

[Paper presented at the Course on 'Meaning and Natural Kinds', Inter university Post graduate Centre, Dubrovnik, Yugo-slavia, 15-25 September 1986.
Revised by the author in 1993.]

CAN A PSYCHOLOGIST MAKE SENSE OF NATURAL KIND TERMS
AS RIGID DESIGNATORS?

U.T. PLACE

As those of you who have attended the last two of these Dubrovnik philosophy courses will know, it is my role on these occasions to introduce a note of dissent into what is otherwise inclined to degenerate into a straightforward exposition of a particular philosophical standpoint. My contribution this year will, I suspect, be no exception to that rule. For reasons which will come as no surprise to those familiar with my approach to the mind-body problem, there is one part of the argument developed by Kripke in his 'Naming and necessity' paper with which I am profoundly out of sympathy. This is his discussion at the end of the paper on the intuition which he himself attributes to Descartes, but which is commonly referred to by others, e.g. McGinn (1977), as "Kripke's intuition". This is the intuition that, as Kripke puts it, whereas God could not have created heat without the heat he created being molecular motion, he could have created pain without the pain having to consist in the firing of C-fibres or whatever physiological event pain is found to be correlated with. Quite apart from the fact that he is using this argument as an objection to the kind of type-identity version of the mind/brain identity thesis to which I have been committed for more than thirty years, I find this appeal to an intuition of the contingency of the mental/cerebral relationship particularly galling in view of the fact that it was precisely this argument from the contingency of the mental process/brain process relation which my 1956 suggestion that it is a reasonable scientific hypothesis that conscious experiences are processes in the brain was designed to meet.

Of course, the necessity/contingency distinction I had in mind in my 1956 paper was what in Kripke's terms would be called necessity/contingency de dicto, whereas what he claims to be talking about is necessity/contingency de re. When it is interpreted as a necessity/contingency de dicto, the contrast between the necessity of the identity of heat and molecular motion in contrast to the contingency of the mental process/brain process relation can be

explained by invoking Sir William Hamilton's distinction between the "intension" and "extension" of a term. By the "intension" or, to use Frege's term, the "sense" or Sinn of a term or expression is meant the criteria we employ when we accept an instance as an instance to which the term or expression in question applies. By the "extension" of the term is meant the class of instances which are accepted as falling under the term or expression in question by virtue of satisfying the criteria in which its sense or intension consists. In terms of this distinction the heat/molecular motion and mental process/brain process cases are both cases in which two expressions either have (in the heat/molecular motion case) or are assumed to have (in the mental process/brain process case) the same extension in the sense that as a matter of fact every case of heat is a case of molecular motion and vice versa and every mental process is some kind of brain process, though in this case not vice versa. The fact that there are some brain processes that are not mental processes is simply a reflection of our ignorance of what kinds of brain processes mental processes actually are. To use an expression like the firing of C-fibres in this connection is simply a cloak for ignorance.

Apart from the ignorance in the mental/cerebral case, though in fact, as I shall argue, arising out of it, the important difference between the two cases is that in the heat/molecular motion case the intension of the term "heat", as used by physicists and by those whose thought on such matters is influenced by their knowledge of physics, includes the notion that heat is molecular motion, whereas the intension of our ordinary notion of a mental process does not, as things stand, include the notion that a mental process is some kind of brain process. What this means is that if a physicist were to encounter a case of apparent heat which, it so turned out, did not in fact involve any kind of molecular motion, it would be rejected as not being a genuine case of heat. Whereas in the mental process/brain process case, we would not, as things stand, reject as a genuine case of mental activity a reported experience which, it turned out, was unaccompanied by any kind of brain activity.

I emphasise "as things stand" here, because it seems to me, as it did more than thirty years ago when I wrote 'Is consciousness a brain process?' that the intension or sense of our words and expressions as reflected in our linguistic intuitions as to what is and is not contradictory to suppose, and hence what is necessary or contingent in the de dicto sense, is extremely sensitive to the state of our knowledge. I illustrated this point in 'Is consciousness a brain process?' by the example of the case of the table which consists of an old packing case turned upside down. Here we have a typical and I should have thought uncontroversial case of contingent token-identity, in which two descriptions "a

table" and "an old packing case" have, as things stand, quite distinct and non-overlapping senses or intensions, but whose otherwise distinct extensions happen to overlap in at least this one case to which both descriptions happen to apply.

I then went on to consider what would happen if we lived in a world in which, instead of being an exception, the only kind of table that existed was a table made from an old packing case turned upside down. I suggested - and as far as I know no-one has come forward to refute that suggestion - that in such a world the sense or intension of the expressions "table" and "old packing case" would no longer be independent of one another. One of the criteria for deciding whether or not a given instance was a table would be whether or not it was an old packing case turned upside down. If it wasn't, it wouldn't be accepted as a table, just as a physicist wouldn't accept a case of heat as genuine if it didn't involve molecular motion.

But this kind of de dicto necessary connection between the intension of two terms depends essentially on their co-extension being a matter of common knowledge. If people didn't know that these things they called "tables" were in fact old packing cases turned upside down, the conceptual connection between the intension of the two terms would not develop. Where the criteria for assigning instances to terms which are in fact coextensive are such that they are never applied simultaneously to the same object, the coextension will be obscured and until it is demonstrated scientifically beyond all doubt, the intensions of the two descriptions will remain distinct and the identity an empirical hypothesis concerning a matter of contingent fact. It was this concealment of coextension by the physical separation of the situations in which the different criteria are applied which I tried to illustrate in 'Is consciousness a brain process?' by the example, first, of the cloud and the mass of water droplets in suspension and, secondly, of lighting and the motion of electric charges.

There is, of course, one important implication of this account which I did not pick up in my 1956 paper, which is that once the co-extension of the two descriptions becomes a matter of common knowledge as a result of scientific research, what begins as a hypothetical contingent identity between two expressions whose factual co-extension is not recognised, becomes, once the co-extension is recognised, a de dicto necessary identity or other analytic connection, such as class inclusion, between the two descriptions.

On this view, the difference between our intuitions in the heat/molecular motion and mental process/brain process cases can be explained completely and, to my mind, entirely satisfactorily by the fact that this change whereby knowledge of co-extension leads to the development of an analytic connection between the intensions

of the two terms or expressions has already taken place in the case of heat and molecular motion, but has not yet taken place in the case of consciousness and the, as yet, unidentified brain process in which it putatively consists.

Kripke's alternative interpretation of this situation in terms of his notion of an a posteriori and de re necessary connection which holds in all possible worlds between heat and molecular motion but not in the case of pain and C-fibre firing seems wholly unsatisfactory by comparison. The notion of an a posteriori and de re necessary connection strikes me as a reactionary move designed to freeze the process of conceptual innovation in science at its present stage by issuing a fiat declaring the present state of conceptual development to be a once and for all time discovery of the real essences of the natural kinds in question. Not only does this rule out the possibility of any change in our conception of the relationship between heat and molecular motion in the future; it also rules out the possibility of any future discovery of the real essence of mental activity in the brain. Indeed the effect of Kripke's account is to render wholly and irremediably mysterious phenomena such as conceptual change and our conceptual intuitions which are quite simply explained on the de dicto theory. How, to use an example I used in my paper on the social relativity of truth two years ago, do we account for the fact that at one time the intensions of the terms "fish" and "mammal" were not, as they are now, mutually exclusive, so that, in terms of the conventions then prevailing, it was correct to classify whales as falling within the extension of both terms. On Kripke's view, as I understand it, the terms "fish" and "whale" qualify as rigid designators of natural kinds whose real essence is a matter of scientific discovery a posteriori. "Mammal" on the other hand is a scientific term of art with a nominal definition - a species of animal of which the female suckles its young by means of mammae or breasts. On this analysis we should have to say that our present a posteriori knowledge of the real essences of whales and fishes shows us that the two kinds are mutually exclusive such that if something is a whale it cannot be a fish and vice versa and that this fact is not unconnected with the fact that the real essence of whales involves being a mammal while the real essence of a fish excludes that characteristic. On this view we should have to say that the Mediaevals who classified whales as fishes, despite the fact that they were known to qualify as mammals, had simply got it wrong. They were deceived by a superficial resemblance of shape and habitat into assimilating two mutually exclusive natural kinds.

Up to this point the story that a Kripkean has to tell is both consistent and plausible. It begins to creak however when we ask how such mistakes as that of supposing that a whale is a fish can arise, if both terms are rigid designators of natural kinds. It is

not that the concept of rigid designation makes no room for error, it is rather that this concept, based as it is on the way in which a proper name designates an individual person or place, does not make room for the kind of error involved in describing whales as fishes. For the only kind of mistake that one can make in using a proper name is to assign it to or use it in an attempt to refer to an individual or place that does not in fact bear that name. The analogy to this in the case of kind terms, whether natural or otherwise, would be a case in which the term "fish" is mistakenly used in place of the term "whale" in a context where the meaning intended is that conventionally conveyed by the latter. But the case we are considering is not like this. The nearest analogy in the case of proper names for the kind of mistake that used to be made when whales were thought to be fishes would be if someone made the mistake of supposing that Dubrovnik and Split are two alternative names for the same place in the way that Dubrovnik and Ragusa actually are alternative names for the same place. But this analogy is not a very good one as is shown by the fact that whereas the mistake of supposing whales to be fishes is one which is very easily made, one indeed that is made by almost every child when first shown a picture of a whale, the mistake of supposing Dubrovnik and Split to be one and the same place is the kind of mistake that is virtually never made, except in the case of people or places whose names are very similar like Dhaka the capital of Bangladesh and Dakar the capital of Senegal.

Similar difficulties arise in connection with Kripke's reliance on intuition as evidence of the existence of a de re necessary connection in the case of heat and molecular motion and the non-existence of such a connection in the case of mental and brain activity. For whereas it is readily intelligible that, in learning a language, a native speaker of that language should acquire intuitive knowledge of the intension or sense of the words and expressions contained in that language, and should be able to tell in this intuitive way whether or not the intensions of two words or expressions in that language coincide, overlap, mutually exclude or are independent of one another, there is no such plausible explanation that can be given of how human beings might be supposed to acquire the kind of intuitive metaphysical knowledge of what is de re necessarily the case in all possible worlds which Kripke's theory requires that we possess.

You will notice that in both these cases, the case of conceptual change over time and the case of intuitive knowledge of necessary truths, the objection I am raising to Kripke's position is that on his assumptions no plausible psychological account can be given of how our beliefs about whales and fishes could have changed as they appear to have done over time or of the possibility of our possessing the kind of intuitive knowledge of necessary truth which his theory

requires. It is this, moreover, which brings me at long last to the question which forms the title of this paper: "Can a psychologist make sense of Kripke's notion of natural kind terms as rigid designators?" In other words, can the difficulties I have raised in making psychological sense of Kripke's theory be overcome or are they so serious that in the face of them the whole theory collapses like a house of cards? Alternatively, are these psychological difficulties, as some would argue, entirely beside the point?

The view that considerations of psychological plausibility are irrelevant to the evaluation of a theory such as Kripke's would, I suppose, be defended on the grounds that Kripke's theory is a purely logical theory and logic is an a priori science which is in no way answerable to a mere empirical science, concerned only with matters of psychological fact. To suppose otherwise would be to commit the fallacy to which Frege gave the name "psychologism", the fallacy whereby logical principles are treated as if they were empirical generalisations about how human beings actually think.

I would want to counter this line of argument as follows: I have no wish to deny that psychologism as defined above is a genuine fallacy. However what makes psychologism a fallacy is not that it fails to separate a priori logical considerations from a posteriori psychological ones, but that it fails to recognise that actual human thinking is frequently illogical and irrational. Logic tells us how we ought to think and reason, not how we actually think and reason in practice. But although Logic does not describe the inferences human beings actually make, it does describe the principles they in fact appeal to in evaluating their own and other people's reasoning. That is not to say that the logician is a kind of empirical sociologist who derives conclusions about the inference-evaluating practices of social groups from observations of those practices. Logic is after all an a priori science like mathematics which aims to and, in part, succeeds in formulating the normative principles which, on pain of incoherence, must govern all reasoning in which the establishment and preservation of truth is an objective. Moreover, in so far as the establishment and preservation of truth is an objective to which no linguistic community and no member of such a community can fail to have an interest in sustaining, it follows that there can be no actual linguistic community within which the principles of logic are not respected in practice, however much they may be despised and rejected in theory.

But the fact that no linguistic community can fail to conform to the principles of logic in evaluating the inferences which are drawn within it, in no way absolves the a priori normative principles of logic from respecting the empirically derived descriptive principles of cognitive psychology. Logic can no more claim an Olympian position outside the body of science in general from which it can pass judgment on other sciences without having to

be answerable to them in matters of overlapping interest than can Epistemology, as argued by Quine in his 'Epistemology naturalized' paper which Kathy Wilkes discussed in her paper to the last course in this series two years ago.

But if logic is answerable to cognitive psychology just as much as cognitive psychology, like any scientific discipline, is answerable to logic, is there any point in trying to make psychological sense of Kripke's doctrine of natural kind terms as rigid designators in face of the evidence already adduced of the psychological implausibility of his account of historically recorded conceptual evolution and our intuitive knowledge of necessary conceptual connections? I think there is. It seems to me that, despite these blemishes, there are three important contributions which Kripke has made to the theory of meaning and the principles of taxonomy which anyone with an interest in these topics, whether from a logical and philosophical or from a psychological standpoint, cannot afford to ignore. I do not include here Kripke's contribution to modal logic, because despite its importance both in putting modal logic on a much sounder footing than hitherto and despite the way it suggests and underpins his peculiar doctrine of a posteriori necessity and of the necessity of all identity statements, I am far from being convinced that it has contributed anything of value to the theory of meaning.

The three contributions which I do value are (1) the conception of proper names as rigid designators which cuts the Gordian Knot into which the Fregean distinction between Sinn and Bedeutung, Sense and Reference, as it is usually translated, gets entangled in relation to the meaning of proper names, (2) the revival and revitalisation of Aristotle's distinction between "real" and "nominal essences" by linking real essences to natural kinds, which has made the debate between Plato and Aristotle over universals into a live issue in contemporary philosophy, and (3) the doctrine of natural kind terms as rigid designators which, although, as we have seen, won't do as it stands, does, I am persuaded, draw attention to an important difference between natural kind terms and other universals.

But if it is true both that the theory of universals must make psychological as well as logical sense and that natural kind terms are in some sense rigid designators, as Kripke holds, we cannot avoid addressing the question at issue as to whether or not sense can be made in psychological terms of the concept of natural kind terms as rigid designators. In addressing the problem of universals from the standpoint of the psychologist, the first point that needs to be made is that a living organism which depends for its survival on moving its limbs and/or its whole body relative to the objects in its environment must of necessity sort the various stimulus events which impinge on its receptor organs into kinds of object,

event and states of affairs. For example, the most primitive of unicellular organisms which depends for its survival on its ability to capture or collect and then digest suitable foodstuffs to be found in its environment must be able to sort the stimuli that impinge on its receptors into (1) those that correspond to the presence of a nutritious foodstuff needing to be digested, (2) those that correspond to the presence of poisonous substances which need to be expelled and subsequently avoided, and (3) those that are neither nutritious nor poisonous which need to be ignored as far as an ingestion is concerned.

Thus, from the standpoint of a biologically based psychology, sorting features of the environment into things of a kind for which a particular response is appropriate is the most fundamental of all behavioural processes. Indeed it is evidently more fundamental than a pattern of behaviour that is reserved for a particular individual. Consequently a theory such as Kripke's which makes the assignment of a proper name to an individual the most fundamental semantic relation makes very little sense to the psychologist. For the psychologist the recognition of an individual person or place is a matter of sorting encounters with objects and places as an encounter with this object or place rather than that object or place. In other words, it is just another way of sorting individuals into kinds, one in which the individuals are encounters and the kinds are those constituted by the different individual persons or places encountered.

Not only is sorting into kinds phylogenetically and ontogenetically more fundamental than the identification of individuals as such, it is also the case, as the example of sorting stuffs into edible and inedible shows, that sorting into kinds according to their utility in relation to the survival of the individual and the species is in this respect more fundamental than any kind of objective assignment of individuals to different natural kinds. Moreover, as this example also shows, where correct sorting is a matter of life and death there is always a discrepancy between what the Mediaeval philosophers used to call the "sensible species", the characteristic sensible appearance by which a thing of a kind is recognised as such, and its essence or defining characteristic. Thus the essence of the edible is that its ingestion and digestion should contribute to the sustenance of the organism that eats it without poisoning it at the same time. But if the organism is to survive it cannot wait to see what the effect of a particular kind of stuff on its physiology will be before deciding whether or not to eat it. It must make that decision before the stuff is swallowed and preferably before it is taken into the mouth. It is forced, therefore, by what Skinner (1975) has called "the contingencies of survival" to rely on the sensible appearance of the stuff in front of it, its look, smell, taste and feel in order to make this

discrimination. Moreover, since mistakes here can be fatal, every living organism is equipped by the processes of natural selection with an innate response mechanism which ensures that by and large the organism is attracted towards stuffs that are nutritious and is repelled by those that are toxic. But on the principle "Better safe than sorry", it frequently turns out that the organism is attracted towards stuffs which, though not poisonous, are not nutritious and is repelled by stuffs which are not in fact poisonous. Where this happens it is invariably explicable by the fact that the effect which the stuff in question produces on the relevant sense organ, in this case primarily the senses of taste and smell, is indistinguishable from that produced by an important food source or toxin as the case may be. At the same time, the probability that the organism will encounter this kind of stuff in its natural habitat is sufficiently low that no advantage from the point of view of the survival of the individual and the species is conferred by the ability to discriminate between the two. A good example of the operation of this principle is provided by experiments which have shown that rats will choose water laced with saccharin in preference to unadulterated water, despite the fact that saccharin has no nutritive value whatever. This is explained (a) by the fact that the effect produced by saccharin on the animal's taste receptors is indistinguishable from that produced by sugar which is an important source of nourishment and (b) by the fact that the chance of saccharin's occurring in the organism's natural habitat is so low that an organism with the ability to make the saccharin/sugar discrimination has no competitive advantage in terms of survival over one that cannot make this discrimination.

This kind of discrepancy between the intension or sensible species by which a kind of thing is recognised and the essence that distinguishes that kind from other kinds is very characteristic of both the categories which, according to Kripke, are rigidly designated by the names by which they are known, i.e. both natural kinds and the individual persons and places to which proper names are assigned. Thus in everyday life we recognise heat in part, as Kripke emphasises, by the sensations which are produced when our skin is stimulated by objects and stuffs at different temperatures; and these sensations can sometimes mislead us. It is perhaps worth noting in this connection that Kripke's account of heat is defective, not only, as Kathy Wilkes has pointed out in her paper on Monday, with respect to his account of its real essence as revealed by modern science; it is equally defective in its account of our ordinary notion of heat which is not confined, as Kripke seems to think it is, to sensations of heat and cold, but includes the idea of heat as a causal agent with respect to effects like the freezing of water, the melting of ice, the evaporation and boiling of water and the condensation of steam. Yet none of these features

is sufficient to give the man-in-the-street any inkling of its real essence in terms of the kinetic energy of the molecules and other particles composing the stuff which possesses this property or to which, in the case of radiant heat, it is liable to be imparted. Similarly we recognise an individual person or place by that individual's characteristic visual appearance, by the sound of the person's voice or, in the case of a place, by the sound of a particular clock chiming the hours, by a characteristic smell, and so on. But the essence of an individual person consists in the continuity of its unique career through space over time; while the essence of a particular place consists in the continuous occupation over time of approximately the same area on the face of the earth; and, as we all know, appearances can often be deceptive as a guide to such continuities. We often mistake one person or place for another or fail to recognise a person or place when their appearance has been mutilated or otherwise changed beyond recognition.

Although these basic biological sortings, typified by the classification of stuffs in terms of their edibility, resemble the things to which Kripke's rigid designators are assigned in the discrepancy that exists in both cases between the intension or sensible species by which a kind is recognised on the one hand, and its defining characteristics or essence on the other, they differ in that the propensity to sort things according to their basic biological significance for the organism is, for obvious reasons, invariably innate, whereas the ability of human beings and other "higher" animals to recognise and thus assign a distinctive set of behavioural strategies to individuals and natural kinds is undoubtedly learned. Moreover, it is when we come to examine the process whereby animals and human beings learn to sort things into kinds, and in particular when we contrast the "discrimination learning" that is characteristic of animals and of human children who are in the process of acquiring linguistic competence with the "concept formation" that is characteristic of linguistically competent older children and adults, that a distinction corresponding to that between sorting on the basis of rigid designation and sorting on the basis of a "nominal essence" or verbal definition begins to appear.

This type of discrimination learning found in animals and which, I am suggesting rather more controversially, is characteristic of a child learning to rigidly designate is distinguished by two features (a) generalisation on the basis of what, following Wittgenstein, we may call "family resemblances", and (b) discrimination on the basis of the contrast between positive and negative instances or, as Skinner (1938) puts it, between S^D , the class of stimuli in the presence of which behaviour is reinforced or punished, and S^A , the class of otherwise similar stimuli in the presence of which no reinforcing or punishing event occurs. We can

illustrate these principles by figure 1 which is taken from Karl Lashley's classical 1938 study of discrimination learning in the rat. At the top marked b you will see the two stimuli used in the initial training session, with the positive stimulus or S^D on the left and the negative stimulus or S^Δ on the right. You will see that the positive stimulus is a white triangle on a plain black ground and that the negative stimulus is a white cross, also on a plain black ground. Once this basic discrimination had been learned, the rats were tested for their ability to generalise the discrimination to the pairs of stimuli shown below. In the case of those pairs marked c, d, f, g, h, j, k, n a majority of the subjects tested showed a significant tendency to prefer the stimulus shown here on the left and which, to our eyes, resembles the positive stimulus in the standard used in training. Where the pairs marked e, i, l, m no such generalisation of the discrimination was apparent. The animals' preference for the one stimulus over the other was no better than would be expected by chance.

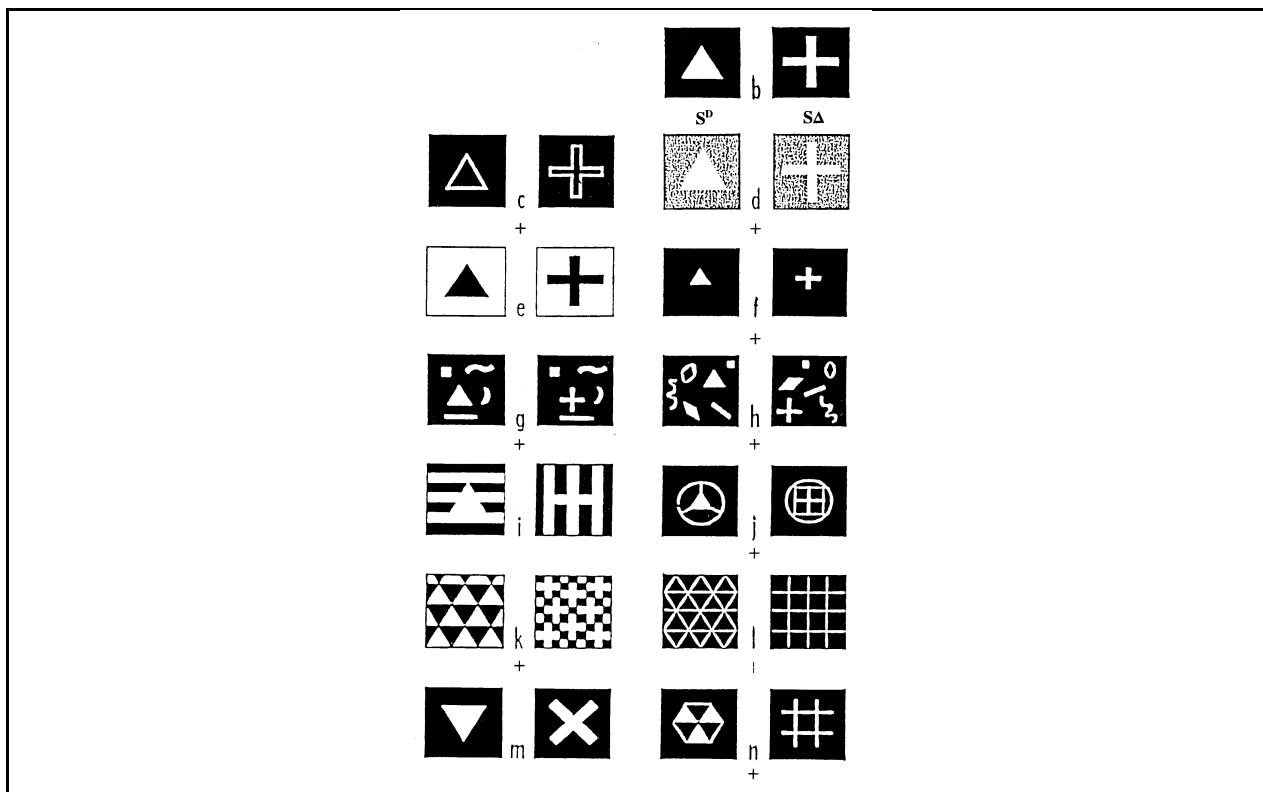


Figure 1. Pattern discrimination learning. Training stimuli b. Generalization to pairs + (Munn, 1950, after Lashley 1938).

Two points to notice about this experiment are (a) the crucial role played by the contrast between the positive instances (S^D) and the negative instances (S^Δ) in establishing this type of discrimination and (b) the fact that the only principle which appears to determine whether or not generalisation occurs to the

test figures is the number of features the pair has in common with the standard pair used in the original discrimination training. The larger the number of respects in which the test pair resembles the standard, the greater the probability that generalisation will occur. This, it seems to me, teaches us two things about this kind of discrimination learning and hence, if I am right, about the psychological processes underlying or involved in rigid designation. The first is that the boundaries between concepts whose intension is established by this kind of discrimination will only be precise in so far as (a) the sense organs of the organism are sensitive to the relevant differences between the two and (b) positive and negative instances of the kind in question which differ in the relevant respect have actually occurred in the past learning experience of the organism. In those respects where the organism's sense organs do not allow it to discriminate or where the contrast between the relevant positive and negative instances has never been encountered, conceptual boundaries will necessarily be fuzzy.

Verbal Identifier	Common element						Pack						72 test characters					
		1	2	3	4	5	6											
oo	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
yer	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
li	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
ta	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
deg	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
ling	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
hui	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
chun	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
vo	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
na	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
nez	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇
fid	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇	𠂇

Figure 2. The 144 Chinese characters used by Hull (1920) to study concept

The second point that arises from a consideration of this experiment is the resemblance which is, to my mind, by no means accidental between the way in which this type of discrimination generalises and the way in which Plato thought universals are related to their instances. Just as for Plato an individual horse does not cease to be a horse because it fails to resemble the Ideal Form of a horse in one respect, e.g. in having only three legs where the Form of horse has four, provided it resembles the Form of horse in other respects, so there is no one respect in which the test

stimulus may differ from the standard which will prevent generalisation from the standard to the test stimulus provided the test stimulus resembles the standard in a sufficient number of other respects.

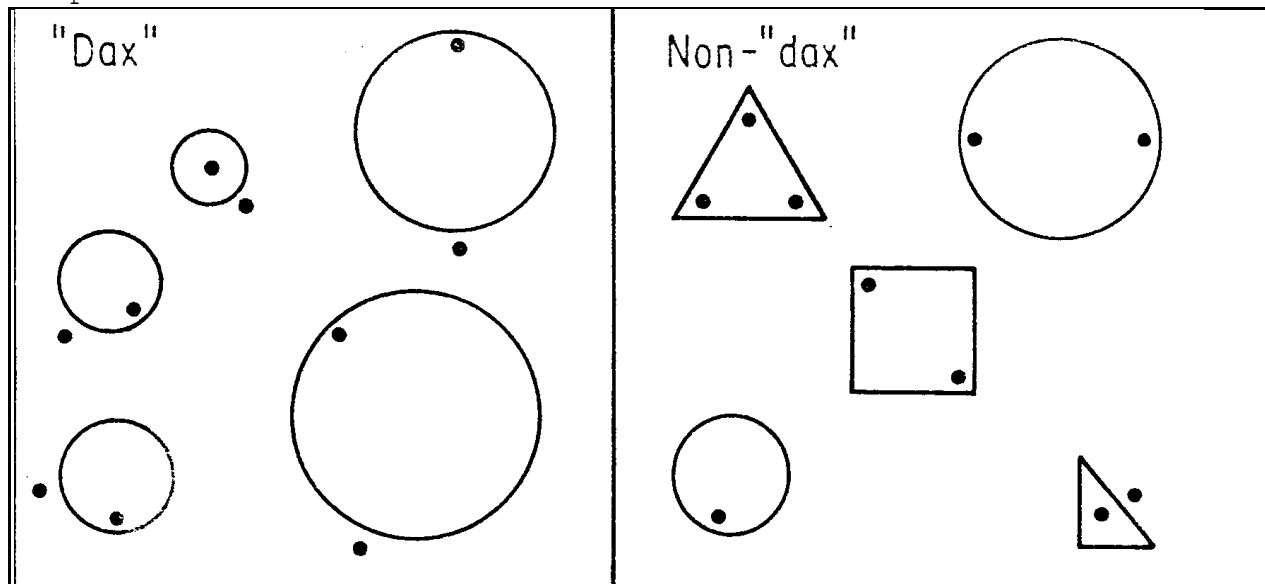


Figure 3. Samples of "dax" S^D 's and "dax" S^A 's (From Kimble and Garmezy, 1963, after Smoke, 1932).

Now let us turn to figure 2. As I shall explain in a moment, this figure is really a kind of academic confidence trick. It illustrates an experiment on what is known in the psychological literature as "concept formation" using linguistically competent human children and adults as subjects. It belongs to an experimental tradition which goes back to Clark Hull's study in the nineteen twenties of the ability of American college students to learn to pick out the common radical in a set of otherwise different Chinese characters to which the same Chinese-sounding nonsense syllable had been assigned. In the variant of Hull's experiment developed by Smoke (1932), Hull's sets of Chinese characters with a common radical embedded in each member of the set were replaced by geometrical figures like the ones on the slide in which a nonsense syllable, in this case the nonsense syllable "dax", was assigned to all those figures with a certain set of features in common, the task being to learn to identify further instances to which the same label could by the same token be assigned. In the Kimble and Garmezy (1963) version of Smoke's experiment illustrated in figure 3, we see the influence of the subsequent (i.e. relative to that of Hull and Smoke) work of Lashley and Skinner on discrimination learning in animals and of the behaviourist's insistence on interpreting human in terms of animal behaviour wherever possible. What this way of presenting the experiment is designed to suggest is that there is no real difference between what the human subject learns to do in this kind of experiment and what a rat learns to do in the Lashley

experiment illustrated by the previous slide, that, like the rat, the linguistically competent human needs the contrast between a set of positive and a set of negative instances in order to learn the discrimination between daxes and non-daxes, and that generalisation in such cases proceeds according to the principle of family resemblance, just as it does in the case of the rat.

In fact, none of these propositions suggested by the layout of the slide is true. If you take the trouble to do what colleague of mine, Dr. Larry Brownstein, did and look up Smoke's original (1932) monograph, you will find that among other things he did an experiment to see whether in fact human subjects need to be presented with negative instances in order to form a concept like that of a "dax"; and what he found, surprise!, surprise!, is that they don't. All you have to do is to show the subject a set of instances like those on the left here to which a nonsense name like "dax" has been (rigidly or non-rigidly?) assigned and after a brief inspection all subjects can reliably pick out daxes from a set of otherwise similar figures without having been shown any instances of non-daxes during the initial training phase. Moreover, unlike the rat who is easily put off by changing the figure in respects which by human standards are irrelevant to the required discrimination, the human subjects are able to recognise daxes when embedded in a wide variety of different contexts.

Nor is this difference between the sorting behaviour of the rat and the human being simply a matter of the difference in intelligence between the two species involved. It is due, as a mounting body of empirical evidence is showing, to the fact that human beings make discriminations of this kind by constructing a verbally formulated definition of the "nominal essence" which unites all the instances to which the nonsense syllable in question is assigned.

Everyone here, I suspect, will agree that you only have to look for a few seconds at the set of figures marked "dax" on the left of the slide, before one is in the position to give a strict verbal definition of a "dax" as a circle with one dot inside the circle and another adjacent to the circle on the outside. This definition, moreover, serves both as a definition of the (nominal) essence of a dax in that it specifies what it is that distinguishes daxes from other kinds of thing and as the intension or sensible species of a dax in that it also provides the criterion whereby the individual distinguishes daxes from other kinds of thing. Where kinds are differentiated by means of verbally formulated nominal essences, the gap which exists in the case of those kinds which are rigidly designated by the names they are known by between the unknown real essence of the kind and the intension by which it is recognised disappears altogether; and with it goes all the fuzziness and imprecision which is so characteristic of the intension, though not

perhaps of the real essence of natural kind terms.

My conclusion, therefore, is this. Yes, there is indeed a way of understanding the notion of natural kind terms as rigid designators which makes psychological sense. Unfortunately, it is not a way of understanding rigid designation which will satisfy a philosopher who subscribes to what we are calling "the new theory of reference" in contrast to the "old" or Fregean theory. The new theory of reference seeks to cut through the difficulties presented to the Fregean distinction between sense and reference in the case of proper names and natural kind terms by proposing that such terms simply refer to their objects without the benefit of any criteria available to the man-or-woman-in-the-street for deciding whether an instance is or is not an instance of a given kind. The problem is that, without such criteria, no rational psychological account of how such discriminations are learned in the first place and applied once learned can even begin to get off the ground.

The alternative theory of rigid designation I am proposing, instead discarding the notion that reference is determined by sense or intension in these cases, suggests that in rigid designation intension diverges from essence whereas in the case of nominal essence concepts essence and intension are one. On this view all the names and general terms that we learn on our mother's knee we learn, in the first instance, as rigid designators. We learn to apply them through generalisation on the basis of family resemblances and discrimination on the basis of the contrast between positive and negative instances. Once a certain degree of linguistic competence has been acquired, it becomes possible to acquire new concepts by the alternative process of generating or learning from others a verbal definition which specifies the "nominal essence" of the concept. At the same time we begin to learn verbal definitions of the (nominal? real?) essences of the terms which we learned originally as rigid designators. Thus we learn to identify the kind "circle" in the first instance in the manner which, on this account, makes the word "circle" a rigid designator. We learn to use the word by being shown cases of circles and contrasting cases that are not circles. Only later, when we learn geometry, do we learn the definition of a circle as a line all points on which are equidistant from a single point and begin to use this definition as a criterion for excluding as cases of genuine circles instances like that which Zoran drew on the board on Monday evening, thus making the definition into the intension as well as the essence of the term. Maybe we should view the process whereby scientists discover the real essence of things like heat and water as just another example of this process whereby gradually increasing verbal sophistication results in the conversion of what begin as rigid designators of individuals and kinds whose real essence is unknown into kinds with a verbally specified nominal essence in which essence and intension are united

in a single formula.

- Hamilton, W. (1860). *Lectures on Logic* (H. L. Mansel and J. Veitch, Eds.). Blackwood.
- Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs*, 28(123).
- Kimble, G.A., & Garnezy, N. (1963). *Principles of general psychology*. Ronald.
- Kripke, S. (1972). Naming and necessity. In G. Harman and D. Davidson (Eds.), *Semantics of Natural Language*, Reidel.
- Lashley, K. S. (1938). The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *Journal of General Psychology*, 18, 123-193.
- McGinn, C. (1977). Anomalous monism and Kripke's cartesian intuitions. *Analysis*, 37(2), 78-80. [doi:10.1093/analys/37.2.78](https://doi.org/10.1093/analys/37.2.78)
- Munn, N. L. (1950) *Handbook of Psychological Research on the Rat: An Introduction to Animal Psychology*. Boston: Houghton-Mifflin.
- Place, U. T. (1956). [Is consciousness a brain process?](#) *British Journal of Psychology*, 47, 44-50.
- Place, U. T. (1991f). [On the social relativity of truth and the analytic/synthetic distinction.](#) *Human Studies*, 14, 265-285. [doi:10.1007/bf02205609](https://doi.org/10.1007/bf02205609)
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis of behavior*. Appleton-Century-Crofts.
- Skinner, B. F. (1975). The shaping of phylogenetic behavior. *Journal of the Experimental Analysis of Behavior*, 7, 117-120.
- Smoke, K. L. (1932). An objective study of concept formation. *Psychological Monographs*, 42, No. 191.