

Introduction

In the Western philosophical tradition, the human being has been characterised variously as *animal politicum*, and as *animal rationale*. Fifty years ago, Cassirer (1944) proposed another characterisation: *animal symbolicum*. Just as we humans are community-living creatures, and just as we are creatures endowed with the ability to reason, so, too – and this has long been acknowledged – we are symbol-producing, symbol-using and, often, symbol-dominated beings; the creation and use of symbols is central and distinctive in our behaviour and in our mental life.

Human behaviour and mental life being the specific concerns of psychology, it would seem reasonable to look to that discipline for an account of symbolism. But mainstream psychology has disappointingly little to say about the subject. As Bertalanffy (1981) observes, ‘In spite of the fact that symbolic activity is one of the most fundamental manifestations of the human mind, scientific psychology has in no way given the problem the attention it deserves’ (p. 42). In contrast, outside mainstream psychology there is no shortage of material on symbolism. Leaving aside psychoanalysis (for the moment), the humanist, phenomenological and existentialist movements, on the periphery of mainstream psychology because of their opposition to science, devote considerable attention to symbolism, and their contributions are joined by an even more extensive body of literature which spans the whole range of the social sciences: general philosophical treatments of the symbol, hermeneutics, sociology, anthropology, semiotics, aesthetics, and so on, each either appropriating symbolism as its own proper subject matter, or, at least, claiming to reveal valuable insights into symbolism.

This vast literature, however, is disorganised, confusing, and riddled with disagreements. Perhaps we should not be surprised by this, given the enormous range of phenomena encompassed by the concept of ‘symbol’, from the consciously formed and completely transparent symbols of, say, logic and mathematics, to the unconsciously formed and quite opaque symbols of the dream. But almost all of the controversial

2 Freud, psychoanalysis, and symbolism

literature on symbolism is concerned with only one end of this continuum; in general, the use of mathematical symbols is not regarded as contentious, or as posing interesting and difficult psychological questions, whereas what *is* contentious, what has been disputed for hundreds of years, is the explanation of symbols in myths, fairy tales, dreams, ritual, religion, art, psychopathological symptoms, and so on. In these areas consensus appears to be limited to two points: firstly, that the lack of a general, unified theory of the symbol, though regrettable, is an inevitable result of the nature of the symbol, which is ‘intrinsically complex’, ‘infinitely varied’, ‘multifaceted’; and, secondly, that the contribution of scientific psychology (including psychoanalysis) to theories of symbolism is necessarily limited, because the symbol is not amenable to scientific investigation; only a broader social-science perspective, a perspective whose eclecticism and opposition to ‘psychologism’ and to ‘psychological reductionism’ can accommodate the complexity of the symbol, holds any promise for the eventual emergence of a unified theory.

If these claims were true, they would indeed explain why mainstream scientific psychology has neglected the subject, and why there is no general theory of the symbol. But inspection of the literature reveals both claims to be unjustified. For one thing, the disorganisation and lack of unity in the existing treatments of symbolism are attributable less to the ‘complexity’ of the symbol than to conceptual confusions and other flaws in the various treatments. Any theory of symbolism, it seems to me, is obliged to respect certain logical constraints and meet certain psychological requirements. The contributions from the various areas which purport to deal with symbolism, or offer insights into it, show little awareness of these requirements, and so fail to meet them. For another thing, although symbolism does belong generally to the social sciences, nevertheless, when the logical constraints are identified, it is clear that the rejection of psychology is misguided. Since symbolisation is a three-term relation, one term of which is a cognising subject, one of the logical constraints on any theory of symbolism is that it *must* be a *psychological* theory. Not only must psychology play a part, that part is fundamental. As for the nature of the required psychological theory, symbolism *is* (as it must be) amenable to scientific investigation. Having said that, however, let me quickly add that, by ‘scientific’, I do not mean ‘positivist’ or ‘behaviourist’ (despite the widespread contrary misconception in contemporary psychological theory), but rather, realist, empiricist, and determinist. Unfortunately, these terms are themselves today much misunderstood. The history of psychology is not just a history of recurrent themes, it is a history of recurrent conceptual con-

fusions. Psychology's attempts to extricate itself from its philosophical roots have long resulted in the neglect of important theoretical issues. Critical thinking and conceptual analysis have been abandoned as esoteric and irrelevant exercises, and looked upon with suspicion and contempt, rather than acknowledged as necessary tools for conducting any kind of scholarship, including scientific inquiry. Today there are signs that theoretical issues in psychology are beginning, again, to receive the attention they deserve. But the signs are far from uniformly encouraging. As part of the broader intellectual *Zeitgeist*, the directing of the critical spotlight in contemporary psychology onto questions of conceptual clarity and coherence is too often advocated only by those who are opposed to science, and welcomed only because it is thought to go hand in hand with 'marginalizing facts' and with recognising that 'the very idea of an "independent" world may itself be an outgrowth of rhetorical demands' (Gergen 1991, p. 23). This attitude is combined with an appeal for psychology to move beyond the sterile, outmoded 'Rhetoric of Scientific Truth' (Ibañez 1991, p. 187), and to embrace instead the post-empiricist, postmodernist 'turn', whose key achievements have been to expose the fallacy of objective science, and to unmask all theory as mere ideology. Similarly, what is identified (and rejected) as 'empiricism' is the supposed scientific aim of discovering indubitable truths, an aim which is illegitimately conflated in contemporary theory with objectivity (realism). This conflation has resulted in the proclaiming of the 'waning of empiricist foundationalism' (Gergen 1991, p. 13), and of the victory of relativism, that is, the victory of 'traditions marginalized within this century by the empiricist hegemony, metatheories of long-standing intellectual currency removed from common consciousness by the prevailing practices' (ibid., p. 16). Consequently, having been told almost three decades ago that 'Philosophically, the heyday of realism is receding into the past' (Palmer 1969, p. 221), those who take the presently unfashionable step of supporting realism find that they have to make an unusual effort to justify that step.

It is not my purpose here to mount a detailed defence of philosophical realism. But, briefly, there is no question that the 'demise' of realism has had much to do with the mistaken equating of it with some kind of self-proclaimed path to the indubitable, and with aspects of positivism and behaviourism. Greenwood (1992) shows how social constructionists, for example, misrepresent scientific realism by assimilating to it a number of features (e.g., operationalism, verificationism, instrumentalism) that are supposed to be associated with empiricism, and he presents realism as a choice of theoretical stance which has been misunderstood, and which, *properly* understood, offers much more than

4 Freud, psychoanalysis, and symbolism

has recently been appreciated. Stove (1991) goes further, suggesting that realism, properly understood, is not even an *option*; it is not contingent, not a scientific *theory*, but a *precondition* of discourse and understanding, the only apparent alternative, solipsism, being unworthy of serious consideration. As for determinism, that is perhaps less controversial. ‘Without a causal structure’, says Hart (1982), ‘the mind should probably be denied to have a nature; and if the mind had no nature, there would be precious little for a scientific psychology to discover’ (p. 193). Conversely, anything which has a nature must be bound by the constraints of that nature. As Anderson (1936) points out, ‘it is a condition of a thing’s existence that it determines and is determined by other things, and . . . to investigate or “give an account of” it involves consideration of such determinations’ (p. 123). Therefore, ‘Those who are interested in mind’s workings will naturally take up a determinist position . . . Theoretical concern with what is the case is, it seems to me, coextensive with determinism’ (ibid., p. 125). Furthermore, any attempt to import a partial non-determinism via the postulation of a particular version of the free-will/determinist ‘interactionist’ position effectively denies *any* determinism; if the same set of physical antecedents leads, on one occasion, to a certain set of physical consequents, and, on another occasion (that of the intervention of a free agent), to a different set of physical consequents, then it is clear that ‘there can be no physical uniformity’ (ibid., p. 124). This illustrates Anderson’s insistence that the strongest defence of a particular theoretical position consists in demonstrating that the opposition must implicitly assume it in the process of explicitly rejecting it, and so can be shown to hold a view which ‘amounts to the same as contradicting the possibility of discourse’ (ibid., p. 123). Arguments such as these suggest that many contemporary anti-scientific movements are self-contradictory – they are logically dependent on the realism and determinism which they explicitly deny. In the case of symbolism, then, only a theory which is realist, empiricist, and determinist will be genuinely explanatory. I am well aware that many readers would not be satisfied by this all too brief nod towards a defence of the philosophical realism of my position. I can only request them to reserve final judgement until the end of the book, for many of the discussions throughout the rest of the book serve as developments of the points I have made here.

Bearing these points in mind, then, let me return to the unjustified claims made in the literature on symbolism. Not only is the disorganisation there *not* the result of the complexity of the symbol, and not only must a general theory of the symbol be both psychological and scientific, but the material for such a theory is in fact available. That material is

to be found in Freud's writings. There, as I shall show, the groundwork for a scientific treatment of symbolism has been laid, in that respect confirming Badcock's (1980) observation that 'the relative failure of the human sciences to provide convincing and exact explanations of cultural phenomena is in large part to be attributed to their failure to take account of Freud' (p. 2). There are two reasons why Freud's contribution to a general theory of the symbol has gone unrecognised. Firstly, as Ricoeur (1970) points out, 'a systematic study of Freud's notion of symbol remains to be done' (n. p. 97). *A fortiori*, there has been no critical evaluation of Freud's writings (direct and indirect) on symbolism. Secondly, as a psychological theory, psychoanalysis is caught between two hostile movements in psychology, each rejecting it, but each, ironically, locating it in the other camp. On one side are the humanist, idealist, phenomenological, and existentialist psychologists, united by an anti-scientific stance, an insistence that symbolism cannot be studied scientifically, and the claim that psychoanalysis – classical psychoanalysis, at least – is hampered by its 'scientistic' restrictions and misconceptions: its narrow determinism, outmoded realism, and 'reductionist' bias. On the other side are the mainstream, largely 'experimental', psychologists, united by their scientific stance, and by the rejection of psychoanalysis (including any psychoanalytically based theory of the symbol) as 'unscientific'. Little wonder that, on the one side Freud alone, and on the other side Freud and symbolism together, have not been given the attention they deserve.

It is my contention, then, that a general theory of the symbol derived from Freud's psychoanalytic theory *is* possible, and that the resulting theory *is* scientific. Investigation of Freud's writings reveals that, in those writings, there is a genuine foundation for such a theory. Admittedly, the identification of that foundation requires considerable extraction and critical textual exegesis, to show that what is of value is not what is usually identified (even by Freud himself) as his theory of symbolism (i.e., the narrow view onto which his ideas converged during the years 1914–17), but, rather, a broader view, for which a schema is discernible in his earlier writings, and whose individual themes were elaborated later, although Freud, for various reasons, did not recognise the unifying role of those themes. Those aspects of the narrow view which are conceptually flawed, and which have left (what is generally regarded as) Freud's theory of symbolism open to easy dismissal, must of course be rejected, but the rest are assimilable into the broader view. As it stands, however, that broader theory will not do. A coherent and defensible general theory of the symbol does not appear until a number of major issues in Freud's writings have been revised and clarified. These

6 Freud, psychoanalysis, and symbolism

revisions are required not only to safeguard the theory from certain potentially damaging criticisms, criticisms which have caused concern to Freud's defenders, but also for establishing the soundness of the theoretical basis of psychoanalysis, upon which the broad theory of the symbol rests. The direction of these revisions is towards maintaining Freud's explicit commitments to realism, empiricism, and determinism, commitments in which he sometimes wavers in his metapsychological treatment of the central concepts of his theory. Those commitments underlie his contribution to one of the major concerns of twentieth-century psychology and philosophy of mind, that of human action and its explanation by means of the traditional 'desire plus belief' model, a model which is currently popular in the 'folk psychological' treatment of humans as 'intentional systems', but which has also come in for much (often justified) criticism. Some of the sounder aspects of Freud's contribution along these lines have been brought out recently by, for example, Maze (1983), Hopkins (1988), Wollheim (1993), and Gardner (1993). But the revisions which I am proposing here are particularly relevant to a successful consolidation of the general theory of the symbol. Once that consolidation has been achieved, I shall be able to show that the theory does meet the logical and psychological requirements which constrain any theory of symbolism, and that its meeting these criteria contrasts with the failure of other approaches, which are left unable to offer any serious challenge, either in the form of a successful alternative, or in the form of valid criticisms of the psychoanalytic approach.

Part One

Exegesis and Extraction

My main concern in Part One is to show, via chronological exegesis of Freud's published writings, that what is of value in those writings for a general theory of the symbol is not what is commonly identified as his contribution to symbolism. I shall first present in detail (in Chapter 2) that 'standard' picture of Freud's position, a position whose easy rebuttal reveals why his contribution has been underestimated. In Chapter 3, I begin the chronological investigation by demonstrating that Freud's early writings, typically dismissed as irrelevant to his theory of symbolism, contain the foundations for a different, much broader, approach to the symbol. Chapters 4 to 6 trace the development of those early themes in Freud's subsequent work, showing how they were continued and elaborated, albeit in a scattered and unsystematic fashion, and alongside the 'standard' narrow view. Before turning to Freud's material, however, let me depict the scene which first confronted me, and which would confront anyone approaching this field with the question: can there be a general theory of the symbol?

1 From disorder towards the focus of inquiry

The problem of definition

‘There is something very curious in semantics’, says Lévi-Strauss (1978), ‘that the word “meaning” is, probably, in the whole language, the word the meaning of which is the most difficult to define’ (p. 12). Perhaps because of its intimate connection with the concept of meaning, the term ‘symbol’, despite an extensive literature devoted to the subject, is almost as difficult. Derived from the Greek verb *συμβάλλειν* (literally, ‘to throw together’), the noun *σύμβολον* (a ‘tally’) referred originally to each of the two corresponding pieces of some small object which contracting parties broke between them and kept as proof of identity (Liddell and Scott 1968). That meaning subsequently expanded to include a diversity of meanings (other kinds of token, seal, contract, sign, code, etc.), which today has mushroomed even further. Many contemporary definitions reflect the mystique originally associated with symbols, and which prompted Whitehead (1927) to comment on the ‘unstable mixture of attraction and repulsion’ (p. 60) in our attitude towards symbolism. But the most frequent observation is that it is impossible to find a general, unifying definition. Bertalanffy (1981), for instance, complains that ‘in spite of the fact that symbolic activity is one of the most fundamental manifestations of the human mind . . . there is no generally accepted definition of “symbolism”’ (pp. 41–2), and Safouan (1982) warns that ‘anyone who tries to study symbolism in all its generality is liable to discover that there is no unity at all that underlies these different uses of the word’ (p. 84). To underscore the point, we are faced with such vacuities as ‘whatever has meaning is a symbol, and the meaning is whatever is expressed by the symbol’ (Radcliffe-Brown, in Skorupski 1976, p. 117), or, worse, ‘wherever we look around us, everything can be expressed by the concept of symbol’ (Ver Eecke 1975, p. 28). Even amongst those who bring some rigour to their treatment of the topic, there is considerable disagreement: disagreement, for example, about how to classify signs and symbols – what

is the difference (if any) between the two, which is the broader term, and to which does language belong; disagreement also about the nature of symbolism – what constitutes symbolism, what activity may properly be described as ‘symbolic’, what are its origins, development, role, effects, and so on.

But this picture of disorder should not lead us to agree too readily that there can be no general theory of symbolism. Instead, by considering the definitions of symbolism from two different perspectives in turn, the first an overview both of the scope of the subject matter and of the extent of the disorder, the second allowing a convergence on the real centres of controversy, we shall find ourselves on a journey which leads through the disorder towards a focus of inquiry.

Perspective one: the broad to narrow continuum

The more obvious perspective is to regard the enormous range of definitions of symbolism as lying along a continuum, from very broad definitions to extremely narrow ones. At the broad extreme we find the **symbol as superordinate category**. Here are located the ‘Bibles’ of symbolism (as Bertalanffy (1981) calls them): Cassirer’s *The Philosophy of Symbolic Forms* (1953, 1955, 1957 [orig. 1923, 1925, 1929]), and Langer’s *Philosophy in a New Key* (1942). Each of these works is neo-Kantian in spirit; philosophical concern with the question of ‘meaning’ intersects with the treatment of symbols from a strongly phenomenological, constructivist perspective. Cassirer’s Kantian debt is the more marked; for him, the ‘symbolic’ is equated with ‘structure’ or ‘form’, and it is the symbolic concept, not the semantic, that is truly universal. Thus, ‘the conceptual definition of a content goes hand in hand with its stabilization in some characteristic sign. Consequently, all truly strict and exact thought is sustained by the *symbolics* and *semiotics* on which it is based’ (1953, p. 86). Langer also says that symbolisation is the *essential* act of thought, and that ‘The symbol-making function is one of man’s primary activities, like eating, looking or moving about. It is the fundamental process of his mind, and goes on all the time’ (1942, p. 41). Unlike Cassirer, however, Langer combines this broad definition with a more modern information-processing view of thinking, according to which ‘the human brain is constantly carrying on a process of symbolic transformation of the experiential data that come to it’ (ibid., p. 43). This notion is still, of course, popular, especially with those involved in computer applications in psychology and artificial intelligence. McCorduck, for instance (in Graubard 1988), suggests that artificial intelligence might be the best hope for discovering that ‘universal

symbolic code', that 'set of universal concepts', which 'underlies all human symbolic expression' (p. 82). Others who appear to follow the general Cassirer–Langer broad approach include: Rapoport (in Royce 1965), for whom symbols are 'products of the human abstracting process' (p. 97), Hayakawa (*ibid.*), for whom symbolism is 'that which shapes the entire psychic life of man' (p. 92), and Whitehead (1927), for whom symbolism 'is inherent in the very texture of human life' (p. 60). Piaget too, though he is not consistent, occasionally treats symbolic behaviour as being almost as broad as what he terms 'operational intelligence', and his philosophical perspective is similarly neo-Kantian and constructivist. In general, what characterises these very broad definitions of symbolism is the view that the 'symbolic' is universal because it is somehow fundamental to the thinking process.

At a little distance from the broad end of the definitional continuum are treatments of the **symbol as a kind of sign**; the sign is the generic term and the symbol is the special case, albeit special in different ways for different theorists. This view is typical of semiologists or semioticians. As Todorov (1982) says, 'if one gives the word "sign" a generic meaning through which it encompasses that of symbol (the symbol then becomes a special case of the sign), one may say that studies of the symbol belong to the general theory of signs or semiotics' (pp. 9–10). Eco (1973) defines a sign as 'anything that can be taken as "significantly substituting" for something else . . . a sign is something (whether a natural or an artificial object) which stands in place of something which is absent' (p. 1149). Hawkes (1977) points to the culmination of the historical development of a general theory of signs in Jakobson's synthesis of the work of Peirce and Saussure – a curious combination, given the radically different views on the concept of 'symbol' held by these two. For Peirce, the American 'founder' of semiotics, the tripartite division of signs produces the icon, the index, and the symbol, the last being the case where the relation between signifier and signified is arbitrary; thus the major systematic manifestation of symbols is in language. Saussure, on the contrary, held that it is the *sign* which is arbitrary, and the *symbol* which is non-arbitrary or 'motivated', and so does not properly belong to the field of semiology (which, of course, locates Saussure's position further along our definitional continuum, in a region where the symbol is no longer a *kind* of sign, and where 'affect' plays a crucial role). In Jakobson's synthesis, the Saussurean fundamental dimensions of language – the syntagmatic and the paradigmatic – will be found in *any* symbolic process or system of signs. Bertalanffy (1965) follows Peirce; for him also the sign is the broader term, deriving from the general notion of 'meaning' (i.e., representation), and symbols are *kinds* of