are of these image schemas. Knox says of the archetypes:

> 'But what exactly are the processes whereby the abstract gestalts of primitive image schemas become elaborated into the rich psychic world of symbolic meaning and metaphor? Whilst image schemas can provide us with an information-processing model of the archetype-as-such, we also need to understand how day-to-day experience is internalized and linked with the image schema to create a pattern of representation-rich meanings… The archetype, as image schema, provides an initial scaffolding for this process, but the content is provided by real experience, particularly that of intense relationships with parents and other key attachment figures'. (Knox, 2004, p.10)

It is unclear to me why Knox (2004) places primary influence of the content of the archetypal transformations experiences with caregivers, particularly when many of the images (as outlined in the solar phallus example) are so far removed from day-to-day experience. If the content were concerning immediate relationships alone, there would have been no need to consider primordial influences; as Palmer (1997) states,

> 'For Jung… a true analysis of the individual can only begin when the contents of the personal unconscious have been exhausted and when the focus of attention can therefore shift to the collective unconscious – to those irreducible and primordial psychic images that exist in their own right'. (Palmer, 1997, p. 101)

As is indicated here, the very nature of the collective unconscious (which Jung conceptualised as the basis of the archetypes) is for those images that seem beyond *personal* experience. This explanation also does little to explain the numinous element of the experience associated with the archetypes and their often outlandish manifestations, and thus seems insufficient in its current state. Further detail is needed to account for the elaborate form and associated affect for each archetype, e.g. the Shadow, versus the Mother, versus the Hero all have distinct affective associations which are not fully explained by spatial relations alone (but which may all begin with the internalisation of spatial relations as a foundation).

It is also important to note that the research Knox (2004) drew upon (i.e. that of the metaphor theory of Lakoff and Johnson (1999)) has developed further since the time of her writing, particularly with the addition of force-dynamics, and a fuller explanation has been provided for the underlying mechanisms of the embodiment.

The reason that spatial information and embodiment are so crucial in the theory of the archetypes, is that we see these basic primitive image schemas shaping every language in the world (Lakoff, 2015). This is indicative of underlying universal patterns of thought that do not shift, largely akin to what was originally theorised to be the case for the archetypes, suggesting that we could be looking at a similar or the very same phenomenon. Spatial information and force dynamics (as illustrated in Table 1) are embodied experiences (leading to image schema) which help to characterise and shape perception and conceptual

representations in thought and language.

Table 1.

Spatial Relations and Force Dynamics (The Embodiment Hypothesis, Lakoff, 2015).

Spatial Relations	Force Dynamics
Deictic Center (viewpoint)	Force
Cardinal Directions (north, east, south, west)	Counterforce
Container (bounded region in space), interior: boundary: exterior; portal	Compulsion
Part-whole	Enablement
Center-periphery	Support
Axis (long, short)	Access
Front-back	Pull
Side	Push
Left, right	Resistance
Top-Bottom	Inertia
Adjacent Region	Momentum
Figure-Ground	Physical Containment
Source, path, goal	Blockage
Line	Blockage Removal
Verticality: Up-down	Restraint
Horizontality: side-to-side, back-and-forth	Restraint Removal
Region	Balance
Link (association, connection)	Stasis
Near-far-distant (scalar)	Thrust
Facing, Opposite	Friction, drag
Intermediate	Compression
Encircling, Curving, Tangential	Tension
Attached-Separated	Centrifugal force
Collected-Distributed	Centripedal force
Merging-Diverging	Entropy-Anentropy
Accessible-Blocked	Application (use of force)

Completed-Open-ended (path with and without endpoint)	Capacity (potential force)
Continuous, intermittent	Dissipation of energy
Length: short-long	Accumulation of energy
Instantaneous-extended	
Change in rate: acceleration-deceleration	

Having a direct understanding of the underlying neurological processes here is important for falsification, but also provides further information about how more complex archetypes could come to be.

Internalisation of Spatial Relations

In order to first establish a physical means through which this internalisation could be possible, (and to gain a more thorough theoretical understanding of the transformations), the first thing to be understood is how exactly external visual experiences and neurological processes are related. Through doing this, it will become more possible to understand the encoding of experience as visuo-spatial relationships. Some of the most fruitful research on this topic can be found in the field of neuroaesthetics.

Neuroaesthetics is a relatively new discipline in psychology which seeks to understand the biological basis of aesthetic experience (Chatterjee & Vartanian, 2014). Di Dio and Gallese (2009) conducted a meta-analysis of findings in this field and outlined a suggested role of embodiment (i.e. how the human brain and body interacts with the environment, particularly via sensory experiences and physiology), in aesthetic experiences. This is highly relevant given the suggested embodiment underlying the development of image-schemas.

In their meta-analysis, Di Dio and Gallese (2009) indicate sensorimotor activation during the viewing of artwork as evidence for the role of embodiment in aesthetic experience. In relation to the sensorimotor centers, aesthetic experience has been found to elicit activation in the parietal regions and posterior parietal cortex. These areas respectively play a role in spatial sense, navigation and the integration of sensory modalities, and the dorsal stream of vision. This indicates a process of visuo-spatial encoding, as well as motor mapping in aesthetic experience (Di Dio & Gallese, 2009), which supports the role of embodiment in

aesthetic experience (and thus of transformations, since they are, as images, inherently aesthetic). This is congruent with the image-schema model previously proposed by Knox (2004), since the very act of viewing begins a process of encoding spatial dimensions of an experience via the parietal and posterior parietal cortex. It is of note here that the posterior parietal cortex is also involved in the mapping of *conceptual* and *imagined* spatial relationships (Whitlock, 2017); i.e. the spatial relation could also be one representing power relations (high and low). This provides us with a potential means through which these spatial experiences can be converted into conceptual cognitive structures.

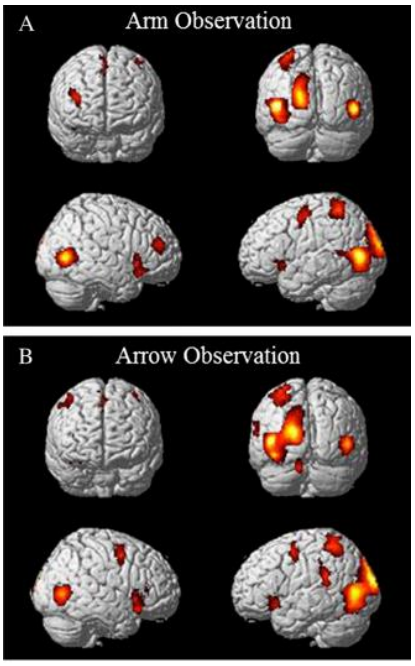

Figure 1: Biological versus non-biological effectors. (Di Dio, Cesare, Higuchi, Roberts, Vogt and Rizzolatti, 2013, p.433).

Another notable contribution from neuroaesthetics, which may help explain internalisation of external phenomena, is the mirror mechanism. The mirror mechanism is the automatic response of something resembling empathy and identification that exists within visual experiences; it essentially acts as a mental recreation of that which is physically observed. The first studies of the mirror mechanism were conducted on monkeys over 25 years ago, with findings showing a visuomotor response to watching another monkey perform an act such as 'grasping', a response which was also present when the monkey itself performed the action (Bonini, Maranesi, Livi, Fogassi and Rizzolatti, 2014). This empathic response within two members of the same species is not altogether surprising when one considers the importance of imitation in learning and adaptation, but what is more surprising is the observation of this visuomotor response when viewing an inanimate object in more recent studies. For example, Di Dio and Gallese (2009) discuss the work of Di Dio, Macaluso, and Rizzolatti (2007), a study in which participants viewed classical sculptures and displayed activation of the ventral premotor cortex (area F5, associated with mirror neurons) and the posterior parietal cortex (which is responsible for both sensory integration and motor planning, and which was also discussed earlier as being involved in imagined and concrete spatial relations.) That is to say, even though the viewers were only viewing a still sculpture, the subconscious response is to then imagine and mirror the implied movement of the artist. As Di Dio and Gallese (2009) put it, this suggests that there is a 'relationship between embodied simulation-driven empathic

feelings in the observer and the visible traces of the artist's creative gestures'. Much of the previous research into mirror neurons has been concerned with identifying their role in social interaction (specifically during observation and execution of an action) (Kilner and Lemon, 2013) but the presence of mirror neurons when viewing inanimate objects has fascinating implications for what it is possible to internalise, and the breadth of possible sources for spatial relations (and the concepts which arise from them). Further studies have also shown that the mirror mechanism activates regardless of if the effector is biological or non-biological (e.g. a hand versus an arrow) provided that the motion itself is biological, as is illustrated in Figure 1 (Di Dio, Cesare, Higuchi, Roberts, Vogt and Rizzolatti, 2013).

This suggests that viewing others, viewing inanimate objects which have been crafted, and viewing non-biological actors may all provide a basis for embodied experiences and thus spatial relations which abstract reasoning may be built upon. However, it is important to note that the mirroring is not present when the motion is non-biological, which may make a dividing line between what it is and what it is not possible to draw a spatial map from simply from viewing.

While this is speculative, it is interesting to consider how this may apply to certain viewed experiences, e.g. the growth of a plant, or the motion of a twig drifting downstream, or even the viewing of a heavily industrialised city. To what extent do we internalise and create spatial maps of the world around us, and how does this contribute to our available image-schemas? It is important not to overstate this speculation, as the possible functions of mirror neurons in themselves have already been vastly overstated in the last decade, but for a theory like the archetypes and transformations which centers so heavily around embodiment and identification, it does seem fruitful to further research the limits of this mechanism; without knowing the full range of what is possible (and what is not), it remains difficult to map and test what can fall within the parameters. However, it does seem unlikely, based on this and the other routes for internalisation outlined above, that image-schemas are rooted purely in early development.

More detailed neurological accounts of internalisation of image schema can also be found in neural dynamic models of spatial relation, such as that by Richter, Lins and Schöner (2017). This model uses dynamic field theory (DFT), which suggests that activation fields

have metric dimensions, with objects of interest continually updated. In DFT, neurons respond to external stimuli and can have inhibitory effects on local neurons when making decisions about key features, e.g. in activation of neurons for a colour field, the presence of red stimuli may have a stimulating effect on a 'red node' and inhibitory effect on other local colour nodes. Peaks in activity are thus processed as indicative of the presence of one item and the non-presence of another. This simulation-based model allows for amodal perception, and also suggests the involvement of the parietal cortex. Here, the parietal cortex is considered as the centre for encoding bimodal signals, i.e. of features and spatial relations, from the ventral stream (Wiebe, Morton, Buss, & Spencer, 2014). This gives a brief idea of some of the neural processes which may potentially underlie the conversion of external stimuli into conceptual structures.

The work of Piaget and Inhelder's (1969) suggested that cognitive development in children begins with sensorimotor experience. Physical experiences (internalised through processes as outlined above) become the first cognitive steps, and these experiences are then used to metaphorically map abstract thoughts and situations. This indicates the potential for a continued relationship between embodiment and abstract thought in later life; by necessity, those things which cannot be understood concretely are metaphorically substituted (as outlined in the earlier discussion of conceptual metaphor theory). This enables them to be understood through analogy, i.e. conceptual representations which, at core, are based on spatial relations (possibly due to the 'conceptual' and 'literal' flexibility of the posterior parietal cortex) which have been encoded through the processes above.

In fact, problem-solving through analogy is an everyday phenomenon, with people drawing frequently on embodiment to help resolve knowledge gaps. For example, when explaining complicated or disconnected ideas, there is an increased use of bodily gesture, which is thought to support problem-solving by facilitating working memory (Keefer & Landau, 2016). Metaphorical structure-mapping, wherein an unrelated concept is mapped onto an abstract concept to aid perceptual representation (as earlier discussed in conceptual metaphor theory), is also common, though it is less automatic and can require active thought or linguistic cues to trigger relevant structural maps (Keefer & Landau, 2016). It is this that I feel leads to the emergence of the katabasis mythology only after the dawn of agriculture; humans had to physically experience seeing something go down into the ground (a seed), and

then being ‘reborn’ (as a plant) before being able to conceptualise the phenomenon within themselves; before agriculture, there was no physical experience to act as conceptual scaffolding for this thought.

Although this is speculative, it therefore seems plausible that archetypal experience arises as a kind of subconscious conceptual metaphor, particularly aimed at resolving subconscious problems involving high levels of affect. Transformations (as the source) may thus emerge as part of a class-inclusion process, being subconsciously selected as a good structural match with the target (this being the problem the individual is facing), due to image-schemas grounded in early sensorimotor experience, or due to compounding associations and affect. This class-inclusion process, in turn, may then shape our attitudes towards, and relationship with, the target and source objects due to their various shared affective and experiential associations. For this to be passed evolutionarily, similar to the basic emotions, it therefore follows that this must either have some adaptive benefit, or else may simply be an evolutionary by-product of our problem-solving faculties which developed to solve material problems.

Spatial Relations and Affect

For image schemas to form the basis for transformations (as a highly affective experience), there must also be a relation between spatial relations and the affect that they are used to represent. It is therefore useful to examine the relationship between aesthetic and emotional experience, to see at which point these two experiences connect. This, in turn, will help provide an idea of which brain areas would be involved in an experience of archetypal transformations, since these are also inherently a visual experience.

To draw again on neuroaesthetics, in the analysis conducted by Di Dio and Gallese (2009), aesthetic preference was shown to be associated with the right anterior insular and right amygdala (Di Dio, et al. 2007). The former here is responsible for emotional awareness (Gu, Hof, Friston,& Fan, 2013) and was also found to be active during figurative metaphor work (Schmidt and Seger, 2009), while the latter is responsible for emotional processing (Janak & Tye, 2015). Di Dio and Gallese (2009) suggest this is indicative of a comparison being made between the aesthetic experience and the observer’s own experiences, which could provide an

avenue for the interplay between spatial relations and affect.

This said, amygdala activity was only found when participants were asked to make aesthetic judgements (Dio Dio, Macaluso and Rizzolati, 2007) which calls into question whether or not activity would be shown in this area when experiencing archetypal transformations. This is because it suggests that the viewer's emotional processing may be a result of an evaluative process of personal associations to determine the beauty of the piece, rather than as an instinctive response to aesthetic experience. However, the transformations are highly affective experiences, and known for the visceral impression they supposedly impress upon the psyche (Jung describes them as having a "mystical aura about [their] numinosity, and it has a corresponding effect upon the emotions" 1947, pp. 205-206). With this powerful affective response and knowing that transformations are a moment of unconscious emotion being made conscious (i.e. being brought into emotional awareness, as is one of the suggested functions of the right anterior insular), it is difficult to imagine that the anterior insular and amygdala would not be involved.

It is interesting to note that these brain areas have also been found to be associated with emotional processing in sleep, which is what we would have expected to find if dreams did indeed serve the compensatory function that Jung (1964) suggested. For example, amygdala activity increases during rapid-eye-movement (REM) sleep (Koyama, Aoyagi and Toyomaki, 2011), which is when dreaming occurs. This amygdala activity is also thought to be responsible for attenuating responses to stress during REM (Wellman, Fitzpatrick, Hallum, Sutton, Williams and Sanford, 2017) and further findings have also found a relationships between REM and mood disorders, e.g. in bipolar it was found that dreams of death are common and precede an upward mood shift (Beauchemin & Hays, 1996). Alternate research has also found that the anterior insular may be involved in sleep-regulation (Chiang, Yugay, Patxot, Özçivit, Hu and Lu, 2016), and also that sleep deprivation (and resultant increased activity in the amygdala-anterior insular pathways) leads to higher rates of post-traumatic stress disorder (PTSD) (Schott, Germaine and Milad, 2015). These studies therefore indicate the importance of REM (and thus dreaming) in normal emotional processing. This arguably suggests that amygdala activity (and in fact dream content) during REM has a direct influence on mood and emotional processing in waking hours, which is also what would be expected for an experience of transformations manifesting as a way to consolidate emotions

during sleep. This is further supported by Malinowski and Horton's (2015) work suggesting that metaphor is used during dream-states to assimilate emotional memories. Combined with the above research from neuroaesthetics, if we do take archetypal transformations to be a kind of metaphor processing it thus seems increasingly probable that archetypal experiences would elicit activation in the amygdala and the right anterior insular.

An area in which we might expect transformation experience to differ to typical aesthetic experience is a lesser role of the cognitive faculties since Jung suggested archetypal experience is most likely to occur when conscious defences are relaxed. This is an important deviation from standard aesthetic experience, as the primacy of cognition in our responses to art has previously been considered as the norm (Di Dio and Gallese, 2009). By contrast, transformations would be subject to less influence from the cognitive faculties because they are not conscious experiences. Mills (2012) elaborates on this, suggesting that the archetypes and collective unconscious, like drives, are not faculties of cognition but are rather a pre-conscious process, from which the transformations arise as highly affective and numinous phenomenon. In terms of falsification, this differentiation may provide an important neurological differentiation between symbolic experience and typical aesthetic experience or other forms of metaphor, reflecting the significant difference in qualia, i.e. that the archetypes often feel that they come from outside the self. It also follows that if this is the case, there should be a distinctly different activity in the right anterior insular before and after exposure to the right anterior insular, due to its role in emotional awareness.

From Spatial Maps to Transformations

The findings in neuroaesthetics give us something of an idea of which neural correlates may underly the experience of a transformation and how they come to be; during aesthetic experience, we are not simply viewing, but also creating corresponding spatial maps which may also call on our personal experiences and affective associations with this map. For example, in viewing a person stood on a cliff, a person who fell down the stairs in early childhood, or from a high wall, may call upon that spatial relationship in tandem with personal experience, leading to an association of that mapping with fear. That is to say, visual experiences and representations are often inherently connected to physically embodied motor and affective responses. For instance, Meier, Hauser, Robinson, Friesen, and Schjeldahl (2005), found that a brightness and verticality are often used to represent abstract ideas

without an explicit physical basis, such as ‘valence’ or ‘power’.

Meier et al. (2005) also suggest persistent associations between various physical states can influence later affective attributions. They suggest that daylight hours are associated with receiving various rewards such as food or social interaction, while at night, we are less evolutionary equipped to deal with environmental dangers or predators. As a result of this, bright perceptual input becomes entwined with positive affect, and dark perceptual input becomes entwined with negative affect. This is a lived truth that arises predictably as a result of our evolutionary history and biology, and again suggests an element of innatism and physiology that shapes our affective response to a visual experience. This is not just true for light levels and distance:

> ‘valence can be represented in terms of brightness, and that divinity is represented on a vertical dimension. There is also evidence that the concept of similarity is represented in terms of verticality, physical force and size. Furthermore, studies have shown that importance is represented in terms of weight, and that social proximity is represented as temperature.’ - Zanolie, Dantzig, Boot, Wijnen, Schubert, Giessner and Pecher (p.57, 2012)

In terms of how this may occur, Lakoff (2015) (in discussing the embodiment hypothesis) suggests that the root of these connections is spike-time dependent plasticity. This is when events that happen temporally close together become bound together as part of a scenario or more technically speaking, when a presynaptic cell fires and a post-synaptic cell fires shortly afterwards, the synapse between them is strengthened. This allows us to develop scenarios and concepts which are moderated by coordinating circuits. Lakoff (2015) mentions this in relation to integration and neural bindings in metaphor (e.g. in early childhood, if you are loved, you are often held close, which leads to warmth, hence ‘love is warmth’) and it is just as relevant here for understanding the metaphorical way in which archetypes are presented.

Because the scenarios and associations that we develop are implicit and can often not be consciously checked, it is possible that they can become over-generalised to content with similar sensory encoding, even if the object itself is not the same. For example, someone with a mother who treats them distantly or coldly may end up with a scenario in which the mother

archetype becomes a collection of the following: [SEEKING], [container], [verticality: up], [distant], [resistance], coupled with the resultant basic affect, e.g. rage or panic. These things may bind together to become a scenario, so in later life, when a similar scenario is evoked, e.g. for a woman of interest who is aloof, although not conscious, the same neural bindings fire and lead to the same affect being generated.

When this emotional conflict happens below the level of consciousness, the result may be spontaneous images, which are a structural match because they utilise the same spatial relations and force dynamics which are relevant for the scenario. This may well be the basis of what we would term 'the Mother archetype', i.e. a common collection of associations that arises due to most people having or experiencing a mother. Lakoff (2015) discusses coordinating circuits, which is the idea that we have neural binding circuits for these scenarios, which are gated – over-generalisation then, may arise from stimuli which modulates the gate and activate these circuits provided enough of the underlying sensory information is consistent.

Internal experience having an interconnected relationship with the external environment is not a new idea only found in the field of psychology, but also appears in fine art, or in literature in the form of pathetic fallacy. Although neuroaesthetics is a relatively new field, this is a phenomenon that has been within human awareness for quite some time, and arguably is part of why it comes as no surprise to us when we read sentences such as 'the clouds were sullen'. This statement is not taken literally, nor seen just as text, but instead gives us an image of the cloud with accordant visual properties to those we associate with 'sullen'. Here, the theory again links back into metaphor; when saying "sullen", it is unlikely that those reading envisioned the clouds as bright and white, nor envisioned that they were moving quickly, for example. It is far more likely that instead, the reader envisioned them as a darker colour, such as a dull grey, moving slowly by if moving at all. This is particularly important to note, as here, we begin to see how although visual properties were not disclosed, they were inferred and encoded in relation to their emotional content.

This indicates that abstract properties are not just assigned to metaphors that we project, but that spatial qualities and physical features such as colour are also encoded for objects even when they were not disclosed or present, based on their affective association. Our ability to

metaphorically group items together seems to be based on these hidden ‘nodes’ of information.

This leads us on to synaesthetic correspondence. Although the above provides a theoretical link between visual experience and affect, another question that arises in explaining the archetypes is why particular substitutions are made, i.e. ‘Why does the Transformation appear like this?’. Why the earth as the Mother? Why the sun and moon as siblings, or husband and wife? It is possible to understand this through synaesthetic correspondence. This is important to understand in order to further falsify the theory and delimit the range of transformations from ‘infinite’ to ‘those which are a match for associated sensory experience’.

Synaesthetic correspondence is an aspect of neuroaesthetics which has recently come to light, showing that unrelated information is often linked by similar sensory encoding. For example, recent work with taste-shape matches amongst non-synaesthetes show that rounded shapes which were also symmetrical with few elements were likely to be judged as sweet, while asymmetrical, angular shapes with many elements were likely to be branded as bitter and unpleasant (Salgado-Montejo, Alvarado, Velasco, Salgado, Hasse, and Spence, 2015). Symmetry corresponding with positive valence is not an entirely surprising finding, as this correlates with human ideas of beauty and harkens back to evolutionary fitness and sexual selection; significant facial asymmetry is often found in chromosomal disorders (Hoyme, 1994, cited in Rhodes, Zebrowitz, Clark, Kalick, Hightower, McKay, 2001). There are also long-standing ideas that beauty has a connection to truth and virtue (Bommel, 2015) which arise from our more recent cultural and religious history (i.e. concepts of God given beauty) which in itself is an association which may arise from synaesthetic correspondence.

Synaesthetic correspondences provide a lens for understanding why symbols come in the forms they do, and also provide a contribution for understanding why there are commonalities in representations. For example, in the study above, both evolutionary and environmental factors could lead to the direct connecting of two pieces of seemingly unrelated information; symmetry and beauty. There are countless other experiences which may also be routinely connected (for example, in the study above, we see ‘few elements’ and ‘sweet’ which more generally may be ‘simple’ and ‘sweet’ or ‘angular’, perhaps akin to sharp, as ‘bitter’), stemming either from evolutionary or environmental factors. This may

eventually be mapped in a more comprehensive way than has been established at present, but these relationships make visible some of the underlying sensory and conceptual connections that exist as a result of the embodied aspect of experiences.

For example, take the sun and moon mythology (which appear frequently as transformations for the God or Mother archetype) which is common across many cultures. Many things exist above us; clouds, birds, the stars – but we seem most compelled to write mythology about the sun and moon, particularly framing them as brother and sister, husband and wife, or both. Why could this be? Perhaps because they correspond with a collection of abstract qualities that arise from early image-schemas. They have verticality, they are distant, and our sensory experience of them is as bright, and they are large (and thus also draw a great deal of visual attention), with the varying heat of the sun (which could be interpreted as nourishing or punishing) being climate dependent. According to the studies outlined above, this psychologically corresponds to valence, divinity and power (Zanolie et al. 2012).

Supporting the implications of this research, the mythological pattern also follows this trend; the Japanese Tsukiyomi (moon god) and Amaterasu Omikami (sun gooddess), who is both Tsukiyomi's sister and wife; the Norse, Sunna (sun goddess) and Mani (Moon God) who is her brother; in Greek, Selene (moon goddess) and Helios (sun god) are siblings, with the addition of Eos (dawn Goddess).

It is of interest to note that Egypt is something of an exception here: Egyptian Ra (the Sun God) was actually married to (or the child of) Hathor who epitomised Egyptian femininity and was considered the 'mother'. Although Hathor occupied the typical function of other moon goddesses, she also had a dual aspect of being Ra's feminine counterpart or a maternal figure, one considered wrathful, and one gentle (which is reminiscent of earlier discussions of the Mother archetype and the Earth mother). Hathor was also typically represented by cattle (which was a symbol of nourishment in Egypt and many other cultures) and it is actually Khonsu (whose name interestingly means 'traveller' and was also referred to as 'pathfinder', which links us back to the ‘path’ image-schema) who acted as the moon God.

Understanding the basis for these deviances is exactly the place where we may find the

precise workings of the archetype; what commonalities is it possible to find, and what causes variance in transformations? When Herodotos (an ancient Greek historian) visited Egypt in the 7th century, he joked that their culture was "in most of their manners and customs, [the Egyptians are] exactly reverse the common practise of mankind" because of their climate, the patterns of the rivers, and the more active position Egyptian women possessed in society (Erman, 1971, p. 1). Could this be in some way responsible for the variation in their transformations? At present it is impossible to say, but this in itself provides one potential avenue for study for those preferring a general systematic approach to the collection of knowledge.

To return to the point of the metaphorical connection between the sun, moon, and parental figures. It is not entirely surprising that one should be attributed masculine and the other feminine; the typical experience of a child throughout the ages has been to grow up with two primary figures above, one masculine, one feminine, both of whom are life-giving and protective forces and much larger than the child. When the developing person is then presented with two unknown objects above (which also correspond with dual figures, with the underlying synaesthetic correspondence of power, valence and divinity), they must then ask 'what is it?'. It is unsurprising then that the person concludes that they must share similarities to their parents as they share a multitude of abstract qualities. The person thus categorises them together, and the sun and moon become a God and Goddess, who are either brother and sister or husband and wife (and sometimes both)**.** They could not be human as a matter of relativity – they are higher, larger, brighter, they must be something *more* powerful than a human, *more* life-giving, *more* divine. This was not merely a weak explanation for something that the generations of the past did not understand, it was a psychological truth about the function of the sun and moon as life-giving forces.

The sun is indeed life-giving; without it, plants would not grow, the cattle could not eat, and humans could not navigate. For a human who reasons through function, this is what the sun is 'for'. In modern secular society, we instead suppose that everything is for nothing, there is no 'purpose', per se, to anything and the causality is in fact reversed; we evolved to adapt to our surroundings, what grew is what *could* grow under these conditions. And this is true also – this is the material truth, logos, but it does not stop the sun having a powerful emotional impression, nor from being life-giving, or from inspiring a sense of the sublime when we

consider, for example, how far it is from us, or how small we are by comparison. The two are not in conflict with each other except when one side insists that the other must submit to it; the lens of the emotional truth is not appropriate for understanding the material history of the sun, and equally, the lens of material truth is not appropriate for understanding the emotional relevance of the sun. Material history and human experience run concurrently, they are interlinked and dynamic – but they are not the same and it is a mistake to view them as so. This brings us back to the earlier discussion of the subjective and objective experience that Humboldt, Goethe and Jung tried to capture.

Cultural Influence on Transformations

As stated above, it is highly plausible that as well as developmental experiences, culture may also have a mediating effect on transformations. This can even take primacy over the physiology as a source domain, as is explored in the work of Kövecses (2005).

Kövecses discusses the physiological basis for universal metaphor and states that general metaphors are often universal. For example, in discussing anger, 'angry person as a pressurised container' is near universal (Kövecses 2000, cited in Kövecses 2005). However, as metaphors become more specific, cross-cultural influences begin to emerge. For example, anger in the stomach is specific to Japan; in Zulu culture it is in the heart; in China, the anger that fills the container is an excess energy called *qi* (gas, a comparison which stems from Chinese philosophy and medicine); and in English metaphor, fluid.

This indicates that while a general metaphor is universal, the higher levels of elaboration are likely to be subject to cultural influence. Due to the neural plasticity needed for this kind of environmental influence, it seems to be the case that subcortical areas would have to work in tandem with 'higher' brain areas to develop metaphors on the specific level. It is difficult to tell at this stage whether this would still be classed as a transformation, as due to the cultural influence, these transformations may be more conscious, and resultantly not as driven by basic emotions and unconscious influences so much as higher order affect.

In relation to embodied experiences, 'anger as heat' also appears across many cultures, and appears to be embedded in the underlying physiological reaction caused by anger; the skin grows hot and there is an increase in blood pressure (Kövecses, 2005), which contributes to

the frequent (and often subconscious) choice to use heat as a metaphorical representation for anger.

This, in part, would lend further evidence to Jung's argument for innateness, as it is true that our physiology is unchanging over our (relatively) short history, and as mentioned earlier, it could well be that our physiology is part of the 'innate base', akin to the 'biological instincts' that comes to colour our symbolic representation. This would be congruent with what Jung suggested towards the end of his career and in his posthumously published work.

> '[The collective unconscious] can no more be a product without history than is the body in which it exists. By "history" I do not mean fact that the mind builds itself up by conscious reference to the past through language and other cultural traditions. I'm referring to the biological, prehistoric, and unconscious development of the mind in archaic man, whose psyche was still close that of the animal.' (Jung, 1964, p.57)

However, as Kövecses (2005) highlights, the picture is not quite so simple; 'heat' has risen and fallen over time with its popular use as a conceptual metaphor for anger. For example, before the year 850, heat-related words for anger in English only stood at 1.59%, sharply rising in the years 850-950 and then decreasing again to 6.22% between the years 950-1050 (Gevaert, 2001, cited in Kövecses. 2005). This suggests the presence of other mediating variables which actually take precedence over physiology. This may again indicate the role of cultural influences at this level, since there is variation emerging over time.

Possible Theory Testing

The aim of this reformulation was to make the theory of the Jungian archetypes more falsifiable and more amenable to scientific enquiry. For the sake of illustration purposes, and in order to demonstrate that this has been achieved, I will now outline an example of the sort of simple and experiment which could be used to test this theory.

An initial test may be to investigate whether there really is an association between transformations and simple image schema. This experiment may be as simple as conducting as a repeated measures study, with one IV (reflective) at two levels; sensory-correspondent or

not sensory-correspondent, and the DV as how often the image is chosen. Participants could be presented with the word for an archetypal pattern, e.g. ‘Hero’, followed by two images of visuo-spatial orientations (or simple animations to better capture force dynamics or transitional image schema, but each using simple black circles, except when shape or colour is relevant to the visual input being tested), one reflective of the underlying sensory encoding and the other not (e.g. a light circle and a dark circle to correspond to valence) in accordance with the work outlined above. The participant will be asked which image or animation most accurately represents the word. It is anticipated that the participant will choose the media with sensory-correspondence more often than the one without sensory-correspondence.

Research Hypothesis: Participants will choose the media with sensory-correspondence with significantly higher frequency than the image which does not have sensory correspondence.

Null Hypothesis: Whether the image has sensory correspondence with the concept or not will have no influence on how often the image is selected.

If the research hypothesis was supported, this could be continued across various sensory inputs and positioning in order to create a physical representation of the archetype made from basic shapes and spatial configurations. For example, testing could be done to see if verticality is significant in the image-schema underlying the ‘hero’ concept, or if path would be. Through this, we could essentially produce an ‘image’ of every claimed Jungian archetype (by gradually combining those sensory impressions which are relevant).

Once a physical representation of the archetype was established, the study could be reversed, with the physical representation being presented first, placing and asking participants to then select between ‘hero’ and ‘trickster’ for example. It is difficult to generate a hypothesis for this without the prior work being done on each variable in isolation, but for example:

Research Hypothesis: Participants will select the word ‘hero’ more often than ‘trickster’ when the image presented first contains the elements of verticality, brightness and path from down to up rather than up to down.

Null Hypothesis: The elements of verticality, brightness and path from down to up will have no effect on which image is selected.

There would be a lot of testing to do in order to even reach this stage. However, this once again opens the door for normal science for the archetypes, as outlined in Kuhn's (1970) work. This would highlight which sensory influences are most important for each archetype, and, following Lakoff's (2015) work on coordinating circuits, we would be able to see how many sensory correspondences are needed in order to modulate the circuit gates for a given archetype and trigger activation. If successful, this may allow us to discern what can and what cannot be an archetypal transformation, as we would know how many corresponding synaesthetic elements would be needed to activate the circuit (i.e. establishing an activation threshold). For example, there may be no difference in selection when attempting to represent Hero simply through brightness; it may need to be brightness and verticality; or brightness, verticality and something to stimulate a particular affective response, in order to modulate the circuit gates, leading to a significant difference in selection rate. Accordingly, we would then be able to say that anything that did not include those features did not represent that particular archetype, because it would not meet the requirements to activate the circuit. If viable, this would allow us to limit the "infinite variety" (Jung, 1954a/1975, p.81, *CW* 9 pt.1, para 156) initially suggested for the archetypes.

Additionally, by presenting the image first, the selection process should be mostly intuitive and based on subconscious inclinations. When switching to the word following the image, this implicit process may be made conscious in a way that ought to be akin to archetypal transformations. In aid of falsification for researchers with access to an fMRI, one might therefore predict that this will elicit activity in the parietal regions, the posterior parietal cortex, the right anterior insular, the amygdala, the angular gyrus, the left-inferior frontal gyrus and potentially the ventral premotor cortex when the emotional scenario given involves implied movement. How this differs from a metaphor is that the process is largely unconscious and acts as a subconscious learning tool, with higher affect involved. This should lead to differing activity in the right anterior insular, and higher activity in the amygdala compared to a standard metaphor.

Conclusions

Early sensorimotor experiences, shaped by reliable basic emotions, environmental factors, and common associations between embodied states (moderated by cultural differences) all come together to shape the way we encode experience. Through conceptual metaphor theory and the processes of internalisation, we can begin to see a subconscious and *continuous* process of identification with our environment, utilising metaphor to bridge the gaps in our knowledge. The work by Kövecses (2005) indicates that even our most general metaphors can have multiple groundings, from our physiological responses to our environment, and in the work by Meier and Robinson (2005) we can see how continued associative links may plausibly be responsible for how we encode affect. Accordingly, here I will differ from Knox (2004) and Merchant (2009) in considering the archetypes as particularly a product of our early developmental experiences; as long as new embodied experiences continue to occur (especially drastic ones, now often facilitated by technology), then it follows that the potential archetypal transformations (and thus conceptual metaphorical structures) would expand as our capacity to identify with our environment expands.

I have posited that the experience of archetypal transformations is a method of subconscious expression and potentially a form of problem-solving. As discussed earlier, the transformations arise as a result of basic emotions, which have evolutionarily ensured our survival as a species. As a result, they emerge with visceral affect and are often associated with a sense of urgency. The transformations are those conceptual metaphors which arise subconsciously, perhaps when these subcortical structures (responsible for basic emotions) are activated, but the problem (i.e. knowledge of that which is subconsciously perceived threat to survival or health) is pre-conscious. This is then symbolically represented (with symbols chosen according to synesthetic correspondence with the target item) when conscious defences are lowered (as in dreams, fantasy or mythology). The process through which this occurs is through the utilisation of metaphorical substitutions which arise from embodied experiences and physiology, and at higher levels of specificity, culture. The reason for the primacy of embodied experiences and physiology is because cognitive development in children begins with sensorimotor experiences (Piaget & Inhelder, 1969). These concrete experiences are then used to metaphorically map non-concrete concepts. Appropriate metaphors are chosen in accordance with a process akin to synaesthetic correspondence,

which is when unrelated information is connected by its underlying sensory encoding. For example, I discussed earlier that the Wise Old Man archetype need not be a philosopher, but is instead any figure which serves to provide guidance and sound judgement. What this can be further broken down into is a series of sensorimotor commonalities and associated affect, arising from a combination of physiology and developmental experiences. For example, in this case, Wise Old Man could be embodied as The Self experiencing primitive 'SEEKING' behaviour and interacting with a 'higher than self' spatial relationship combined with 'CARE'. Commonalities in the metaphors used to represent these affective and spatial relations are explained by a combination of physiology, the lack of plasticity in the mammalian brain (which both account for elements of what one might refer to as innatism), as well as universal commonalities in embodied experiences which arise from environmental constants. Examples of these include the laws of physics and revolutionary societal changes (such as the shift from hunter-gatherer to agriculture), which lead to similar sources for the subconscious representation of psychodrama. Deviations in the metaphors can occur at the specific level and these changes are often drawn from long-standing cultural trends, as suggested by the work of Kövecses (2005). The purpose of these transformations (and the archetypal process underlying them) is to prompt reflection or moderate behaviour in order to ensure health or survival. This separates them from other metaphors, as their associated affect is much stronger because of their evolutionary function to ensure the immediate survival of the organism, as opposed to the domestic reasoning associated with other conceptual metaphors.

This is similar to Knox's (2004) theory of the archetypes but differs in that I see no reason why the process of acquiring embodied models would be primarily shaped by early relationships with caregivers. As long as new embodied experiences occur and new affective associations arise, the capacity to draw new analogies will also expand. This is due to a continued process of identification and internalisation of our external environments. Physical experiences thereby provide new pathways for problem solving; myths and dreams are not filling gaps in logos, they are ways of discussing and resolving emotional truths as we, as conscious and reflective beings, navigate a psychological landscape coloured by (sometimes maladaptive) emotions which once acted as a guide to ensure our survival. Archetypes and transformations are not an accident but an essential part of our lives as humans living in a

society, acting as a way of representing and reflecting on the universal patterns and scenarios that make up our existence as human beings.

There are various other issues which are not touched on here. For example, I do not discuss Jung's theory of synchronicity. This theory concerned 'meaningful coincidences' and was thought in some way to provide explanation for the archetypes. The reason for not touching on this is that it would be easy for this to become a discussion of physics, evolution and entropy. Synchronicity is also well-known for having many falsification issues of its own, which would need re-evaluation in light of modern evidence. In a paper which already touches on a wide range of topics, I feel it would be unwise to go further afield and include this here, although seeing how it might work in conjunction in synchronicity may be something to consider for future research.

Another construct I do not particularly mention by name is the collective unconscious. In the original theory, Jung suggests the collective unconscious as a kind of store for predetermined contents and behaviours, with the archetype as templates as fixed types which would emerge (e.g. Trickster, Shadow, etc), and their manifestations as the observable phenomenon (which I have referred to here as the transformations). The reason for not discussing this explicitly is that there is no apparent need for the collective unconscious as separate from the archetypal process, if one takes the view that genetically fixed brain structures and common environmental factors are responsible for the commonality in the transformations. In the theory outlined above, I feel everything necessary to create the observable phenomenon is accounted for by a combination of universal environmental and innate factors. Perhaps these, combined, could be considered a kind of 'collective unconscious' – a universally shared set of spatial relations, force dynamics and basic sensory associations, which the transformations arise from. However, I do not see any need for this term, especially since it is suggestive of a kind of archaic collective pool of knowledge that I do not see evidence (or need) for in explaining the archetypes and their transformations.

Finally, I have not much mentioned individuation here. Individuation is the Jungian telos achieving mental balance, where neither conscious or unconscious content are causing undue stress on the other. Although not discussed here, I feel above all that individuation may be an awareness of the processes above and could therefore be reframed as a metacognitive process

whereby individuals become aware of these implicit learning processes. Perhaps similar to the phenomenon of mindfulness, individuation may be a process wherein individuals become able to consciously shift executive function to be aware of themselves, using transformations as tools for reflection and establishing personal well-being. However, again, how exactly this fits in with the theory above may be a matter for further research.

I recognise the failure to fully explore these as a limitation of this work, but my modest hope is that what I have stated is simply a step in the right direction. The above work is an elaboration on the theory put forward by Knox (2004), with the inclusion of comparisons with neuroaesthetics (and suggestions about possible shared neural-correlates), ideas about the potential role of synesthetic correspondence, further detailing on metaphor theory and embodiment based on the more recent work of Lakoff (2015), suggestions of how archetypal transformations could arise, and a proposal for scientific enquiry to begin narrowing the range of outcomes which could be attributed to an archetype. I thereby hope to have aided falsification and furthered the grounds of scientific enquiry for this theory.

Bibliography

Armstrong, K. (2005). A Short History of Myth. Edinburgh: Canongate.

Beauchemin, K. M., & Hays, P. (1996). Dreaming away depression: The role of REM sleep and dreaming in affective disorders. *Journal of Affective Disorders, 41*(2), 125-133. doi: 10.1016/s0165-0327(96)00080-8

Beck, R. (2000). Ritual, Myth, Doctrine, and Initiation in the Mysteries of Mithras: New Evidence from a Cult Vessel*. *The Journal of Roman Studies, 90,* 145-180. doi: 10.2307/300205

Beeman, M.J. & Bowden, E.M. (2000). The right hemisphere maintains solution-related activation for yet-to-be-solved problems. *Memory & Cognition, 28*(7), 1231-1241. doi: 10.3758/BF03211823.

Binder, J., Desai, R., Graves, W., Conant, L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19 (12) 2767–2796.

Bonini, L., Maranesi, M., Livi, A., Fogassi, L., & Rizzolatti, G. (2014). Space-Dependent Representation of Objects and Other's Action in Monkey Ventral Premotor Grasping Neurons. *Journal of Neuroscience, 34*(11), 4108-4119. doi: 10.1523/JNEUROSCI.4187-13.2014

Bonner, M.F., Peele, J.E., Cook, P.A., & Grossman, M. (2013). Heteromodal conceptual processing in the angular gyrus. *Neuroimage, 1* (71), 175-186. doi: 10.1016/j.neuroimage.2013.01.006.

Bommel, B. v. (2015). *Between "Bildung" and "Wissenschaft": The 19th-Century German Ideal of Scientific Education*. Retrieved from http://ieg-ego.eu/en/threads/models-and-stereotypes/germanophilia-and-germanophobia/bas-van-bommel-between-bildung-and-wissenschaft-the-19th-century-german-ideal-of-scientific-education

Bornstein, R. F. (2001). The impending death of psychoanalysis. *Psychoanalytic Psychology, 18*(1), 3-20. doi: 10.1037/0736-9736.18.1.2

Bornstein, R.F. (2005). Reconnecting Psychoanalysis to Mainstream Psychology: Challenges and Opportunities. Psychoanalytic Psychology, 22 (3), 323-340. doi: 10.1037/0736-9735.22.3.323.

Cambray, J. (2014). Romanticism and Revolution in Jung's Science. In R . A. Jones (Ed.) *Jung and the Question of Science*. (pp. 9-30) . East Sussex: Routledge.

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

Chalmers, D.J. (2007). The Hard Problem of Consciousness. In Velmans, M. (Ed.) The Blackwell Companion to Consciousness (pp. 223-235). Oxford: Wiley Blackwell.

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences, 19*(7), 370-375. doi: 10.1016/j.tics.2014.03.003

Chen, M., Chiang, W. -. Y., Yugay, T., Patxot, M., Özçivit, İ. B., Hu, K., & Lu, J. (2016). Anterior Insula Regulates Multiscale Temporal Organization of Sleep and Wake Activity. *Journal of Biological Rhythms, 31*(2), 182–193. doi: 10.1177/0748730415627035

Chomsky, N. (2013). Power Systems. London: Hamish Hamilton.

Clifford, C.W.G. (2006). Binocular Rivalry. *Current Biology, 19* (12), 1-2.

Clodfelter, M. (2017). Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1492-2015 (4th ed.). Jefferson: McFarland & Co Inc.

Cohen, M. X. 2011. It’s about time. *Frontiers in Human Neuroscience* 5:2.

Cohen, M. X. (2014). Analyzing Neural Time Series Data: Theory and Practice. Massachusetts: The MIT Press.

Di Dio, C., & Gallese, V. (2009). Neuroaesthetics: a review. *Current Opinion in Neurobiology, 19*(6), 682-686. doi: 10.1016/j.conb.2009.09.001

Di Dio, C., Macaluso, E., & Giacomo Rizzolatti, G. (2007). The Golden Beauty: Brain Response to Classical and Renaissance Sculptures. *Plos|One*, *2*(11), doi: 10.1371/journal.pone.0001201

Di Dio, C., Di Cesare, G., Higuchi, S., Roberts, N., Vogt, S., & Rizzolatte, G. (2013). The neural correlates of velocity processing during the observation of a biological effector in the parietal and premotor cortex. *NeuroImage, 64*(1), 425-436. doi: 10.1016/j.neuroimage.2012.09.026

Dieter, K. C., Melnick, M. D., & Tadin, D. u. j. e. (2015). When can attention influence binocular rivalry?. *Attention, Perception, & Psychophysics, 77*, 1908–1918. doi: 10.3758/s13414-015-0905-6

Einstein, A. (1920). Relativity: The special and general theory. Penguin.

Ekman, P., Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review, 3*, 364–370.

Erman, A. (1971). *Life in Ancient Egypt.* New York: Dover Publications.

Feyerabend, P. (1975). Against Method. London: Verso Books.

Freud, S. (1913). Totem and Taboo: Resemblances between the psychic lives of savages and neurotics. London: George Routledge & Sons, Limited.

Goodwyn, E. (2010). Approaching Archetypes: Reconsidering innateness. *The Journal of Analytical Psychology, 55*(4), doi: 10.1111/j.1468-5922.2010.01862.x

Goodwyn, E.D. (2012). The Neurobiology of the Gods. Hove: Routledge.

Gordon S. (2015) Alan Watts and Neurophenomenology. *Self & Society. 43* (4) 311–321.

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *The Journal of comparative neurology, 521*(15), 3371–3388. doi:10.1002/cne.23368

Guendelman, S. (2017). Varela and Contemporary Psychoanalysis, an Epistemological Gesture: From Representations to Co-emergence. *Studies in Gender and Sexuality, 18*(4), 235-243. doi: 10.1080/15240657.2017.1383029

Henderson, L (2019). *The Problem of Induction*. Retrieved from https://plato.stanford.edu/archives/win2019/entries/induction-problem/

Hjørland, B. (2005). Empiricism, rationalism and positivism in library and information science. *Journal of Documentation, 61*(1), 130-155. doi: 10.1108/00220410510578050

Hogenson, G. B. (2019). The controversy around the concept of archetypes. *The Journal of Analytical Psychology, 64*(5), 682-700. doi: 10.1111/1468-5922.12541

Huntington, S.P. (1993). The Clash of Civilizations?. *Foreign Affairs. 72* (2), 22-49.

Izard, C. E. (2011). Forms and functions of emotions: Matters of emotion–cognition interactions. *Emotion Review, 3*, 371–378.

Janak, P. H., & Tyre, K. M. (2015). From circuits to behaviour in the amygdala. *Nature, 517*, 284-292. doi: 10.1038/nature14188

Jones, R. A. (2014). Romanticism and Revolution in Jung's Science. In R . A. Jones (Ed.) *Jung and the Question of Science*. (pp. 1-6) . East Sussex: Routledge.

Jung, C. G. (1977). The Tavistock Lectures: Lecture II (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 18, pp. 36-57). Princeton, NJ: Princeton University Press. (Original work published 1935)

Jung, C. G. (1975). The Archetypes and the Collective Unconscious: The Concept of The Collective Unconscious (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 9 pt. 1, pp. 42-54). Princeton, NJ: Princeton University Press. (Original work published 1936/37)

Jung, C. G. (1975). The Archetypes and the Collective Unconscious: Psychological Aspects of the Mother Archetype (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 9 pt. 1, pp. 75-112). Princeton, NJ: Princeton University Press. (Original work published 1954a)

Jung, C. G. (1975). The Archetypes and the Collective Unconscious: Archetypes of the Collective Unconscious (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 9 pt. 1, pp. 3-42). Princeton, NJ: Princeton University Press. (Original work published 1954b)

Jung, C. G. (1947). On the nature of the psyche. In Collected works, vol. 8. London: Routledge & Kegan Paul; Princeton: Princeton University Press (pp. 159-234)

Jung, C. G. (1977). The Archetypes and the Collective Unconscious: Foreword to Harding (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 18, pp. 518-521). Princeton, NJ: Princeton University Press. (Original work published 1948)

Jung, C. G. (1977). The Archetypes and the Collective Unconscious: Foreword to Jacobi: “Complex/Archetype/Symbol.” (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 18, pp. 532-533). Princeton, NJ: Princeton University Press. (Original work published 1956)

Jung, C. G. (1967). Symbols of Transformation: The Song of Moth (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 5, 2nd. Ed., pp. 79-121). Princeton, NJ: Princeton University Press. (Original work published 1952a)

Jung, C. G. (1967). Symbols of Transformation: Symbols of the Mother and of Rebirth (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), *The collected works of C.G. Jung* (Vol. 5, 2nd. Ed., pp. 207-273). Princeton, NJ: Princeton University Press. (Original work published 1952b)

Jung, C. G. (1977). Freud and Psychoanalysis: Answers to Questions on Freud (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), The collected works of C.G. Jung (Vol. 18, 2nd. ed., pp. 438-440). Abingdon-on-Thames: Routledge and Kegan Paul LTD. (Original work published 1953a).

Jung, C. G. (1977). The Archetypes and The Collective Unconscious: Foreword To Wickes: "Von Der Inneren Welt Des Menschen." (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), The collected works of C.G. Jung (Vol. 18, 2nd. ed., pp. 527-528). Abingdon-on-Thames: Routledge and Kegan Paul LTD. (Original work published 1953b).

Jung, C.G. (1964). Approaching the Unconscious. In Freeman, J. (Ed.) Man and His Symbols (pp. 1-94). London: Dell Publishing.

Jung, C. G. (1966). The Spirit in Man, Art and Literature: Paracelsus (R. F. C. Hull, Trans.). In H. Read et al. (Series Eds.), The collected works of C.G. Jung (Vol. 15., pp. 3-12). Abingdon-on-Thames: Routledge and Kegan Paul LTD. (Original work published 1929).

Jung, C. G. (1993). Three Versions of a Press Conference in Vienna. In W. McGuire, & R . F . C. Hull (Eds.) C.G. Jung Speaking: Interviews and Encounters. (3rd ed., pp. 38-46) . New Jersey: Princeston University Press. (Original work published 1928)

Kilner, J. M., & Lemon, R. N. (2013). What We Know Currently about Mirror Minireview Neurons. *Current Biology, 23*(23), 1057-1062. doi: https://doi.org/10.1016/j.cub.2013.10.051

Kirscher, T.T.J., Leube, D.T., Erb, M., Grodd, W., & Rapp, A.M. (2007). Neural correlates of metaphor processing in schizophrenia. *NeuroImage, 34* (1), 281-289. doi: 10.1016/j.neuroimage.2006.08.044

Keefer, L. A., & Landau, M. J. (2016). Metaphor and analogy in everyday problem solving. *WIREs Cognitive Science, 7*(6), 394-405. doi: 10.1002/wcs.1407

Knox, J. (2004). From archetypes to reflective function. *Journal of Analytical Psychology, 49* (1), 1-19. doi: 10.1111/j.0021-8774.2004.0437.x.

Knox, J. (2010). Responses to Erik Goodwyn's 'Approaching archetypes: reconsidering innateness'. *Journal of Analytical Psychology, 55*(4), 522–549. doi: 10.1111/j.1468-5922.2010.01863_3.x

Koyama, Y., Aoyagi, T., & Toyomaki, S. (2011). Phasic activity of the amygdala neurons during REM sleep. *Neuroscience Research, 71*, e171. doi: 10.1016/j.neures.2011.07.738

Kövecses, Z. (2005). *Universality and Variation in the Use of Metaphor*. Cambridge: Cambridge University Press.

Kuhn, T. (1998). Logic of Discovery or Psychology of Research?. In M. Curd & J . A. Cover (Eds.) *Philosophy of Science: The Central Issues* (pp. 11-21). London: Norton & Company Ltd. (Original work published in 1970).

Lakatos, I. (1998). Science and Pseudoscience. In M. Curd & J . A. Cover (Eds.) *Philosophy of Science: The Central Issues* (pp. 20-37). London: Norton & Company Ltd. (Original work published in 1973).

Lakoff, G. & Johnson, M. (1999). Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought. New York: Basic Books

Lakoff, G. (1997). How Unconscious Metaphorical Thought Shapes Dreams. In *Cognitive Science and the Unconscious*. UC Berkeley: Retrieved from: http://escholarship.org/uc/item/5f85g4bf

Lakoff, G. (2015, April 7). *George Lakoff: How Brains Think: The Embodiment Hypothesis* [Video file]. Retrieved from https://www.youtube.com/watch?v=WuUnMCq-ARQ&t=3682s

Laughlin, C.D. & Tiberia, V.A. (2012). Anthropology of Consciousness. Archetypes: Toward a Jungian Anthropology of Consciousness, 23 (2), 127 157. doi: 10.1111/j.1556-3537.2012.01063.x.

Lawrence, P. K. (1999). Enlightenment, modernity and war. *History of The Human Health Sciences, 12*(1), 3-25.

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review, 3*, 379–386.

Ma, L. & Liu, A. (2008). A Universal Approach to Metaphors. Intercultural Communication Studies, XVII (1), 260-268. Retrieved from http://web.uri.edu/iaics/files/26-Lin-Ma-Aihua-Liu.pdf.

Malinowski, J. E., & Horton, C. L. (2015). Metaphor and hyperassociativity: the imagination mechanisms behind emotion assimilation in sleep and dreaming. *Frontiers in Psychology, 6*(1132), 1-19. doi: 10.3389/fpsyg.2015.01132

McCarthy, M. (2012, December 12). Extravagant, yes, exotic, certainly – but black swans aren't as rare as you may think. *The Independent.* Retrieved from https://www.independent.co.uk/voices/comment/extravagant-yes-exotic-certainly-but-black-swans-arent-as-rare-as-you-may-think-8411679.html

McGuigan, R. (2009). Shadows, Conflict, and The Mediator. *Conflict Resolution Quarterly, 26* (3), 349-364. doi: 10.1002/crq.

Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's "Up" With God? Vertical Space as a Representation of the Divine. *Journal of Personality and Social Psychology, 93*(5), 699-710. doi: 10.1037/0022-3514.93.5.699

Merakchi, K. & Rogers, M. (2013) The translation of culturally bound metaphors in the genre of popular science articles: A corpus-based case study from Scientific American translated into Arabic. *Intercultural Pragmatics, 10* (2), 341-372. doi: 10.1515/ip-2013-0015

Merchant, J. (2009). A reappraisal of classical archetype theory and its implications for theory and practice. *Journal of Analytical Psychology*, 54, 339-358. doi: 10.1111/j.1468-5922.2009.01784.x.

Merchant, J. (2016). The image schema and innate archetypes: theoretical and clinical implications. *Journal of Analytical Psychology*, 61 (1), 63-78. doi: 10.1111/1468-5922.12194.

Mills, J. (2012). Jung's metaphysics. *International Journal of Jungian Studies, 5*(1), 1-25. doi: 10.1080./19409052.2012.671182

Moxnes, P. & Moxnes, A. (2016). Are We Sucked into Fairy Tale Roles? Role Archetypes in Imagination and Organization. *Organization Studies, 37* (10), 1519-1539. doi: 10.1177/0170840616634135.

Pace-Schott, E. F., Germain, A., & Milad, M. R. (2015). Effects of sleep on memory for conditioned fear and fear extinction. *Psychological Bulletin, 141*(4), 835-857. doi: 10.1037/bul0000014

Palmer, M. (1997). Freud and Jung on Religion. London: Routledge.

Panksepp, J. (1998). Affective neuroscience: The foundations of human and animal emotions. New York, NY: Oxford University Press.

Panksepp, J., Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review, 3,* 387–396.

Petitmengin, C., NaVaro, V., & Levanquyen, M. (2007). Anticipating seizure: Pre-reflective experience at the center of neurophenomenology. *Consciousness and Cognition, 16* (3), 746-764. doi: 10.1016/j.concog.2007.05.006

Piaget, J. and Inhelder, B. (1969) The Psychology of the Child. Basic Books, New York.

Popper, K. (1998). Science: Conjectures and Refutations. In M. Curd & J . A. Cover (Eds.) *Philosophy of Science: The Central Issues* (pp. 3-10). London: Norton & Company Ltd. (Original work published in 1963).

Rapp, A.M., Leube, D.T., Erb, M., Grodd, W., & Kirscher, T.T.J. (2004). Neural correlates of metaphor processing. *Cognitive Brain Research, 20* (3), 395-402. doi: 10.1016/j.cogbrainres.2004.03.017.

Ramachandran, V.S. & Hubbard, E.M. (2003). Hearing Colors, Tasting Shapes. Scientific American, 288 (5), 52-59. Retrieved from http://cbc.ucsd.edu/pdf/SciAm_2003.pdf.

Richter, M., Lins, J., & Schöner, G. (2017). A Neural Dynamic Model Generates Descriptions of Object-Oriented Actions. *Topics in Cognitive Science, 9*(1), 35-47. doi: 10.1111/tops.12240

Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., & McKay, R. (2001). Do facial averageness and symmetry signal health?. *Evolution and Human Behavior, 22*(1), 31-46. doi: 10.1016/S1090-5138(00)00060-X

RSPB. (2007). *I saw a black swan on new years day whilst walking by the river. Are they rare in the UK?*. Retrieved from https://ww2.rspb.org.uk/birds-and-wildlife/bird-and-wildlife-guides/ask-an-expert/previous/blackswans.aspx

Ruse, M. (1998). Creation-Science Is Not Science. In M. Curd & J . A. Cover (Eds.) *Philosophy of Science: The Central Issues* (pp. 38-45). London: Norton & Company Ltd. (Original work published in 1982).

Saban, M. (2014). Romanticism and Revolution in Jung's Science. In R . A. Jones (Ed.) *Jung and the Question of Science*. (pp. 30-49) . East Sussex: Routledge.

Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., & Spence, C. (2015). The sweetest thing: The influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Frontiers in Psychology, 6*, 1382. doi: 10.3389/fpsyg.2015.01382

Saroglou, V. (2016). Is Religion a Cultural Cause of Intergroup Conflict?. *Intergroup Conflict, Religious Fundamentalism, and Culture, 47* (1), 33-41. doi: 10.1177/0022022115621174.

Schmidt, G.L. & Seger, C.A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition, 71* (3), 375-386. doi:10.1016/j.bandc.2009.06.001.

Schmidt, G.L., DeBuse, C.J., & Seger, C.A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. Brain and Language, 100, 127-141. doi: 10.1016/j.bandl.2005.03.002.

Shank, J. B. (2008). *The Newton Wars and the Beginning of the French Enlightenment*. Chicago: University of Chicago Press.

Shibata, M., Toyomura, A., Motoyama, H., Itoh, H., Kawabata, Y., & Abe, J. (2012). Does simile comprehension differ from metaphor comprehension? A functional MRI study. *Brain and Language, 121* (3), 254-260. doi: 10.1016/j.bandl.2012.03.006

Shutova, E., Teufel, S., & Corhonen, A. (2013). Statistical metaphor processing. *Computational Linguistics, 39* (2), 301-353. doi: 10.1162/COLI_a_00124.

Singh, K. (2013). The Unstruck Sound: Archetypes of Rhythm and Emotion in Indian Alchemy and Jungian Analysis. *Culture & Psyche, 7* (2), 35-61. doi: 10.1080/19342039.2013.787887.

Searle, J. (1979): Metaphor. In: Ortony, A. (ed.): Metaphor and Thought. First Ed., Cambridge: Cambridge University Press

Sotirova-Kohli, M., Rosen, D.H., Smith, S.M., Henderson, P., Reece, S.T. (2011). Empirical study of Kanji as archetypal images: understanding the collective unconscious as part of the Japanese language. *Journal of Analytical Psychology. 56* (1), 109-132.

Stevens, A. (2006). The Archetypes. In R . K. Papadopoulos (Ed.) The Handbook of Jungian Psychology: Theory, Practice and Applications (pp. 74-93). Hove: Routledge.

Storr, A. (1973). *Jung*. London: Fontana Press.

Stringaris, A.K., Medford N.C., Giampietro V., Brammer M.J. & David A.S. (2007). Deriving meaning: Distinct neural mechanisms for metaphoric, literal, and non-meaningful sentences. *Brain and Language, 100* (2), 150-162. doi: 10.1016/j.bandl.2005.08.001.

Swick, D., Ashley, V., & Turken, A.U. (2008). *Left inferior frontal gyrus is critical for response inhibition*. BMC Neuroscience, 9 (102). doi: :10.1186/1471-2202-9-102.

Thagard, P.R. (1998). Science: Why Astrology Is A Pseudoscience. In M. Curd & J . A. Cover (Eds.) *Philosophy of Science: The Central Issues* (pp. 27-37). London: Norton & Company Ltd. (Original work published in 1978).

Three Initiates, (1908). *The Kybalion*. Chicago: Yogi Publication Society.

Tracy, J. L., & Randles, D. (2011). Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review, 3*(4), doi: 10.1177/1754073911410747

Trumpp, N. M., Traub, F., & Kiefer, M. (2013). Masked Priming of Conceptual Features Reveals Differential Brain Activation during Unconscious Access to Conceptual Action and Sound Information. *PLoS ONE, 8*(5), e65910. doi: 10.1371/journal.pone.0065910

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies, 3*(4), 330-49.

Von Franz, M.-.L. (1964). The Process of Individuation. In Freeman, J. (Ed.) Man and His Symbols (pp. 159-254). London: Dell Publishing.

Whitlock, J. R. (2017). Posterior parietal cortex. *Current Biology, 27*(14), 691-695. doi: 10.1016/j.cub.2017.06.007

Wiebe, S. A., Morton, B., Buss, A. T., & Spencer, J. P. (2014). Dynamic field theory. *Monographs of the Society for Research in Child Development*, **79**, 26–44.

Zanolie, K., Dantzig, S. v., Boot, I., Wijnen, J., Schubert, T. W., Giessner, S. R., & Pecher, D. (2012). Mighty metaphors: Behavioral and ERP evidence that power shifts attention on a vertical dimension. *Brain and Cognition, 78*(1), 50-58. doi: 10.1016/j.bandc.2011.10.006

Zeev-Wolf, M., Faust, M., Levkovitz, Y., Harpaz, Y., & Goldstein, A. (2015). Magnetoencephalographic evidence of early right hemisphere overactivation during metaphor comprehension in schizophrenia. *Psychophysiology, 52* (6), 770-781. doi: 10.1111/psyp.12408.

Appendix

Appendix 1: Schmidt and Seger (2009) Table illustrating metaphor imaging literature.

Table 1
Metaphor imaging literature. All known functional neuroimaging studies of metaphor processing published to date. Studies are listed with stimulus types grouped together, based on our reading of each paper. In addition, studies that found activation in right hemisphere areas are listed first (1–8).

Source	Experiment design	Stimuli	Main findings
1. Bottini et al. (1994)	*n* = 6, PET, block design	**Complex sentences**: ***novel*** metaphors, uncommon literal sentences	***Right*** frontal and temporal activations for metaphors compared to literal sentences
2. Sotillo et al. (2005)	*n* = 24, ERP LORETA spatial analysis	**Complex/poetry style phrases**: Spanish ***unfamiliar*** metaphorical phrase followed by metaphorically related or unrelated word	Higher N400 activation for metaphoric words localized to ***right*** MTG and STG
3. Mashal et al. (2005)	*n* = 15, 1.5 T fMRI block design, principal component analysis	**Word pairs**: literal relationship, conventional metaphorical relationship, or ***novel*** metaphor relationship	Novel metaphor processing: large network including left frontal and temporal areas, and ***right*** Wernicke's area, precuneus and insula
4. Mashal et al. (2007)	*n* = 15, 1.5 T fMRI block design, ROI analysis	**Word pairs**: literal relationship, conventional metaphorical relationship, or ***novel*** metaphor relationship	***Bilateral*** inferior frontal gyrus, middle frontal gyrus and superior temporal activations for metaphors, including ***right*** Wernicke's area
5. Arzouan, Goldstein, and Faust (2007)	*n* = 29, ERP LORETA spatial analysis	**Word pairs**: literal relationship, conventional metaphorical relationship, or ***novel*** metaphor relationship	***Right*** temporal and superior frontal involvement in novel metaphor processing
6. Pobric et al. (2008)	*n* = 12, rTMS	**Word pairs**: literal relationship, conventional metaphorical relationship, or ***novel*** metaphor relationship	rTMS of ***right*** posterior superior temporal sulcus disrupted processing of novel but not conventional metaphors
7. Stringaris et al. (2006)	*n* = 12, 1.5 T fMRI, event related, non-parametric analysis	**Simple sentence–word combinations**: conventional metaphors and literal sentences with related or unrelated words as targets	***Right*** ventrolateral prefrontal cortex activation to target words in the metaphoric but not literal condition
8. Ahrens et al. (2007)	*n* = 8, 1.5 T fMRI, block design	**Simple sentences**: Mandarin Chinese anomalous (***novel***) and conventional metaphors and literal sentences	***Bilateral*** middle frontal gyrus and precentral gyrus, left inferior frontal gyrus and fusiform, ***right*** superior frontal gyrus activations for anomalous metaphors compared to literal
9. Stringaris et al. (2007)	*n* = 11, 1.5 T fMRI, event related, non-parametric analysis	**Simple sentences**: ***conventional*** metaphors and literal sentences of the form 'Some X are Y'	Left inferior frontal gyrus and thalamus active for metaphors, no right activation specific to metaphors
10. Eviatar and Just, (2006)	*n* = 16, 3 T fMRI event related, ROI	**Simple sentences: *conventional*** metaphoric, ironic and literal sentences, following two sentence 'story'	Metaphors recruited mainly LH areas: inferior frontal and inferior temporal
11. Rapp et al. (2004)	*n* = 15, 1.5 T fMRI event related design	**Simple sentences**: German metaphors and literal sentences with the form an X is a Y, may have been moderately ***familiar***	Virtually no right hemisphere activations when comparing metaphors to literal sentences or grey screen baseline
12. Rapp et al. (2007)	Reanalysis of Rapp et al., 2004 data. ROIs subjected to laterality index calculation	**Simple sentences**: German metaphors and literal sentences with the form an X is a Y, may have been moderately ***familiar***	No differences in laterality patterns between literal and metaphorical sentences
13. Mashal, Faust, Hendler, and Jung-Beeman, (2008)	*n* = 14, 1.5 T fMRI	**Sentences:** novel metaphors, literal sentences, and unfamiliar non-sensical sentences	Stronger left dorsolateral prefrontal and posterior middle temporal activation for novel metaphors than for literal and non-sensical sentences
14. Lee and Dapretto (2006)	*n* = 12, 3 T fMRI	**Word triads**: literal relationship or ***conventional*** metaphorical relationship	Only left prefrontal and tempo-parietal activations, no right activations for metaphorical relationships