Memories, fantasies, archetypes: an exploration of some connections between cognitive science and analytical psychology

Jean M. Knox, Oxford

Abstract: The value of cognitive science as a means of investigating psychodynamic theory and practice is discussed and the limitations of this approach are described. Research findings from cognitive science are drawn on to clarify the nature of memory, which is seen to be a mixture of reproduction and reconstruction and the concepts of true and false memory are explored in this light. The part played by implicit memory and internal working models in producing transference is also examined. New ways of conceptualizing fantasy, which describes it as another facet of internal working models, and the role of transgenerational transmission of attachment patterns in creating internal working models are explored.

The nature of archetypes is considered in the light of cognitive science research and a minimalist model is proposed, in which they can be likened to image schemas, that is, primitive conceptual structures that exist in a form which can never be experienced directly or indirectly.

Key words: archetypes, cognitive science, fantasy, image schema, internal working model, memory, transference.

Introduction

In the first half of this century, the classical theory of behaviourism dominated academic psychology; it was a theory which represents the low point of any attempt to understand the human psyche, because it simply did away with mind altogether. One of the most famous behaviourists, B. F. Skinner actually said, ‘The question is not whether machines think but whether men do’ (Pinker 1997, p. 62). Behaviourism treated humans as mindless automatons whose essential spirit, our hopes, fears, desires and beliefs were all treated as an irritating irrelevance. It was a psychology of stimulus response, reflex arcs and salivating dogs which alienated generations of potential psychologists and can only have been attractive to people too out of touch with their own emotions to allow that those emotions might be a vital part of understanding the way we all function as human beings.

© 2001, The Society of Analytical Psychology
Published by Blackwell Publishers Ltd, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.
An information-processing model of the mind

In the 1950s and 1960s a revolution began to take place in academic psychology, a revolution essentially linked to the development of the computer. A computer-based model for the human mind may initially seem to be just as reductive as behaviourism because it also might seem to imply a predetermined set of rules governing the way the mind works, like a form of 'Windows' software installed in our heads, but this is very far from the truth about the computational theory of mind, or cognitive science.

The computational theory of mind is revolutionary in relation to behaviourism and crucially important to us as analysts because it is a theory about the way the mind deals with information and information can include anything that goes on in the human mind: hopes, beliefs, emotions, desires, dreams, fantasies, bodily sensations, perceptions; all these are information, all of equal value in the computational theory of mind even though some of them may be more difficult to investigate than others.

Cognitive science is a description of all the ways in which scientists try to understand the way the mind computes or processes information. It is therefore apparent that cognitive science covers many of the areas which we as psychotherapists are exploring clinically; we do aim to be therapeutic with our patients, but we also draw on our clinical work to attempt to gain an ever more accurate understanding of how the human mind works, an understanding based on the remarkable pioneering work of Freud, Jung, Klein and those who followed them. In our attempts to gain a better understanding of human mental functioning, psychotherapists are cognitive scientists as well, investigating the way our patients process the information of their dreams, fantasies and memories.

I would not, however, want to give the impression that I am suggesting that we can ultimately reduce the mystery of the human spirit to a series of computations, the processing of packets of information. We cannot and should not try to explain away the numinous or demonic aspects of our experience, painting, poetry, religion, our dreams or intuitions, the overwhelming experience of passionate love or hate, the tenderness of intimacy and the symbols which express these emotions. As therapists much of our skill lies in being attuned to the transforming symbols which our patients unconsciously communicate to us in the session; as therapists we cannot be merely scientific objective observers of our patients' mental processes but must also allow ourselves to be drawn in and sometimes taken over emotionally; we must be able to feel love and hate, sometimes towards and sometimes with or on behalf of our patients. It is this intuitive response to symbolic communication which plays as important a part in therapy as objective understanding.

Clinical illustration

A patient, whom I shall call Anna, had been telling me over the course of several sessions about her childhood and her relationship with her parents, but
in a way which left me feeling puzzled and confused because she seemed so objective and emotionally detached even when describing events which sounded very traumatic, such as the time when her father, a wealthy man, took the family up in his private aeroplane, with himself at the controls. He deliberately flew the plane in such a terrifying way that my patient’s mother became panic-stricken and tried to grab the controls out of his hands, whilst my patient watched, completely helpless, with both her parents out of control and all of them in real danger several thousand feet up in the air. However, Anna’s account of such events seemed very difficult to respond to because I always had the strong impression that she was not really interested in finding out what they meant to her but instead she was always watching my face closely to see what her descriptions meant to me – what effect they had on me. This pattern went on for some time, with her recounting various forms of emotional and physical abuse but with neither of us really knowing what was going on between us in the session. I certainly felt that I did not know what was required of me in response to the material she was giving me.

However, one day Anna told me the story of the family’s return from holiday, driving across Europe, once again with her father at the controls of the car. On this occasion, his omnipotence showed itself in his refusal to stop for refreshment in spite of the fact that it was very hot. They travelled hour after hour without stopping and Anna, aged about 7 or 8, found a bottle of wine in the back of the car which she opened and drank because she was so thirsty; neither parent stopped her and she became semi-conscious. Her mother subsequently told her that she was really worried that Anna might die but still did nothing, perhaps because she was unable to oppose Anna’s father. The car finally had to stop in a queue waiting for a toll bridge and Anna opened the car door and vomited red wine over the road in view of a horrified stranger who seems to have smelt the alcohol and protested strongly to her father, who even then lied and said she had eaten too many cherries. I do not remember the rest of that story, but I did recognize with great clarity at that point that she was telling me she needed me to be the stranger who could confront the mad omnipotence of her father, just as the other car-driver had done; she was also desperately afraid that I would be completely useless to protect her, as helpless as her mother had been, that I would not even realize how desperately ill she was or even worse, that if I did, I might not care and might even want her to suffer, which is why she had to watch my face so carefully whilst she told me about her past suffering.

This fear of hidden sadistic feelings in me, as in her mother, emerged in a dream some time later. She dreamt that she was in a large room in a house, which seemed pleasant and welcoming, but then a woman came into the room, came over to her and suddenly stabbed her with a knife, saying, ‘Let’s see how much pain you can bear’.
The scientific method and its relevance to psychotherapy

It is possibly true that some understanding of attachment theory and patterns of attachment might be helpful at some point in making sense of Anna’s accounts of her past life, but at that moment in the session, what mattered was that I was able to hear her, to allow myself to be put into the role of the stranger who could express in words the truth about her father’s abuse of her. Scientific understanding has no part to play at moments such as these in psychotherapy. They represent the aspects of our work which are outside the realm of science and depend upon our professional training in the art of being therapists, in which we develop our ability to use the countertransference to relate to our patients symbolically, to accept and work through the material they project onto and into us.

However, psychotherapists do need to be clearer about the value of the scientific approach in relation to the theoretical models that underpin our clinical practice. This value lies in the discipline of the scientific method which is always trying to narrow down its field of investigation, to make each problem it tries to solve into as small and precise a one as possible in order to exclude all the variables which might confuse the answer; if successful, scientists can produce the evidence with some confidence that the experimental results are relevant to the question being asked.

In contrast, all psychodynamic theories fall, at times, into the trap of trying to offer an explanation for all mental phenomena. We run the risk of working with increasingly outdated and inaccurate models of the human mind if we avoid subjecting them to the rigour of scientific scepticism, for fear that the numinous or spiritual will be destroyed by the scientific advances in understanding the way the mind actually works, the ways in which it processes information. Richard Dawkins argues in his book Unweaving the Rainbow that scientific understanding of the physics of the rainbow can produce just as much wonder, or as we might say, can be just as numinous as the belief that it is a symbol of God’s covenant with Noah after the flood. Dawkins says:

The feeling of awed wonder that science can give us is one of the highest experiences of which the human psyche is capable. It is a deep aesthetic passion to rank with the finest that music and poetry can deliver. It is truly one of the things that makes life worth living and it does so, if anything, more effectively if it convinces us that the time we have for living it is finite.

(Dawkins 1998, p. x)

I think that the more scientifically accurate our models of the human mind are, the better therapists we become because we can be much more relaxed about the fact that there is so much we do not know. Richard Dawkins vividly illustrates this point in his description of a lecture he attended as an undergraduate:

One of the formative experiences of my Oxford undergraduate years occurred when a visiting lecturer from America presented evidence that conclusively disproved...
the pet theory of a deeply respected elder statesman of our zoology department, the
ty that we had all been brought up on. At the end of the lecture, the old man
rose, strode to the front of the hall, shook the American warmly by the hand and
declared, in ringing emotional tones, ‘My dear fellow, I wish to thank you. I have
been wrong these fifteen years’. We clapped our hands red. Is any other profession
so generous towards its admitted mistakes?

(Dawkins 1998, p. 31)

Scientists do not by any means always achieve the altruistic ideal conveyed in
this anecdote, but the requirement to support theories with empirical evidence
means that a scientist does learn to live with uncertainty, with the possibility
that new evidence may at any moment cast doubt on a model which he or she
has spent a lifetime developing. In contrast, rigid analytic dogmas, like rigid
religious dogmas, can at times, seem to demand absolute conviction and cer-
tainty from their supporters, so that new evidence which might require revision
of the model is seen as a threat. This authoritarian strand can be traced back
to Freud’s inability to tolerate dissent from his view of the psyche and his need
for ‘disciples who would accept his doctrine without reservation’ (Ellenberger
1970, p. 669); it re-emerged with full force in the rejection of Bowlby’s revolu-
tionary ideas by the majority of the psychoanalytic community in his day
(Holmes 1993, p. 6).

A therapist whose mind is closed to new discoveries and developments in
theory or practice will miss valuable opportunities to increase his or her thera-
peutic effectiveness. Daniel Stern has vividly described how crucial this open-
mindedness is clinically; he has identified unpredicted ‘moments of truth’ which
suddenly arise in a session, moments which are laden with emotional significance
and with potential importance for the immediate or long-term future, a ‘kairos’
moment which must be seized or be lost. Stern says that in one of these ‘now’
moments, the patient and therapist are surprised and taken off guard by its
unpredictability. Both may defend against the anxiety this creates by resorting
to established technique. Stern says, ‘if the therapist “knows” what to do, he
has probably missed the “now” moment’ (Stern 1998; Stern et al. 1998).

I will now illustrate the general points which I have made by focusing on
some specific areas where scientific understanding has advanced enormously,
an understanding which as therapists we must incorporate into our theories,
whether we are Jungians, Kleinians or Freidians. In fact, I would even suggest
that these scientific discoveries may begin to break down these rather outdated
divisions between therapeutic schools; we all use the Freudian concept of
repression in our clinical work as well as the theoretical model of dissociation
which was Jung’s preferred explanation for the fact that some mental contents
are kept separate from others. We also almost universally work with object
relations, the idea that we all have internal mental objects which we relate to,
an idea which was most clearly expounded originally by Melanie Klein. We
learn to be reasonably comfortable as therapists with the experience of shifting
between these ways of understanding the clinical material of a therapeutic
session, but we certainly need the help of cognitive scientists and developmental psychologists to correct some of our mistaken ideas about how these clinical phenomena arise developmentally. To my mind, the key issue in this respect is the place of instinctual drive theory in our theoretical models; one of the most clear-cut distinctions between Freud and Jung and their early followers was the sharp division over the place of instinctual drive in shaping psychological processes. Jung rejected the sexual definition of libido, instead identifying it as a neutral psychic energy which is not attached to a specific instinct. He wrote: ‘I cannot see the real aetiology of neurosis in the various manifestations of infantile sexual development and the fantasies to which they give rise’ (Jung 1916, para. 574).

Jung’s view gains support from the work of developmental psychologists such as Daniel Stern and Joseph Lichtenberg who have shown fairly conclusively that instinctual drive theory should be discarded as an explanation for human psychological development and if that is the case then the bedrock of the distinction between Freudian and Jungian metapsychology has gone. On the basis of his and others’ empirical research, Stern said:

[I]nfants from the beginning mainly experience reality. Their subjective experiences suffer no distortion by virtue of wishes or defenses, but only those made inevitable by perceptual or cognitive immaturity … the capacity for defensive – that is, psychodynamic – distortions of reality is a later-developing capacity requiring more cognitive processes than are initially available … reality experience precedes fantasy distortions in development.

(Stern 1985, p. 255)

Memory

Let me therefore illustrate some of these general themes by focusing on the question of memory, a topic of central importance to the psychotherapeutic endeavour and one which has shown therapists how vulnerable we are to public censure when our theoretical models for understanding memory processes are shown to be faulty or inadequate by scientific investigation.

Two statements illustrate our difficulty:

- A memory may be true but may be repressed for long periods of time and then recovered.
- A memory may be held with great clarity and conviction but is actually false.

‘True’ means that the memory reflects real events which took place whilst ‘false’ means that the events which seem to be remembered did not in reality occur.

Both these statements are accurate and as therapists we must have a theoretical model of memory which allows for both these possibilities. John Morton, who
Chairled the British Psychological Society Working Party on recovered memories, summarized this in the statement:

Sometimes we remember events pretty much as they happened, sometimes we remember fabrications as if they were reality and sometimes we do a bit of both. Even when we do recount things as they happened, we are, simultaneously, likely to be remembering details incorrectly. All memories are a mixture of reproduction and reconstruction'.

(Morton et al. 1995)

There has been considerable scientific research conducted on the whole area of true and false memory which provides evidence that false memories can be constructed quite easily and then held with great conviction particularly when strengthened by frequent rehearsal. A beautiful example is given by Piaget himself from his own childhood:

One of my first memories would date if it were true, from my second year. I can still see, most clearly, the following scene, in which I believed, until I was about fifteen. I was sitting in my pram, which my nurse was pushing in the Champs Elysées, when a man tried to kidnap me. I was held in by the strap fastened round me whilst my nurse bravely tried to stand between the thief and me. She received various scratches and I can still see vaguely those on her face...

When I was about fifteen, my parent received a letter from my former nurse... she wanted to confess her past faults, and in particular to return the watch she had been given as a reward. She had made up the whole story... I must therefore have heard as a child, the account of this story, which my parents believed, and projected into the past in the form of a visual memory.

(Piaget 1962)

Subsequently there has been considerable experimental evidence for such false remembering. One such study showed that the test subjects could be made to ‘remember’ that they had been lost in a shopping mall when young; as part of this study, one 42 year-old subject was falsely convinced by his sister that he had been lost as a small child. She gave a vivid account of the event to him and a day later the test subject was sure that he could remember some of the events she described. Loftus describes five other subjects who were led quite easily to ‘remember’ events that had never taken place (Loftus & Coan 1994).

However, Loftus, who has conducted a whole range of experiments on the memory process, has also said that she too believes that ‘it is possible to lose contact with memories for a long time’, a statement from one of the most authoritative scientists who study memory and which contrasts with the conclusion of the Royal College of Psychiatrists’ Working Party that recovery of long-forgotten memories of abuse was most unlikely. However this report was widely criticized for failing to take into account research which demonstrates that while it is possible for false memories to be firmly implanted by suggestion, it is also possible for memories to be unavailable to recall for long periods of time and then recovered.
There are both experimental studies and clinical studies which are relevant to this issue. Carefully conducted experimental studies have demonstrated a phenomenon called state-dependent retrieval, which essentially means that we are more likely to remember an event which occurred under certain specific conditions if those conditions are reproduced at the time we try to recall the event. One particularly striking experiment showed that deep sea divers showed better recall of material they had learnt underwater when they were again underwater as opposed to recall when they were on land (Bower 1981). Other studies have shown that, although alcohol has a generally adverse effect on memory, experiences which occurred when a person was drunk are more likely to be recalled when that person is drunk again than when they are sober (Eich 1980). Changes in mood can contribute to state-dependent retrieval, so that events which occurred in one particular mood or frame of mind are more likely to be recalled when in the same state of mind again.

This phenomenon of state-dependent retrieval is clearly highly relevant to the issue of lost, then recovered, memories. A child, who is being sexually abused by a parent, experiences intense and distressing emotions. The experience is often kept totally secret, usually at the bidding of the abuser who calls it 'our secret', and may remain unintegrated into the rest of that child's life as (s)he grows up; (s)he may form satisfactory emotional and sexual relationships as an adult and the abuse remains either unavailable to recall, or if it is remembered it may be recalled without any re-experiencing of the intense emotional pain felt at the time of the abuse. However at any time that person may suddenly find him/herself vividly re-living the abuse in the most painful and terrifying way if something happens to trigger state-dependent retrieval. Patients of mine have described this experience as a result of watching the programmes or advertisements which warn about child abuse; the advertisements have actually been experienced as abusive themselves because the words and images, the picture of bedclothes moving, the words 'come and sit over here' have actually triggered an acutely painful reliving of their particular experience of abuse with all the distress they felt at the time.

Clinical studies have also demonstrated that memories may be lost and then recovered, although strenuous efforts have been made by a number of scientists to discount such findings as evidence of repression (Pope & Hudson 1995). The most frequently quoted study is one by Williams who followed up children with documented evidence of childhood sexual abuse into adult life; she found that adults who had documented experience of childhood abuse sometimes do not report these experiences at interview even when asked about such events; sometimes they would report the general experience of maltreatment but would seem not to remember specific episodes, of which there was documentary evidence from childhood notes (Williams 1994). Most strikingly of all, 12% of these adults with documented evidence of childhood sexual abuse completely denied that they had ever been abused by anyone at any time. Simple forgetting does not seem an adequate
So far, I have been talking about remembering specific events from the past, the kind of memory we are all familiar with and which is usually called declarative or explicit memory by cognitive scientists. However, in an editorial in the International Journal of Psychoanalysis, Peter Fonagy questions the role of this kind of memory in relation to therapeutic effectiveness; he says:

The aims of psychoanalysis have been greatly elaborated over the hundred years since Freud’s original model of undoing repression and recovering memory into consciousness. But these advances have not brought with them an updating of the role of memory in the therapeutic process, nor a clear and consistent theory of therapeutic effect. Some still appear to believe that the recovery of memory is part of the therapeutic action of the treatment. There is no evidence for this and in my view to cling to this idea is damaging to the field.

(Fonagy 1999)

So, if psychotherapy does not bring about change by the recovery of autobiographical memories, what is the role of memory in therapy? There has been an explosion of scientific research on the nature of memory, research which has demonstrated that there are several different types of memory and that some forms of memory can operate completely outside conscious awareness without any psychodynamic process such as repression or splitting being involved. A most striking illustration of this is the phenomenon of blind-sight, where damage to the occipital cortex results in the loss of the conscious sense of vision; in other words the patient apparently cannot see. However Weiskrantz found that when a light was flashed in the part of the patient’s visual field affected by the brain damage, the patient said he could not see anything, but when asked to guess the location of the flash he could do so quite accurately as though he were capable of some sort of unconscious perception (Weiskrantz 1986). The word ‘unconscious’ can simply mean outside awareness rather than dynamically repressed and as therapists we do need to have some understanding of this kind of research; otherwise we will be looking for psychodynamic explanations for such phenomena where none exist.

There is another way in which memories can exert an active influence, but without that person being aware of any conscious memory. Implicit memory is a form of memory in which past events influence our thoughts, actions and perceptions but outside consciousness. This can be quite a simple perceptual priming effect such as the fact that subliminal messages or advertisements flashed on to a screen so briefly that we are not aware we have seen them, can be shown to influence our judgements and choices. However, implicit memory may also involve complex conceptual knowledge, an unconscious memory, not for specific events but rather a kind of generalized knowledge about how to do things, procedural knowledge, and what things mean, semantic knowledge. Implicit memory is a kind of abstract generalized set of unconscious rules,
guidelines and expectations about how the world works and how one functions in it (Schacter 1996, chap. 6). Christopher Bollas’s formulation of the ‘unthought known’ suggests that he is describing aspects of implicit memory, although his concept also implies that there is innate instinctual knowledge which is activated by facilitating parenting, which is not a feature of implicit memory (Bollas 1987, 1989).

Implicit memory as the basis for the transference

In a previous paper I explored the relevance of the concept of implicit memory to the theoretical framework which informs our clinical work, in particular showing that it offers us new insights into the concept of the ‘complex’ (Knox 1999). However, the concept of implicit memory also has direct implications for our clinical practice, particularly in the way we conceptualize the transference and I would like to explore these applications further here.

Implicit memory forms the basis for most of the unconscious patterns of attitude, behaviour and expectations which our patients bring to their sessions; out of the experiences of a life-time, some of which the patient consciously remembers, but most of which he or she does not, a whole set of models of the world are constructed in implicit memory and these models structure or pattern the way that person relates to any new experience. Thus, implicit memory is the basis for the transference. The transference does not arise as an expression of instinctual drives which somehow spontaneously produce complex mental imagery and fantasy, which is then projected out onto real people; transference arises out of the internalization of actual people and real events in the world and gradually produces an unconscious pattern of generalized expectations about relationships.

The person who most clearly understood this was John Bowlby, whose views were unacceptable to many of his psychoanalytic colleagues at the time. I have already described the way in which the concept of implicit memory helps us to understand how internal objects develop, without our needing instinctual drive theory to explain why internal objects can differ so much from the real world. Implicit memory helps us to see how multiple real-life events become aggregated, mixed up together and that the fears or hopes which a person has at the time can also become incorporated into the memory of events. The end result of this process of internalization of multiple experiences was described by Bowlby as an ‘internal working model’; internal working models can be thought of as internal maps organizing our perception of the world (Bowlby 1988, p. 129). A securely attached child will have internal working models of other people as reliable and predictable and will bring this generalized expectation into the transference, whereas an anxiously attached child will have working models of other people as rejecting, dangerous or otherwise unreliable. Internal working models are the most obvious clinical examples of implicit memory; they are really forms of unconscious memory of an abstract kind and
demonstrate that it is possible for memories to be unconscious without being dynamically repressed. Working model ‘memories’ can never become a conscious memory of an event. The memories of events have been turned into an abstract pattern which can only be consciously discovered when it is relived, for example in the transference, and can then be analysed and modified. The pattern may, from time to time, be illustrated by specific memories from childhood, but it is not the remembering of these particular events which brings about change in therapy, but the slow modification of the working models through the relationship with the analyst; this creates its own working models which are hopefully more positive and which can be gradually integrated with existing working models or exist alongside them as alternatives.

The concepts of implicit memory and the internal working model have profound implications for our understanding of psychopathology and for the theoretical models which inform our clinical practice. The overcoming of repression and dissociation still have a part to play in analytic practice but there is an increasing recognition amongst analysts that our task is also one of constructing new internal working models of relationships with our patients. This approach has been explored by psychoanalysts such as Owen Renik, Robert Emde and the psychoanalytic ‘Process of Change’ study group, who suggest that analysis creates new implicit patterns of interpersonal relationship (Renik 2000; Emde 1999; Stern et al. 1998). For Jungians this ‘synthetic’ or constructive function of analysis is very familiar because Jung himself proposed that: ‘The aim of the constructive method, therefore, is to elicit from the unconscious product a meaning that relates to the subject’s future attitude’ (Jung 1921, para. 702).

Clinical illustrations

1) A patient has given me many descriptions of her parents, both doctors, as people who are unable to live with anxiety, uncertainty or helplessness and always look for someone to blame. She has many clear memories of occasions when her parents have blamed her or anyone other than themselves when things went wrong, but remembering these examples has not produced any change in the constant non-specific anxiety which she lives with every day and which was one of the reasons she came into therapy. However, during the course of the therapy she has gradually realized how much her anxiety arises because she waits to be blamed, including by me; she is constantly unconsciously expecting someone to say ‘it’s your fault’; much of the work in the therapy has involved a gradual recognition that quite often we are all helpless to prevent or solve problems, that may be just the way life is at times and it is not her task to put things right for her internal parents. The working model in this case seemed to be of a pattern of relationship, in which she, the child, was made responsible for allaying parental anxiety and would be blamed if she failed to do so. This pattern gradually emerged in her relationship
with me and making this conscious in our interactions is what has brought about change rather than the memories of actual childhood events.

2) Another patient has described the fact that her mother had numerous illnesses and sudden unexpected hospital admissions when my patient was a small child, which she remembers as experiences of abandonment. Her father was emotionally unable and unwilling to provide the secure parenting she needed at these times but an uncle who lived with them did offer her care and security; however after a year or so, this uncle started to abuse her sexually when he was looking after her.

As an adult she married a man whose work made it necessary for him to be away from home for periods of days or weeks without any prior warning and once again she felt profoundly abandoned, so much so that she was unable to function independently and would spend days in the house without going out, often curled up on her bed in a depressed state. She then met a man who was caring and protective towards her, helping her to manage practical aspects of her life whilst her husband was away; a sexual relationship developed between them, but this soon became abusive.

It seemed to me that unconscious ‘working model’ memories of abandonment were activated in adult life by her husband’s absences and this led to the recreation of a pattern of relationship with an initially caring man who then became sexually abusive, like her uncle. The ‘internal working model’ acted as an unconscious script, which possibly even influenced her initial attraction to a man with a similar personality to her uncle. In the analysis, she has become aware that she may have also unconsciously evoked the abusive attitude which her lover developed towards her, through her own attitude of neediness, seductiveness and passive compliance towards him. She realizes that she was relating to him on the basis of an unconscious belief that the only people who wanted to care for her would also abuse her and that she had to accept this because the alternative was to be abandoned. An unconscious ‘working model’ such as this might have led her to evoke abuse as a necessary, though painful, proof that she would not be abandoned.

**Fantasy**

Some aspects of the scientific investigation of memory can also help us when we turn to the question of the nature of fantasy.

Drive theory, which proposes that unconscious fantasy is a direct expression of instinctual drives, is a theory which many psychoanalysts believe in with great passion partly, I think, because it provides an explanation for the distortions of external reality which are so evident in our clinical work. We do need to be able to account for the way in which internal objects are formed and why they can differ so much from the actual people, usually parents who have played such a central part in the formation of our internal world. Instinctual drives give us an easy answer by proposing that the external reality is relatively
unimportant and that the magical or terrifying ‘good’ or ‘bad’ internal object arises as a direct expression of instinctual drives. In its purest form, external reality is not considered to play any part in the formation of unconscious fantasy, at least in the earliest stages of development. Thus Melanie Klein said that ‘the child anticipates, by reason of his own cannibalistic and sadistic impulses, such punishment as castration, being cut into pieces, eaten up, etc. and lives in perpetual dread of them’ (Klein 1927); one of her closest supporters, Joan Riviere said, ‘Psychoanalysis is Freud’s discovery of what goes on in the imagination … it has no concern with anything else, it is not concerned with the real world’ (Holmes 1993, p. 130).

This is the essence of the Kleinian position and is one which has had considerable influence on the metapsychology of developmentally orientated Jungians. This does create certain theoretical incompatibilities which as Jungians we have not really addressed because Jung himself emphatically rejected instinctual drive theory, arguing that the innate structures of the human mind do not come with prepackaged mental contents and are simply predispositions which organize information coming from the environment. The fact that there are certain similarities between his concept of archetypal polarization and the good and bad polarization of the Kleinian model is not enough to overcome the fundamental differences between the two theoretical frameworks. In Klein’s model of fantasy, the contents of the fantasy arise from within as though certain images pre-exist fully formed in the brain and are waiting to be released by a good feed or an experience of frustration. Jung on the contrary understood that mental imagery always has its origin in external experience which is then internalized and modified by innate or archetypal expectation.

In contrast to the Kleinian position which I have just outlined, we can see that cognitive science research really draws us towards the conclusion that memory can become fantasy and fantasy can become memory. In a symposium on the subject of recovered memory, Peter Fonagy, professor of psychoanalysis at University College London, made a similar point:

We are setting up truth against falsehood, history against phantasy, fact against desire ... (but) these pairs of opposites do not exist independently ... the dialectic of fact and desire is that fact makes desire and then desire makes fact, in an interminable sequence of events and thoughts that are repeated throughout life.

(Fonagy 1997, p. 126)

If, as I am suggesting, we should reject the idea that unconscious fantasy arises directly from instinctual drives, then we need to investigate what alternative explanations for the development of unconscious fantasy are offered by research based on attachment theory and cognitive science. I think that implicit memory can also give us new ways of thinking about the nature of fantasy as well as adding to our understanding of memory itself. Wishes and desires, fears and defences not only influence and distort the way we experience events but also form part of that experience and so themselves become incorporated
into the memories of events. Internal working models, the unconscious patterns of beliefs and expectations about relationships that are built up through the process of internalization, are not only a form of memory, but they are also a new way of conceptualizing unconscious fantasy. The child’s own emotions and the imaginative narratives he or she constructs to make sense of the world or to maintain a positive sense of identity become included in unconscious ‘working models’ as they develop. Working models of relationships are internal objects which include all the child’s conscious and unconscious imaginative material and his or her sense of identity as well as being generalized memories of real experience.

For example, a child who is subject to random and unpredictable violence from a parent will feel not only pain and terror but also a sense of complete helplessness, with no power to influence the parent’s behaviour or to have any control over the situation. In addition, without any apparent cause for the parent’s cruelty, the child has to face the intolerable fact that the parent is at that moment hostile, malevolent and sadistic towards the child. I think that for any child to feel this is unbearable and that it may be preferable in that situation for the child to construct a belief or fantasy that (s)he has done something to cause the parent to behave in this sadistic way; such an imaginative belief would allow the child to retain some sense of cause and effect, some belief that (s)he actually does have some control over the situation because (s)he did something wrong which provoked this violent response. The belief that (s)he caused the parent’s violence by some bad behaviour also allows the child to retain the belief that the parent will love her/him again, that it is the bad behaviour that is being punished and not that the parent really hates the child.

This kind of defensive belief becomes part of the unconscious ‘working model’ of relationships with key attachment figures and may emerge in relationships in adult life in the form of an unconscious fantasy that (s)he is responsible for others’ bad behaviour and should be punished for it; it might be quite easy for this person to become a battered spouse as a kind of enactment of the unconscious fantasy and/or memory. The internal working model produces a pattern of implicit beliefs and expectations that determine, for example, the choice of a partner and the nature of the relationship that subsequently develops. A person’s unconscious fantasy that (s)he is to blame if her/his partner is abusive could evoke the same belief in him/her and easily lead to the recreation of the childhood experience.

Attachment theory also offers us another way to understand the development of fantasy in the growing child’s mind, namely that of intergenerational transmission of attachment patterns. The research done by Ainsworth, Main, Fonagy and others has demonstrated clearly that the internal working models of the parents powerfully influence the growing child’s internal working models, reflected in the patterns of attachment which the child shows; a mother who has a dismissive rating on the adult attachment interview, a pattern which demonstrates an avoidance of emotionally painful memories about her own
life, is most likely to have a child who shows an avoidant pattern of attachment (Steele et al. 1996).

It seems that the parent’s internal working models are communicated to and internalized by their children, becoming part of that child’s internal world, in other words, part of that child’s fantasy. There is considerable research evidence that demonstrates this kind of intergenerational transmission of attachment patterns and, by implication, the internal working models which underpin that behaviour. An experiment by Broussard (1970) explored mothers’ fantasies about their babies soon after birth by asking them to rate their first born babies as better than average or not better than average at the end of the babies’ first month. Babies whose mothers had rated them as ‘not better than average’ were three times more likely to show clinical psychological problems at the age of four than the babies whose mothers had rated them better than average at one month. The predictive power of the mother’s initial fantasy of the worth of their infants continued to the age of 10 when the negatively viewed infants still had markedly greater diagnosable mental disorder than the more positively viewed babies.

To conclude this discussion of fantasy, I would emphasize that we do not need to lose the richness of the internal world of object relationships and unconscious fantasy if we give up the explanation that they derive from instinctual drives. For most of us as psychotherapists, the core of our work, which differentiates our profession from other forms of psychological treatment, is the fact that we explore unconscious fantasy and the internal object world of, and with, our patients. Attachment theory and cognitive science do not in the least take that away from us; on the contrary they offer us exciting new ways of demonstrating the existence of that internal world through the interpersonal research tools which are being developed. The key difference is that in attachment theory, the internal world is formed by a process of internalization of real experiences and real relationships, which are then imaginatively reconstructed and endlessly reworked internally, into generalized patterns of belief and expectation about relationships between oneself and key attachment figures, about oneself in the world, patterns which then powerfully influence all subsequent relationships.

I have suggested previously that Jung’s concept of the complex is very close to that of the ‘internal working model’ of attachment theory and that Jung does seem to have recognized the concept of unconscious patterns which influence us without our being aware of them, long before the name of implicit memory had been used to describe this. I have attempted in this paper to apply the theoretical insights which cognitive science offers us more directly to the clinical situation and the experience of the transference. As a clinical manifestation of internal working models in action, the transference can be described as both a form of memory and a form of fantasy; however the concept of fantasy can be integrated with contemporary scientific accounts of the ways in which the mind processes information; instinctual drive theory becomes superfluous.
as an explanation for the distortions which our patients bring to the consulting room.

**Archetypes**

I will finally turn to the subject of archetypes to see whether cognitive science can also help us to define and describe them better. I did not discuss archetypes in any depth when talking about fantasy and many Jungians would say that they do not use Klein’s model of fantasy at all and that, like Jung, they do not regard instinctual drives as the source of fantasy. Instead, many Jungians consider archetypes to be the source of the distortions of external reality and of our memories which fantasy creates.

There has been considerable debate in the *Journal of Analytical Psychology* on the nature of archetypes and some Jungian scholars such as Pietikainen would like to discard the biological basis of archetypes altogether and regard them as cultural symbolic forms (Pietikainen 1998). I think the problem with this view is that archetypes then would cease to have any explanatory value in relation to individual human psychology and become interesting anthropological abstractions, which cannot have any claim to being part of the universal human psyche. We are biological animals and we think, fantasize and produce mental imagery with our brains. Any theory of universal human mental processes has to be compatible with the working of the human brain; if archetypes are part of the universal human psyche, their imagery must arise out of the neurological processes of the brain.

The problem this creates for us as Jungians is I think the same as the problem that instinctual drive theory creates for Freudians and Kleinians. It seems as though archetypes are often thought of as pre-formed packets of imagery and fantasy, waiting to pop out fully formed given the right environmental trigger; on the contrary, the growing body of research from cognitive science is that the human brain just does not work in this way and young infants are simply not capable of the complexity of imagery and conceptual thought which both instinctual drive theory and this form of archetypal theory seem to require.

Cognitive science does offer us a way of thinking about archetypes which to my mind is actually closer to Jung’s predominant view. There is increasing scientific evidence that innate structures of the mind certainly do exist but that, like the internal working models of implicit memory, they exist in a form which is itself inaccessible to consciousness; there are no collective thoughts or images waiting to emerge into consciousness like butterflies from a chrysalis when activated by the right environmental cue. Innate structures of the mind are abstract patterns or sets of instructions which orient attention towards and select external events and experiences. The crucial difference between innate structures and internal working models is that the abstract content of innate structures is a direct expression of genetic codes, a set of hard-wired information-processing instructions about how to attend to and prioritize stimuli; on the
other hand, internal working models contain much more information-rich
generalized patterns which are created from accumulations of experiences and
so represent the meaningful patterns which have been learnt about the real
world. Innate structures are pre-experiential potentialities and internal work-
ing models are post-experiential constructions.

Antonio Damasio, a distinguished neurologist who has conducted extensive
research on the nature of consciousness, has linked both these implicit struc-
tures in his concept of ‘dispositions’. He points out that humans, as complex
organisms, have to respond to a large variety of complex and varied tasks,
saying,

The machinery needed to perform these demanding tasks is complicated and requires
a nervous system. It requires a vast stock of dispositions, a substantial part of which
must be provided by the genome and be innate, although some dispositions can be
modified by learning and additional stocks of dispositions can be acquired through
experience.

(Damasio 1999, p. 139)

Can research help us to identify more accurately the nature of innate cognitive
structures (innate dispositions) and draw on this knowledge to develop a con-
temporary view of archetypes which is compatible with such scientific models?
In a previous paper, I discussed the similarities between Jung’s concept of arche-
types and the primitive structures described as image schemas by cognitive
scientists such as Johnson and Mandler (Knox 1997). I discussed the function
of image schemas in providing a set of conceptual and spatial meanings and
illustrated this with a description of a series of experiments which have shown
that infants’ capacity to recognize faces depends on an innate mechanism which
contains primal structural information about the pattern and nature of a human
face. These experiments showed that it is the activation of this primal mechanism
which makes human faces so much more interesting to infants than many
other stimuli. In itself the mechanism does not distinguish between faces, but
by focusing the infant’s attention on any face it provides maximum oppor-
tunity for the infant to form an image schematic representation of the human
face. The image schema is not itself innate, but its formation depends upon
innate mechanisms.

The similarity between image schemas and the concept of archetypes has also
been noticed by Kotsch, who regards archetypes not as concrete images but as
unconscious structures which antedate conscious experience and as irrepressible
sources of images and ideas; archetypes ‘order experience without appearing
in it’ (Kotsch 2000). This view contrasts with that of Maloney who recently
suggested that archetypes can appear as themes which influence test subjects’
preferences for certain images (Maloney 1999). However his identification of
certain themes as archetypal is arbitrary, in spite of his use of factor analysis.
His claim that the factors he has identified do actually represent innate mental
content needs more substantiation than is given in his paper; it is a large step
from identifying certain mathematical factors to identifying that these factors are archetypal.

The concept of image schemas has mainly been experimentally investigated in relation to infancy because they offer a way of understanding the infant’s pre-verbal and pre-conceptual structures for organizing and making sense of experience. However Johnson does suggest that image schemas ‘exist at a level of generality and abstraction that allows them to serve repeatedly as identifying patterns in an indefinitely large number of experiences, perceptions and image formations for objects that are similarly structured in the relevant ways’ (Johnson 1987, p. 28). One illustration he gives is that of an ‘in-out’ schema which organizes a whole range of experiences, from physical movements, such as reaching into a cupboard to metaphorical uses of the schema, such as entering into a conversation. The problem remains that it is extremely difficult to provide experimental evidence to demonstrate this extension of the image schema into the realm of language and metaphor.

Johnson also begins to explore the crucial question of the biological nature of image schemas. He says that although image schemas are definite structures, they are dynamic patterns rather than fixed and static images. By dynamic, Johnson means that image schemas are the means by which we construct order and are not passive receptacles into which experience is poured. They are thus relatively fluid patterns that become altered in different contexts.

The newly evolving field of evolutionary psychology has taken the investigation of the biological basis for such mechanisms a little further. Cosmides and Tooby describe the artificial intelligence research which has dramatically demonstrated the vital part that innate knowledge plays in problem solving in the simplest tasks set for computers or robots. This innate knowledge is called a frame in artificial intelligence and Cosmides and Tooby draw on this research to clarify the nature of the frames in the human mind:

A frame provides a ‘world view’; it carves the world into defined categories of entities and properties, defines how these categories are related to each other, suggests operations that might be performed, defines what goal is to be achieved, provides methods for interpreting observations in terms of the problem space and other knowledge, provides criteria to discriminate success from failure, suggests what information is lacking and how to get it and so on’.

(Cosmides & Tooby 1992, p. 106)

How can such complex innate structures have developed in evolutionary terms? A debate rages currently among the giants in this field, such as Gould and Lewontin, who argue that many complex biological features such as language may be the emergent products of complex interactions between genetic predispositions and the environment, and Dennet, Pinker, Dawkins and Cosmides and Tooby, who are convinced that these are the direct products of natural selection ‘which creates reliably developing architectures that come equipped with the right frames and frame-builders necessary to solve the
adaptive problems the species faced during its history’ (Cosmides & Tooby 1992, p. 107; Dennet 1999, p. 51). In other words, natural selection is the process that shapes biological form to match function.

I suggest that although, like Freud, Jung contradicted himself at times, there is a lot of evidence that he viewed archetypes as innate predispositions with an organizing function. I think it is one of the possible models for archetypes that we need to research extensively by investigating the ways in which an understanding of ‘frames’ and image schemas may advance our concept. We need to gain a clearer understanding of the extent to which archetypes are ‘hard-wired’, genetically inherited structures or, on the other hand, self-organizing patterns that emerge as a central part of the development of meaning by the human brain as Saunders and Skar (2001) have recently proposed. Of equal importance are the ways in which the concept of archetypes may enhance our understanding of scientific models of innate mental structures and processes.

Conclusion

In summary therefore, scientific evidence suggests to us that the human mind works in the following way:

Innate mechanisms focus the infant’s attention onto features in the environment which are crucial to the infant’s survival; these mechanisms are biologically based and have arisen by the process of natural selection because they improve chances of survival. Innate mechanisms are activated by environmental cues, interacting with them and organizing them, leading to the formation of primitive spatial and conceptual representations (image schemas or archetypes). These form the foundation on which later, more complex representations can be built.

The environment provides information which is then taken in and stored in different forms in memory, some as conscious memories of specific events and some as unconscious patterns in implicit memory. In particular, patterns of relationship between self and other people are stored in implicit memory in the form of internal working models, which can be considered to be both a form of memory and a form of fantasy because the wishes, fears, dreams which we bring to an experience become incorporated into the memory of that experience. These internal working models offer us a new way of understanding object relationships and they pattern our expectations and perception of the present. They therefore form the basis of the transference in therapy. Although conscious memories of specific past events are often very useful illustrations which can activate internal working models, it is not the recovery of repressed memories in itself which brings about change in psychotherapy, but the gradual modification of the unconscious working models as they are enacted and worked through in the transference.
Cet article met en discussion la valeur des sciences cognitives en tant que moyens d'investigations de la théorie et de la pratique psychodynamiques et décrit les limites de cette approche. Les résultats des recherches appuyées sur les sciences cognitives sont dirigées dans la direction d'une clarification de la nature de la mémoire, vue comme un mélange de reproduction et de reconstruction. Les concepts de vrai et faux souvenirs sont explorés à la lumière des résultats de ces recherches et le rôle joué par les modes de travail intérieur et celui du souvenir implicite dans la constitution du transfert sont examinés.

Sont explorées de nouvelles façons de conceptualiser la production d'images, en décivant les représentations comme étant une autre facette modelant le travail intérieur; est aussi exploré le rôle de la transmission transgénérationnelle des structures de l'attachement dans la création des modes de travail intérieur. La nature des archétypes est considérée à la lumière de la recherche en science cognitive, et un modèle minimaliste est proposé, par lequel les archétypes peuvent être vus comme des schèmes d'images qui sont des structures conceptuelles primitives existant dans une forme qui ne peut jamais être vécue directement ou indirectement.


Neue Wege zum Konzeptualisierung von Fantasie werden untersucht, indem Fantasie als andere Facette innerer Arbeitsmodelle beschrieben wird; auch wird die Rolle der transgenerationellen Übermittlung von Bindungsmustern bei der Schaffung innerer Arbeitsmodelle untersucht.

Das Wesen von Archetypen wird im Lichte der Forschung in der kognitiven Wissenschaft betrachtet. Ein minimalistisches Modell wird vorgeschlagen, in dem Archetypen verglichen werden können mit Schemata von Bildern, primitive konzeptionelle Strukturen, die in einer Form existieren, die niemals direkt oder indirekt erfahren werden kann.

Vengono qui descritti l'importanza della scienza cognitiva quale strumento di investigazione della pratica e della teoria psicodinamica e i limiti di tale approccio. Vengono riportate delle scoperte fatte dalla scienza cognitiva volte a chiarire la natura della memoria, che sembra essere un insieme di riproduzione e ricostruzione. Alla luce di tali scoperte vengono esplorati i concetti di memoria vera o memoria falsa e viene esaminata la parte giocata dalla memoria implicita e dai modelli di funzionamento interno nella produzione del transfer.

Vengono investigati nuovi modi di concettualizzare la fantasia, che descrivono la fantasia come un altro aspetto dei modelli di funzionamento interno; viene anche
esplorato il ruolo svolto dalla trasmissione transgenerazionale di schemi di attaccamento nella creazione di modelli di funzionamento interno.

Viene considerata la natura degli archetipi alla luce delle ricerche della scienza cognitiva e viene proposto un modello minimale, nel quale gli archetipi possono essere paragonati a schemi di immagini che sono strutture concettuali primitive esistenti in una forma che non può mai essere sperimentata né direttamente, né indirettamente.

Se discute la validez de la ciencia cognitiva en la investigación de la teoría y la práctica psicodinámica y se describen las limitaciones de esta aproximación. Se describen los hallazgos de la ciencia cognitiva para clarificar la naturaleza de la memoria, la cual pareciera ser una mezcla de reproducción y reconstrucción. Se exploran los conceptos de real y falsa memoria desde la perspectiva de estos hallazgos de investigación y se examina la parte que juega la memoria implícita y los modelos de trabajo internos en la producción de la transferencia.

Son investigadas nuevas formas de conceptualización de la fantasía, describiéndose la fantasía como otra faceta de los modelos internos; así mismo son explorados los papeles de la transgresión en los modelos de relación en la creación de modelos de trabajo interior.

La naturaleza de los arquetipos es considerada a la luz de la investigación de la ciencia cognitiva y se propone un modelo minimalista, en el cual los arquetipos pueden ser equiparados con imágenes esquemáticas las cuales son estructuras conceptuales primitivas que existen en tal forma que nunca pueden ser experimentadas directa o indirectamente.

References


[MS first received July 2000; final version April 2001]