

Andy Blunden 2011

## The Psychology of Concepts

(from *Concepts. A Critical Approach*.)

### Part I. Contemporary Theory, Chapter 1)

Theories of the psychology of concepts are as diverse as psychology itself. But I will limit myself to a fairly general review of the most well-known theories, and without paying attention to the evolution of the views of particular writers.

The dominant trend of the psychology of concepts in the Anglophone world at the moment is the American current which originated from the Cognitive Revolution of the 1950s with people like Jerome Bruner, Jacqueline Goodnow, George Austin and others. The most well-known names in recent work on concepts are Eleanor Rosch, Douglas Medin, Lawrence Barsalou, Edward Smith, Gregory Murphy, Susan Carey, Eric Margolis, Stephen Laurence and the philosopher Jerry Fodor. There are others who have reacted critically to this current, and I will come to these later.

These writers seem to agree that nothing was known of the psychology of concepts prior to the Cognitive Revolution, and apart from gestures to Gottlob Frege, Ludwig Wittgenstein and Plato, nothing of value had been established in the science beyond their own work. Even though they are rediscovering ideas which are centuries old, the *laboratory methods* of investigation they have brought to bear on the theory of concepts are novel and have brought about a rapid turn-over in theories of the concept.

The Psychology of Concepts (when I use capital-P and capital-C I am referring to the current of Cognitive Psychology) locates itself squarely within the traditions of experimental natural science and a variety of analytical philosophy which accepts a naïve variant of Cartesian dualism. (Natural scientists always strenuously deny their Cartesianism, but usually because they know only a caricature of Descartes' position. Misunderstandings often hinge on taking the word 'substance' in the vulgar sense of referring to some kind of stuff or simply failing to grasp the categorical difference between consciousness and matter). Although a number of different theories have been developed, a common philosophical base has meant that this group of investigators share a number of key assumptions.

The first of such assumptions in the Psychology of Concepts is a pervasive Cartesian dualism, for example, Gregory Murphy says in the introduction to "The Big Book of Concepts":

In general, I try to use the word *concepts* to talk about mental representations of classes of things, and *categories* to talk about the classes themselves (2004, p. 5).

And the identical form of words is used by Medin, Unsworth, and Hirschfeld in their contribution to the "Handbook of Cultural Psychology" (Medin et al, 2007). It is taken as given that concepts are mental representations, that is, that concepts are entities or images of some kind inside the head. So we have two distinct worlds: a world of mental objects and a world of material objects. It is further assumed that these concepts sort objects in the material world into categories, with each object individually isolable in some unspecified way prior to being grouped into categories. The field of investigation is thereby narrowed to deal only with those concepts which can be adequately represented by the process of 'pigeon-holing' things, as we say, in the hope that this can be usefully extended to the solution of more realistic problems.

The second assumption is that the objective world is arbitrarily atomistic, being composed of individual entities, and further that these individual entities in turn can be conceived as objects which are exhausted without remainder by their attributes (or 'features'). Some writers observe with some curiosity that their experimental subjects, on the contrary, seem to believe that objects have an essence transcending their contingent attributes. Ruth Millikan put it well when she said: "A dog is a member of the species dog because it was born of a dog, not because it is like other dogs" (Margolis & Laurence 1999, p. 525). One could just as well have said that the President is not President because he looks presidential, but because he was elected.

Generally speaking, these writers accept that people acquire concepts through social experience, but in line with this whole current of thinking, in their experimental work, they strive so far as possible to isolate their subjects from their normal living conditions and create laboratory conditions artificially insulated from social life. The view is that even if people have acquired the concepts they have in the course of life experience, that has no significance for the nature of the concepts themselves and there is no suspicion that tests carried out in the laboratory using made-up concepts will fail to reproduce what is essential to the psychology of concepts. But does the means a person uses to make a selection in a multiple-choice questionnaire reflect the way they act in the course of some social interaction?

Finally, for this current of cognitive psychology, the archetypal concept is the concept of a thing, typically an artificial object with no special social significance, which can be exhaustively described in terms of its visual properties, or a common artefact or a natural kind, depending on what the researcher is trying to prove. Everything that is to be learnt of the psychology of concepts is to be developed from such models. As Gregory Murphy put it:

Although the field (and hence this book) concentrates on common object concepts, the principle involving concept formation and use are thought to be to some degree generalizable across different domains and settings (Murphy 2004, p. 3).

So, essentially we have a theory of object-categorisation, rather than a theory of concepts. Consider for a moment how far the practice of pigeon-holing reflects the understanding of concepts like those mentioned in the introduction. Categorisation is relevant only at the margins, when, pressed to make a categorisation decision, you seize upon some criterion to be made the litmus test in the immediate instance. But locating the border lines of a phenomenon does *not* tell you about the concept of the thing itself. For example, if you were to show a subject a series of figures graduated from a circle to a square with rounded corners of successively smaller radius, you could determine at what point subjects took the figure as a square and ignored the rounded corners. But such figures do *not* give us the concept of a square, a figure which has *no* rounded corners, even minuscule ones. You could take a statue of the Virgin Mary and mess around with her features until it was no longer recognisable by a Catholic as a representation of the Holy Mother. But this gets you no closer to the concept of the Virgin Mary. Exploration of borderline cases does have points of interest for psychology, but these need not concern us here, because borderline cases do not shed very much light on the psychology of concepts. Even when the exploration of borderline cases allows us to identify the feature which the subject is taking as definitive, we can do no better than replace the problem of the entity with the problem of the feature and thereby enter into an infinite regress. For example, if people judge the quality of fruit by features such as size, colour and freedom from blemish, they may well select fruit which has been subject to irradiation and hormone injections but are of poor quality. We often rely on inessential features as fallible guides to what we know to be essential.

The world is not only seen atomistically, but as *arbitrarily* atomistic, because no justification is ever given for what is taken as the units of analysis in any particular study. In one study the minutiae of calligraphy distinguish the digit “1” from the 12<sup>th</sup> letter of the lower case alphabet and the 9<sup>th</sup> letter of the upper case alphabet. But in another study, “1” is a simple concept, while in yet another study, the number 1796 is taken as a “chunk” and may be treated as an atom, rather than each of its 4 digits. Why not take every dot from the laser printer as the unit of analysis? Why an entity may be taken to be the composite of elements of some arbitrary size is never discussed. And yet the *whole problem* is contained in the selection of what are taken to be the units of analysis, or ‘chunks’. And in fact the most important aspect of cognition is just this: what do people take to be the essential units in a given instance?

## **The Classical Theory of Concepts**

The Psychology of Concepts needs to be understood in terms of its own origins myth. According to Cognitive Psychology, from antiquity up to about 1970, the psychology of

concepts was generally ruled by what is called the “Classical Theory of Concepts” and the contemporary Psychology of Concepts originated from the collapse of the Classical Theory. \*

The Classical Theory is the theory of concepts which is embodied in a simple dictionary, that is, that every concept is given by its definition. A word is taken as the sign for a concept, and the meaning of each word is explained in terms of other words. In their interpretation, the definitions of the Classical Theory specify the necessary and sufficient features of any thing coming under the definition, potentially arranging words in a hierarchical taxonomy. So according to the Classical Theory, when a person uses a concept, they do so having the dictionary definition of the concept in mind, and likewise, interpret their perceptual field by reference to this internal dictionary. The question was just “How?”

The first problem with this as an approach to a psychology of concepts is that an internal dictionary may be fine in so far as you are already fluent in the language and have acquired all its basic concepts and your internal dictionary is filed away in your brain. Two questions immediately arise. Firstly, the underlying ‘atomic’ concepts in terms of which other words are supposed to be explained are what are really of interest in a psychology of concepts, but it is these which are left unexplained, or consigned to an infinite regress. The Psychology of Concepts supposes that ‘composite concepts’ can be composed through the intersection or union of these ‘atomic’ concepts. But no dictionary, whether psychic or material, works like that. A dictionary structured like Linnaeus’ *Systema Naturae* is a convenient fiction. The claim that a concept is given by its definition is not at all the same as the claim that a concept is given by its position in some hierarchical *Systema Artificiae et Naturae*.

But secondly, a look into dictionaries soon makes it evident that dictionaries which simply provide definitions give the reader no real idea of the concept. Only those vast works like the 20-volume Oxford English Dictionary which give multiple meanings replete with quotations, and etymology can bring you close to a real understanding of the concept and such dictionaries do not suggest a hierarchical categorisation scheme.

Eleanor Rosch (Murphy 2004, p. 16) discovered that people who knew a concept and were quite definite in classifying objects under a concept, when asked, were often unable to define the concept, or did so inconsistently. At any rate, people could categorise objects under a concept far more quickly and reliably than they could give a verbal definition of the concept. Moreover, it was found that different people at different times would give different definitions, and experts (especially) would fail to agree on how to define concepts which people readily applied in categorisation tests. Few real-life definitions could be found that lacked internal contradiction, inconsistency, ambiguity and ‘fuzzy edges’. This result served to put an end to an idea that people consulted some kind of internal dictionary or look-up table in order to solve categorisation tasks. Thus, whatever the classical theory represented, it did not represent the psychological reality of giving verbal responses to categorisation tasks. The tests sufficed to produce this negative result, but the failure in producing formal dictionary-like definitions of a word, or completing any such formal literary task, in a laboratory setting outside the context of everyday life, could never have given any real insight into the subject’s ability to use concepts. It only proved that people’s knowledge of and facility with concepts is *not* mediated by an internal dictionary. Nonetheless, the results challenge us to explain exactly *how* people *do* solve categorisation tasks, and beyond that, how people acquire, recall and use true concepts, something quite different from solving categorisation tasks in a laboratory. Despite all the attendant confusion in the conception of the ‘Classical Theory’, this is a salutary observation.

The inability of subjects to provide consistent, stable and clear definitions replicating a systematic taxonomy, though, is not a problem of psychology. It is in the nature of *the concepts themselves*. Or I could say, the problem lies in the object, not the subject. As Gregory Murphy points out, it is well-known to lawmakers and those in the judiciary that it is impossible to frame a law that will not sooner or later run into ambiguities or self-contradiction, and as laws are subject to endless revision and interpretation, a time never comes when that ambiguity disappears. When lawmakers and judges set down the principle of justice that they intend, no amount of definition of terms, qualification and explanation can

---

\* I shall not examine the “Classical Theory of Concepts” as a current in the history of philosophy at all. I am concerned only with the role the idea plays as a foil for the Psychology of Concepts. When I look at the history of philosophy in Part 2, I will not touch on this idea either.

reliably represent their concept. Tests (Margolis & Laurence 1999, p. 444) involving novices and experts in the sciences showed that the concepts of experts were fuzzier than those who actually knew nothing about the topic; the more developed the concept, the fuzzier the boundaries. All this goes to show that there is more to any concept worthy of the name than can be set down in a few dead words. This is not a problem of psychology, it is *in the nature of concepts themselves*. Concepts are not pigeonholes and concepts which conformed to expectations of these researchers would be very poor concepts.

The reference of a concept to a dictionary definition is but one aspect of a concept, and can only be tested by means tailored to that task. Laboratory categorisation tasks are not suitable means for exploring this aspect of concepts. Over and above this, dictionary definitions are rarely set-theoretical catalogues of the contingent attributes of things. To believe otherwise is to misunderstand the so-called Classical Theory of Concepts, insofar as any such a theory ever existed in psychology.

Also, were we to accept that as thought-forms, concepts can be entirely accounted for within psychology, then there could be no logic, no mathematics, no science, and in fact life would be nothing more than some kind of solipsistic dream. Concepts have a content which is objective, and insofar as concepts reflect the material world, they will be inconsistent, unstable and contradictory.

When Kant (2007) showed that the “Antinomies of Reason” – concepts such as free will, the infinite divisibility of matter and the boundedness of the universe – fell into contradiction with themselves, he claimed that this was because the concepts went beyond the bounds of possible experience and were therefore meaningless. Hegel (1816/1969) responded however by showing that, not just four, but *all* concepts eventually fall into self-contradiction. Hegel showed that “There is absolutely nothing whatever in which we cannot point to contradictions or opposite attributes” (1830/2009 §89). Concepts inevitably overflow their own boundaries, but it is not really a problem with definitions, for it is in the nature of concepts themselves to fall into self-contradiction. “Zeno, who first showed the contradiction native to motion, concluded that there is no motion (1830/2009 §89). But Zeno was clearly mistaken: motion is real, but it is contradictory nonetheless.

So, the fault of the Classical Theory here is only that it reproduces the very ‘problem’ which is inherent in concepts themselves. Hegel argued against Kant that such contradictions should not be regarded as failures, but rather that they simply express the vitality of concepts. Consequently, any attempt to eliminate this falling into contradiction and overflowing of its own limits, can only serve to kill the concept, which is exactly what a dogmatic insistence on rigid, unambiguous definitions achieves. But on the other hand, if instead of recognising this tendency to overshoot their own boundaries as something in the nature of concepts themselves, we seek to avoid such contradictions by abandoning definitions and steering away from the classical theory of the psychology of concepts, we shall inevitably *miss the concept* altogether. Either way, we will fail to understand concepts. This was the result when recent investigators in the Psychology of Concepts wrongly interpreted the failure of *their interpretation* of the Classical Theory as a problem of *psychology*. Nevertheless, by testing out the psychological reality of the Classical Theory using laboratory methods, the modern Psychology of Concepts produced results which are of interest in their own right.

One of the problems with the cognitivists’ critique of the Classical Theory was that they took it for granted that the conceptual world adheres to a hierarchical taxonomy in which things can be defined according to a list of attributes, which conjointly determine whether any given object is in or out of the category. In the light of the review of concepts I gave in the introduction, this is an excessively restrictive assumption. It is particularly suited to laboratory work though, despite focussing attention on boundary problems rather than the core or essential issues. Set Theory propositions of the form “all *x* are *a* & no *x* are *b* etc., etc.” are taken to be how the mind holds knowledge about the world. When the Classical Theory lost support among psychologists, they still hung on to similar conceptions such as ‘Fuzzy Set Theory’. Even visual images were taken to be a means of encoding of look-up tables of features. The abandonment of the Classical Theory without any revision of the atomistic, dualist and abstract world-view underpinning the cognitivist interpretation of the Theory, meant that they turned their backs once and for all on any theory which could cope with true concepts such as those listed in the Introduction.

If we let go of the idea that Cognitive Psychology is some variety of Psychology, and instead regard Cognitive Psychology as a branch of Engineering Science, then all this makes abundant sense. But then surely Cognitive Psychology loses its very *raison d'être* if it stops paying attention to what is distinctively human and unlike a machine in human behaviour? If so, then Cognitive Psychology needs to learn more from the humanities, rather than simply ignoring the very characteristics that they want to investigate.

### Some Reflections on Aristotle

Before moving on to the theories which Cognitive Psychology produced from its critique of the Classical Theory, it is worth mentioning Aristotle's view of concepts, because Aristotle is often cited (Murphy 2002, p. 39) as a proponent of the classical theory of the psychology of concepts. Such a claim would be absurd because Aristotle did not have a theory of psychology in the modern sense; his concern was how the world itself was constituted, and he took it for granted that the structure of the mind was the structure of the world as grasped in language. So Aristotle's theory of concepts took the form of an examination of the categories of things that exist in language and the world. According to Aristotle, the world had subjects (ultimately substances) and predicates; the predicates were what could be said of subjects. The subjects themselves were irreducibly just what they were, whilst the various predicates which could be said of them were contingent attributes. He thus made from the outset a clear distinction between a subject and those attributes which may or may not be attached to the subject, without affecting what it is, its *essence*. What is required of a theory of concepts is precisely this, to see past the contingent attributes of a thing, to get to its essence, to the *concept* of thing. And it is precisely this which is systematically ignored in the Psychology of Concepts. According to Cognitive Psychology, when all the attributes of something are taken away, there is nothing left. This accords with many contemporary philosophies, such as post-structuralism, but not with science and not with Aristotle.

Talk of 'essence' may be troubling for people who are accustomed to taking it that things are exhaustively determined by their attributes, but this discomfort is really unwarranted. According to the cognitive model in question, the essence of a concept would be some combination of attributes which are deemed to be core or non-core, i.e., essential or inessential, possibly including its place within some social relation, or social practice. Every real thing has contingent attributes which are inessential to it being what it is, and on the other hand, something which makes it what it is, and this essence, in the great majority of non-trivial cases, is not given in the perceptual field.

The philosopher, Ruth Millikan (Margolis & Laurence 1999, p. 525), has an interesting take on Aristotle's theory. A 'substance' is a fundamental component of the world, and according to Aristotle, as a citizen of ancient Greece, the substances were individual things (e.g. Mama), natural kinds and kinds of stuff (e.g., dogs and water). These are the various things which are irreducibly what they are and which "something can be said of." This apparently rather chaotic collection of categories is also a fair representation of the ontology of the infant mind. On the basis of her interpretation of Aristotle's categories, Millikan makes a critique of the dominant psychology of concepts for which concepts are nothing but contingent attributes. Aristotle also counted as substances the 'natures' or 'essences', which he took to be 'that towards which a thing moves', and it seems odd that Millikan has passed over this third kind of substance, since this is more general than her own explanation of the nature of things. Millikan holds that concepts are not constituted by the bundle of attributes we ascribe to them, which are simply means of recognising them, whilst the real ground of a concept is "knowing how to use a thing" (Margolis & Laurence 1999, p. 530). Whilst this capacity may be acquired socially, her approach is otherwise rather strongly biological. Millikan's is a useful critique of the cognitivist interpretation of the Classical Theory.

The Psychology of Concepts should be credited with proving that definitions of concepts have little or no psychological reality in people's immediate interactions. Accessing the lexical definition of a concept is a higher order psychological function and the reproduction of such a definition is a task of a yet higher order. A definition expresses the place of a concept in an infinite semantic network, not a hierarchical taxonomy, and just like the real world, concepts and their definitions are mobile, unstable and internally contradictory. There is every reason to think that such dynamic semantic networks have a place in the understanding of concepts, and must therefore be indispensable to understanding their psychology.

## Prototypes, Exemplars and Ideals

Following up their critique of the Classical Theory, Eleanor Rosch and Carolyn Mervis proposed what is known as the Prototype Theory. Whereas the Classical Theory adopted a model of categorisation along the lines of a Linnaean taxonomy, the Prototype Theory uses categorisation theories based on a typology of distinct clusters of concepts rather than a hierarchical taxonomy. Here it is presumed that a person has in their mind, not a universal definition, but a mental representation (i.e., concept) of an *individual* thing which may be most typical or ideal or perhaps the original *individual* representative of a category of things. All other things are included or not included under the concept according to how much they *resemble* the prototype. This idea has the apparent merit that it allows for the concept to be held in the mind in just the way any individual thing given to the senses is apprehended. There is good reason to argue that all recognition entails individual phenomena like this in some way, rather than noumena. It also explains the observation, which was inexplicable under the cognitivist interpretation of the Classical Theory, that some things were *more typical* of the category than others, as well as borderline cases and so on. These phenomena of typicality can be better reconciled with a model in which a category is defined by a single instance, than in the supposedly all-or-nothing criteria of a lexical definition.

Rosch *et al* were able to demonstrate that this approach had psychological reality in the sense that the so-called Classical Theory did not. Like any good scientific theory, the Prototype Theory immediately opened up a lively field of research, as people searched for prototypes for different concepts under different test conditions, mapped degrees of typicality in terms of attributes, core and non-core attributes, and so on, especially under the conditions of *instant response*, rather than reflective thought. Instant responses are deemed to demonstrate what is 'really on your mind', whereas once a person begins to think about the problem, the question becomes more complicated, and incidentally, of less commercial interest. Whichever way we go here, it seems that instant responses to either categorisation or definitional tasks manifest a psychological function which differs from reflective responses, and differs again from the socially mediated responses which constitute the real life of concepts.

It very soon emerged however that the typicality phenomena pointed not to one prototype in most cases, but a family of *exemplars*, and thus emerged the Exemplar Theory. The Exemplar Theory uses a model of categorisation somewhat like notions of genre. The field of phenomena falling under a concept, the concept's 'extension', contains a number of exemplars which may be of quite different kinds and bear no particular systematic relation to one another (on the contrary in fact), but together mark out a field. With time and experience the number and variety of exemplars which are recognised under the concept grows and the person's knowledge exhibits corresponding maturity and stability.

George Lakoff (Margolis & Laurence 1999, p. 398ff) has an interesting take on the idea of exemplars and prototypes. Rather than seeing them as bundles of common attributes, Lakoff sees prototypes and exemplars as 'cognitive models', illustrating this with a variety of definitions of motherhood (biological, nurturing, adoptive, genetic, protecting, etc.) each with its own *ideal* of what it means to be a mother. Here the same universal symbol is associated with different 'definitions' and correspondingly different expected bundles of attributes as exemplars.

This is not dissimilar to an interesting phenomenon which turned up in work on prototypes and exemplars: that subjects often had an *ideal* of the concept which did not pass as a typical instance of the concept, but rather one which was ideal in terms of the person's professional or otherwise practical relation to the concept (Medin et al, 2007). Objects would then be judged better or worse instances of a concept according to how many of the features of the ideal they exhibit. This implies that the features which exemplars manifest do not reflect typicality in an abstract, contemplative sense, but rather practical, teleological relations to the object, relations that were implicated in some *system of practice*. Where no such motivated relation exists, then the ideal would become by default, a vague and unstable Galton composite of some kind.

Another interesting observation arising from this work is the idea of the 'basic level'. In the hierarchy of generalisation from most particular to the most general (for example: Featherston, armchair, chair, seat, furniture, goods) there is a level of generalisation which is "the highest level at which a single mental image can represent the whole category, and the level at which most of our knowledge is organised" (Lakoff & Johnson 1999, p. 23). I remember a comedian

once explaining how *never* to use the basic level in referring to common objects: refer to your car by its make and model, where you are by the actual street name, what you're eating by a brand name and visual description, etc. The comic effect arises from the resulting cognitive dissonance.

All these ideas produced fairly robust experimental results, but do they say anything at all about concepts? It seems that they tell us about how people recognise an object as of this or that kind before they have had a chance to think about it, a problem which the so-called Classical Theory was not in a position to explain. But it is also quite possible that such representations also play a role in reflective and incidental thought. They also make sense of the phenomena of typicality, which in the absence of any understanding of concepts other than that offered by boundary conditions also seemed inexplicable. But it remains at a level suitable for displaying goods on supermarket shelves or programming robots, but unsuitable for understanding human cognition. The Pope is the Pope perhaps because he was elected by the Cardinals, because he is God's representative on Earth or because he is head of the Catholic Church, but *not* because he wears that big hat by which you can pick him out in a crowd. Whilst escaping from the logical difficulties posed by the cognitivist interpretation of the Classical Theory, cognitive psychology has lost the concept of concept altogether, building instead a theory of object-categorisation. *Not the same thing*. A child can correctly sort the pieces of a chess set into groups without the slightest understanding of the concept of chess or its various pieces. Further, the underlying assumptions have remained the same: a world of psychic images mirroring a world of things through phenomenal similarity based on bundles of features.

### **Theory Theory and Semantic Networks**

A more recent theory which has been produced by this current of Cognitive Psychology is what is called the Theory Theory. A concept, it is said, draws meaning from the context of a larger theory of which it is a part. This opens up the possibility to get away from uncritical dualism, in that a concept is now "picking out" objects from a field of perception which may be already conditioned by a theory which is mediating perception. This not only brings us closer to Kant, but also brings into play the work of people like Thomas Kuhn, Edwin Hutchins and Bruno Latour who have studied the sociology of knowledge. But it is by no means necessary to presume that a person has a 'theory' in the sense of a theory of natural science. Every person perceives the world through the lens of their own, personal ideology, so to speak, which conditions their expectations for what they expect to find in any given situation. Social theory reflects social practice. Further, following Stephen Toulmin (1972), I would say that we do not need to follow advocates of the Theory Theory in positing a comprehensive theory of everything in a person's psyche. It is enough that a person has some conception relating some phenomenon to a larger set of phenomena, or can create one when confronted with a novelty. As diSessa (2006) has pointed out, people characteristically understand the world through an eclectic mixture of 'theories' even if in their own professional work they use an institutional system of concepts. This opens the possibility of approaching a concept 'from above', so to speak, from the system of which it is a part, as an alternative to the former approach 'from below', by building the concept up from contingent attributes or bundles of features.

From this promising start however, Cognitive Psychology reverts to type. A subject recognises an object according to its attributes, but the attributes it expects for objects of a certain kind are given by the relevant theory rather than some other theory, that's all. A person recognises the little fluffy birds following the big duck around as baby ducks, even though they've never seen ducklings before, because they know about creatures having young which are smaller and cuter, etc. So again we miss the concept, because the problem of a concept of duckling is replaced by features and categorisation rules. Nonetheless, the introduction of the idea of a category depending on the use of a more comprehensive theory of the objects and categories in the field is a step towards mediation of categorisation tasks, and consequently a step towards a real study of concepts. It is to be hoped that further work in the field will begin to ask questions about where theories come from and how are they acquired and sustained? what mental form does a theory take? why is one theory rather than another evoked and brought to bear in some circumstances but not in others? how does a theory suggest attributes? and so on. The Theory Theory does answer some questions, but it asks more questions than it answers.

Implicit in the psychology of concepts as developed by Cognitive Psychology is that perception of any object is necessarily the sum of perceptions of simpler features. However, evidence to the contrary comes from Gestalt Psychology, numerous optical illusion experiments, and phenomena such as our ability to read past typographical errors and even absurd jumbling and distortion of the letters of words. Evidence also comes from child development which shows that infants can perceive only general impressions and the ability to differentiate objects and from their background is only gradually achieved as the child develops. And even if we lay to one side the mistake of taking the perceptual field as the sum of “pixels,” it remains the case that problems of perception still do not go to the psychology of concepts properly so called. The overwhelming majority of results reflect responses people have given in the laboratory to tests using word lists. This is a very limited domain of activity, albeit one appropriate to life in a post-industrial bureaucratic/capitalist society.

The contribution of the Theory Theory was to reflect the fact that we recognise something because we have an *expectation* for it before we discern any of its features. The idea of having a theory of a process which would allow you to fill in a blank is not restricted to what would normally be seen as a ‘theory’. *Connectionism* is the idea that concepts are nodes in an extended semantic network and the use of any concept ‘activates’ other concepts which are ‘near’ to it on this network. For example, the mention of ‘kitchen’ activates relevant concepts so that when we hear the word ‘knife’ we think of a kitchen knife not a hunting knife. Such a *semantic network* is a way of visualising a theory at the simplest possible level, which explains how a concept can be approached ‘from above’, rather than ‘below’ by adding up its perceptual features.

### **What has been Learnt from Analytical Approaches**

Despite the conceptual problems with the mainstream Psychology of Concepts, their ideas have been the framework within which a great deal of experimental work has been done. Many of the results are of interest.

The critique of the Classical Theory, for example, established that in the case of everyday concepts, definitions have no psychological reality; people do not ‘consult’ lexical definitions when they use such everyday concepts, and usually cannot even produce a rigorous definition given time for reflection. It seems that such concepts are generally learnt and understood in use, not from a lexical definition. As Rey (Margolis & Laurence, 1999, p. 296) pointed out, the most familiar example of this fact is that native speakers of a language may use it perfectly with no knowledge of the rules of grammar. But we still have no reason to believe that academic concepts like ‘tensor’ or ‘prolepsis’ are used without first acquiring a definition. In any formal institutional setting, it is far more likely that use and acquisition of a concept entails definitions, than would be the case with everyday concepts which are originally acquired by spontaneously conforming to semantic norms while interacting with others. Learning to define concepts is a high level psychological task, generally demanding language skills not called upon by use of the concept. It is not necessarily the case that everyday concepts and technical concepts simply lie on opposite ends of a continuum of some kind. Perhaps they are of qualitatively different kinds, or are pure types of which all concepts are hybrids in some way? Or perhaps the distinction simply reflects the social and institutional context?

The observation that arose from the cognitivist critique of what they call the ‘Classical Theory of Concepts’, that concepts in general do not have clear and unambiguous definitions, and that any concept sooner or later falls into contradiction with itself, coming from *these* writers, firmly situated in the analytical tradition of philosophy, is immensely helpful. Even though 50 years ago Stephen Toulmin (1953) had told us that Formal Logic did not reflect how science worked, and Hegel had told us 200 years ago that all concepts inevitably fall in to contradiction, there is considerable resistance to this idea within the analytical tradition. But it has not stopped the cognitivists from taking formal symbolic logic (which is blind to the internal contradictoriness of concepts) as the gold standard for reason both as a psychological function and as a research method. The cognitivists took the observation that people do not adhere to the laws of symbolic logic in their thinking as a *psychological* discovery, rather than as reflecting an objective limitation on the scope of symbolic logic.

Even if the Prototype and Exemplar Theories have yet to provide an adequate description of concepts and their typicality effects, the evidence that prototypes and exemplars provide a

good model at least for *reflex categorisation* is useful. It seems that the mainstream Psychology of Concepts is moving inexorably towards some variety of the Theory-Theory, some 'top-down' conception of how the mind holds an array of concepts which are activated by sensuous-practical interactions. However, so long as it is tied to the model of a concept as a bundle of attributes, even with core/periphery distinctions, it can make little progress. The observation that prototypes and exemplars may represent a practical *ideal* for people, according to a special interest they may have in some field of activity is of particular note, since it introduces into the process of concept formation a person's participation in relevant forms of social *practice*, as opposed to exposure to experience.

The problem of getting beyond 'descriptivism' is posed quite well with the problem of gender identification: Medin et al introduce the idea of "psychological essentialism,"

People act as if things (e.g., objects [and people]) have essences or underlying natures that make them the thing that they are. ... For example, people in our culture believe that the categories male and female are genetically determined, but to pick someone out as male or female we rely on characteristics such as hair length, height, facial hair, and clothing that represent a mixture of secondary sexual characteristics and cultural conventions (Medin et al 2007, p. 8).

This problem is a rock on which many theories of society and psychology have broken. Even if there is no final or unambiguous essence, even if we have to take essence as a process rather than an entity, *the distinction between essence and appearance is indispensable* for a theory of concepts. We need a theory of essence as a socially constructed process, rather than a metaphysical entity as it had been for Aristotle. To abandon any concept of essence in favour of phenomenalism leads to absurdity. What after all is the difference between a Van Gogh and a very good fake?

### **The Problem with the Analytical view**

The idea of cognitive psychology is that the human mind can be modelled as an information processing machine: the mind acts *as if* it were executing an information processing function.

In itself this is an undeniably powerful approach to understanding the psyche. However, it is the baggage which comes with the model which undermines the potential benefits of this approach. Machines do not exercise free will, lust after others, experience loyalty, guilt, hatred or fear, have intentional dispositions or friends or *understand* what they are doing. So in order to model some human function as a machine function it is first necessary to *construe* the given function as something which a machine could do. Human functions which cannot be modelled by machines, such as free will, may even be deemed to be illusions to legitimate the approach. Consequently, what is modelled is never a human function, but rather an outwardly humanlike machine function.

Secondly, the main function of the mind is for human beings to manage their own relation to the world, but since machines cannot do this, the operator manages the machine-world interface for it. Among other things this means that every internal operation has to begin from an input which the operator controls. Thus we have the absurd and unjustified dogma that every operation of the mind begins from arbitrarily small "chunks" of information. There is no foundation for this in human biology, let alone psychology.

Many cognitive psychologists would acknowledge that people can massively improve psychological functions by the use of cues from the world around them. For example, memory can massively exceed its natural limits if the subject is presented with some system of reminders connecting the stimuli. But it never seems to occur to them that these systems are an integral part of every person's normal cultural environment and that the psychological functions which people exhibit in real life reflect the cues embedded in their cultural environment rather than the underlying natural functions.

The result is a guide as to how a machine could be programmed to do outwardly humanlike things. The knowledge of psychology exhibited by cognitivists is often naïve. For example, introducing the section on knowledge, Lawrence Barsalou characterises education as the "teachers provide information that students incorporate into existing knowledge" (1992, p. 152) apparently blissfully unaware of all the problems which torture the minds of educational psychologists. This is not dissimilar to the knowledge of psychology exhibited by some neuroscientists. The workings of the mind are outside their subject area. And the same goes for

most of the cognitivists' contributions to linguistics, pragmatics, anthropology, political science and so on: the degree to which a machine can participate in the relevant practices is extremely trivial, and observations about the topic based on the machine metaphor are correspondingly trivial. Nonetheless, the idea of modelling the mind as an information processor remains a worthwhile project for engineering purposes.

It has been necessary to spend some time examining the psychology of concepts as developed from the standpoint of analytical science partly because this view corresponds broadly to educated 'mainstream' consciousness in the English-speaking world. The view to be developed here, on the other hand, may be rather confronting from this point of view.

To give experimental subjects a series of categorisation tasks, ticking boxes on a survey form: "Is a car seat furniture? yes/no. Are curtains furniture? yes/no," tells us nothing about the subjects' concepts, but a lot about the concepts of the researchers and their mode of activity. The idea of seeing the world, whether psychic or material, as made up of discrete elements, which have only to be sorted into groups and counted in order to make a decision, is deeply engrained in contemporary culture. Even our system of government is organised along these lines. Any other approach may have difficulty finding a place in today's academy, especially if it cannot be fitted into just one department. The fragmentation of science has gone way beyond division of labour, since division of labour presupposes at least some form of cooperation, exchange and shared objectives. The branches of science today have developed such distinct views of the world that they do not even ask questions, the answers to which would be of interest to those outside the discipline. Universities are organised along lines resembling Set Theory with an intellectual life that increasingly resembles niche marketing.

The Psychology of Concepts, as it stands, has collected data about how people recognise things and categorise them, but it has no theory at all about concepts. As Medin et al (2007, p. 18) point out, even if people pick out the same entities with the same category, this does not at all mean they have the same concept, as the sense might differ widely while reference remains constant. Categorisation is not conceptualisation. In first reducing the Classical Theory to a view of concepts as a catalogue of 'features' and then dismissing it, this current of thought has lost track of concepts altogether and deals only with responses to artificial, laboratory-based, categorisation tasks. As Smith & Medin put it:

categorization theories relying exclusively on similarity relations are insufficient to provide a theory of concepts. We have argued that a coherent concept is one that we have a good theory about and that fits well with our other knowledge. ... Future research on concepts and categories can help answer these questions [about concept acquisition] not by controlling the effects of world knowledge and experience, but by exploiting them – by bringing the concepts into contact with the whole cognitive system that created them (Margolis & Laurence, 1999, p. 455).

The fact that I can quote these same writers in making these points is evidence that these problems are indeed immanent in the practice of this science. But how can they be resolved?

Lakoff and Johnson made an interesting observation about what properties of a gun make it a fake gun (Lakoff & Johnson 1980, p. 121). They determined that all the perceptual properties (what it looks and feels like), the motor-activity properties (how it may be handled) and even the 'purposive properties' (what you can use it for) can and even should be preserved in a 'fake'. 'Functional' properties (*how* it effects its use) and its 'history' (what it was *originally* made as), however, must be negated for it to be a fake. This points to a very subtle distinction in the properties of an object which inhere in its concept.

If we are going to leave the laboratory and study concepts in real life, then we need a theory oriented to social life. The fact is that different people see the world in quite different ways, and this is tied up with the social activity they are involved in and the culture they have acquired. And we can't look into their head and study their concepts in an MRI machine. Thinking has to be studied in the situations in which it is realised. Laboratory experiments produce laboratory results, but in the study of human behaviour and thinking, it is unlikely that what happens in the laboratory will tell us much about what happens in life outside the laboratory. The world is not a word list.

In summary, a concept is not equivalent to a normative list of features. There is a distinction, albeit not absolute, between phenomenon and essence. Perception and the performance of

literary tasks entail distinct psychological functions. Laboratory tasks are practices in their own right and do not 'model' other forms of practice. In fact, if 'concept' is to be a meaningful concept, we have to allow that the same concept manifests itself quite differently according to the practical conditions of its realisation. One and the same concept may be realised as a reflex response to a recognition task reflecting an ideal image, or as a lexical definition in a written essay.

Further, the methods which work very well for the scientific study of material objects do not work so well when the objects under study are thought-objects, if indeed 'thought-object' is a meaningful concept at all. You just cannot succeed in a science of thought-forms while ignoring the results of philosophy, which anticipated the problems experienced by the Psychology of Concepts by at least two hundred years, although without the benefit of experimental methods.

## Analysis

Is it possible to counter the deep-seated conviction of cognitive psychologists that images can only be composed by aggregation of 'pixels'? The pervasive influence of the digital image in our lives seems to reinforce this idea, but have you noticed that when there is some disturbance in transmission and the digital image momentarily breaks into much larger pixels? This is because digital images are not stored, transmitted or created one pixel after another, but rather by a 'successive approximation' process which divides larger pixels into smaller ones.

Consider how the ear works. The brain does not receive a series of air pressure measurements, but pulses from sensitive follicles along the length of a little cantilever in the inner ear, which respond to the vibration of the cantilever. This effectively makes a running Fourier transform of the sound pressure variation, which is more like a recording of the movement of the keys a piano keyboard, than the vibratory movement of the strings it controls.

Consider how you recognise a strange person or object. At first, probably based on a couple of cues, you *misrecognise* it, and see a whole image, but the wrong one. Then you pick up a discrepancy and correct the image, thus approaching a true image of the whole by successive correction of the *whole images* of things you already know.

So there is no basis in nature, psychology or technology for the prejudice that perception is atomistic.

## The Sociocultural Turn

A number of cognitive psychologists picked up on the sociocultural turn. Edwin Hutchins showed how complicated reasoning processes are accomplished differently using the resources of different cultures, and how in fact problem-solving tasks may be distributed amongst a group of people cooperating thanks to the use of technology. However, Hutchins is very concerned with analysing reasoning processes down to lines of computer source code, as if understanding a reasoning process meant being able to reconstruct it in software. Consequently, he tends to overlook or take for granted those nodal points through which all human reasoning must pass: concepts. The signs for the concepts of the culture then appear as *shorthands* for various combinations of atomistic sense impressions. Machines, after all, perform input-output tasks, and it is the job of the machine operators to decide what to input. It is remarkable that in his study of reasoning among the much-studied Trobriand Islanders, he went to great lengths to disclose the rational processes underlying formal discourse around property disputes, but gave no significance at all to the complex and subtle range of concepts which are entailed in Trobriand property relations. Nonetheless, Hutchins' work on distributed cognition and how people create and use artefacts to organise the collaborative activity which is normalised within a community, did open a window on the need to understand tools and symbols indigenous to a culture as having an important place in cognition, and many other things besides.

Bruno Latour, who has disclosed the cultural assumptions and practices underlying natural scientific ideas, also regards concepts as a kind of shorthand for nested subroutines linking atomistic sensations and actions to concepts properly so called. These writers are reminiscent of the philosopher Rudolf Carnap who also regarded a range of philosophical problems as solved by demonstrating that all concepts can be reduced to combinations of atomistic sense impressions, using mathematical formulae to make the device look plausible.

I very much doubt that ideas of concepts and words as shorthand for long chains of atomistic percepts and set-theoretical relations has any more psychological reality than dictionary definitions. What does it prove to reduce a concept to a chain of atomistic percepts rather than a series of successive approximations to the whole? Though such ideas do not reflect a psychic reality, they may be useful for programming machines. But even then, the concept of nested subroutines is an already-outdated concept of how computer and information technology works. The general structure of information technology is better represented by a large number of successive layers of systems each of which operate as a self-contained, stable whole, each with its own language and substrate, while passing information about their state to layers above and below. The metaphor of nested subroutines became obsolete with the advent of object-oriented programming in the early 1960s.

It is not the analysis of complex processes into simpler units which is the problem – such an analysis is always necessary in approaching a concrete situation from outside – but that *no thought is given as to what the unit of analysis may be*. In fact, a simple concept is itself a unit, and the problem is not to break a simple concept into parts (which always destroys the concept) but rather to determine *which* concept is analytically primary. For example, why should we assume that each of the letters in the word “rabbit” is a conceptual unit relevant to reading a line of printed text, with the word itself being regarded as a composite? We could equally well say that each black dot in each letter is the conceptual unit or alternatively that the word, once learnt, is a conceptual unit from which it is possible to proceed to spell out its letters. This is a question which can be resolved experimentally, but all too often it is unwittingly decided in advance by the choice of experimental procedure.

Psychologists study individual actions, whilst sociologists study group behaviour; linguists, on the other hand, study language, the paragon of all those constellations of artefacts with which we organise our thoughts and behaviour, leaving the study of other aspects of material culture to still other disciplines.

Actually, concepts exist only through the correlation of all these domains, and can only be understood through at least a study of psychology, social theory and linguistics, informed by a knowledge of philosophy.

## **Links**

[Concepts. A Critical Approach. Table of Contents](#)

[Andy Blunden's Home Page](#)