To appear in:- Concepts: New Directions, Eds. E. Margolis & S. Laurence

Concepts in the Semantic Triangle[1]

James A. Hampton

City University London

Looking at the range and variety of contributions to this volume it is evident that "concept" is a term that means many things to many people, a fact that has been widely acknowledged (Dove, 2009; Machery, 2009; Weiskopf, 2009). Probably the most central issue of all, the one over which there is least agreement, is how to explain or describe the meaning or content of concepts. What information do they carry and how do they do so?

The aim of this chapter will be to describe three different ways in which it has proved useful to talk about the meaning or content of concepts. The first comes primarily from philosophy and uses the notion of *reference*. When we think a thought, or utter a statement, the words that we use refer to particular things in the world. Each concept term has its *denotation,* the class of things to which it refers. A second way to think about conceptual meaning is by looking at *language use*. When we use a word we can do so either appropriately in a way that others will readily understand, or inappropriately in a way that others will object to. For this approach, meaning is a matter of social convention or practice. The third way to look at meaning is to ask what information is represented in a person's mind at the time that they have a thought or express an idea. As well as asking what a concept refers to, and what the appropriate use of the term in language might be, we can also ask *how* an individual mind achieves this result. What must we assume about the representation of information in the mind that enables a possessor of a concept to use it in these ways?

These three approaches to concepts and meaning lead to very different accounts of what a concept is. While not promising to unravel the tangled knot of how the accounts relate to each other, I hope at least to lay out some of the issues involved.
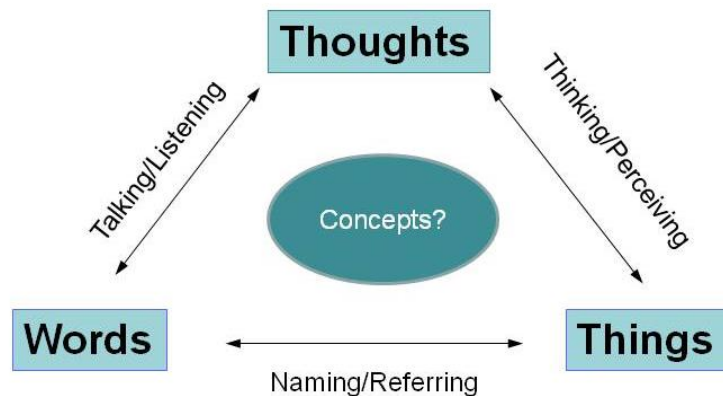
1. The Semantic Triangle

An early view of the relation between different forms of meaning was the classic "*Meaning of Meaning"* by Ogden and Richards (1923). Figure 1 uses their famous semantic triangle to illustrate

James A. Hampton, Department of Psychology, City University, Northampton Square, London EC1V OHB
Email: hampton@city.ac.uk
Web: www.staff.city.ac.uk/hampton

the complex relations that we are considering here.

Ogden and Richards were more interested in illustrating the ways in which misunderstanding can occur as people communicate their thoughts through the use of words. For the present purpose I have relabelled the links to make a rather different point. The three corners of the triangle turn out to represent the three accounts of concept identity that I will be discussing. Taking the accounts in reverse order, at the top of the triangle is the psychology of the individual. In psychology, researchers investigate concepts by interrogating individuals about their conceptual knowledge and intuitions. Psychologists aim to uncover the "meaning" of a concept by direct and indirect measurement of how the individual represents the concept and what function it plays in directing their behaviour. This behaviour will include (among other things) memory, judgments of similarity, inductive reasoning, categorization, naming and language comprehension. For the psychologist, then, a concept is a postulated entity whose character can be inferred from its effects on behaviour. Like many other theoretical terms in psychology since the behaviorist revolution (intelligence, personality and motivation would be cases in point) the theoretical notion of concept is defined operationally in terms of its observed effects on measured behaviours.

The approach that focuses on Words on the left of the triangle is to be found in linguistics, as represented for example by lexical semantics. The study of lexical semantics takes a language such as English as its basic object of study. Pinning down the meaning of a concept, and differentiating it from others, is a matter for cultural study – in effect developing a natural history of the concepts deployed in a given linguistic-cultural group. For example one might try to understand how conceptual terms (the substantive words in a language) can be differentiated into classes such as verbs of motion, animate nouns, adjectives of degree and so forth, in a way that explains their behaviour in different linguistic contexts. The "English Language" is of course an abstraction or idealisation based on empirical observation of texts and the intuitions of speakers averaged across a given language group. So whereas the psychology of concepts is primarily about individuals trying to understand and negotiate their way through the world, the linguistic approach is primarily about the net effect that social communities speaking a common dialect or language and achieving shared meaning have on the standard use and meaning of linguistic terms.

Finally, the bottom right corner of the triangle emphasizes the crucial role that concepts play in our interaction as a social-cultural species with the real world. Thoughts and utterances can have a

semantic value, in the sense that they can be *true* or *false*. Broadly speaking, being true means corresponding to the way things actually are. So an equally important part of the triangle is to provide a "reality check" or constraint on thoughts and words keeping them in touch with the way the world actually is. The development of science, in the broadest sense, is not just the accumulation of facts. Equally important is the evolution across the centuries of the appropriate concepts with which to articulate hypotheses and theories (Kuhn, 1962). In parallel with this positive evolution of "correct" concepts there will be the gradual extinction of concepts that don't do the job of providing us with terms for understanding the world.

Writing in 1851, Herman Melville has his narrator Ishmael define the whale thus:

> *How shall we define the whale, by his obvious externals, so as*
> *conspicuously to label him for all time to come? To be short, then, a*
> *whale is A SPOUTING FISH WITH A HORIZONTAL TAIL. There you have*
> *him. However contracted, that definition is the result of expanded*
> *meditation.*
>
> (Melville, H. (1951) *Moby Dick, Ch. 32.*)

Classification of animal and plant species into the now widely accepted scientific taxonomy taught in schools is a relatively recent achievement of biology. Classifying a whale as a mammal and not as a fish provides a classification that makes sense of a range of data – facts about the anatomy and physiology of the whale such as that it breathes air and suckles its young, but also about the evolutionary history of species and how and where they branched off from other life forms in the past. Humans are free to create whatever concepts or classifications they choose. In the 19th Century, the fact that whales and cod were in a similar category for the purposes of commercial fishing meant that it was intuitively evident that they should both be classed as fish (at least to those who had direct experience of them). Subsequently, increasing education and biological understanding in the population has led to a shift in classification (but see Dupré, 1999, for an alternative view).

The third approach of considering the real world of things serves to tie together the other two corners of the triangle. If we first of all assume that there is a real class of things – for example whales – in the world out there, then we can use that to say both what someone is thinking about (their idea or conception of whale as it plays a role in whale thoughts), and what someone is talking about (the meaning of the word "whale" in the language). It has therefore proved tempting to rely heavily on this third corner of the triangle as the corner-stone for fixing the meaning of concepts. Note, however, that having first to assume the real existence of the class may raise problems for this approach (Wikforss, 2005).

2.1 Psychology and concepts

My own interest in concepts has followed the path taken by psychology. As readers may be aware, interest in how people categorize and conceptualise the world took a great leap forward with the work of Rosch and Mervis in the 1970s (Rosch, 1975; Rosch & Mervis, 1975). Previous work on concepts in psychology had followed two main paths. The Piagetian tradition concentrated on conceptual development and saw concepts as logically structured schemas that organized the perceptions and actions of an individual (Piaget, 1953). While consistent with modern approaches to

understanding concepts as complex representations, Piaget's theory of adult cognition was effectively an idealisation. In Chomskyan terms it was an account of adult competence, in that it sought to explain and elaborate the developmental processes that lead to the full power of hypothetico-deductive thought in the adult. The overly strong focus on logic and the over-elaborated descriptive aspect of the theory meant that with the advent of cognitive experimental psychology in the 1960s, developmental and reasoning psychologists tended to see Piagetian theory as a target for criticism. New methodological developments revealed more about the conceptual abilities of infants and young children, while at the same time the reasoning of adults was discovered to be at odds with the Piagetian account (Wason, 1960). The continuation of this tradition has seen children portrayed as ever more intelligent at ever younger ages (see for example, Gelman, 2003), and adults as ever more likely to fall into logical fallacy and error (e.g. Kahneman, Slovic and Tversky, 1982).

The other way in which concepts were traditionally considered in psychology was as an offshoot of learning theory in the behaviorist tradition, with concepts treated as classification rules that were learned through a process of hypothesis formation and testing (e.g. Bourne & Restle, 1959). Another classic set of studies from this era was by Bruner, Goodnow and Austin (1956) who explored how people set about discovering categorization rules of a given logical form through entertaining hypotheses and sampling exemplars in order to test them.

Whereas psychologists had previously been mainly content to define concepts in terms of simple conjunctive rules, Rosch and Mervis set out to discover how actual concepts were represented within individuals. The most direct way to do this is to ask them, and Rosch and Mervis did just this. Their participants described their understanding of terms like lemon, trout, chair or hammer by listing descriptive properties. From the results of this listing exercise came the idea of prototype concepts. As I have argued in Hampton (2006), the key insight here is that people represent the central tendency of a class rather than its boundaries. It is as if when defining London you specified the centre of the city (Trafalgar Square, say), and gave an indication of its approximate size. The rest is left vague. What you don't do is to trace on the map the actual boundary of the city. Indeed people's understanding of geographical features such as mountains often works in much this way (Fisher et al., 2004). In a similar way people's concept of a hammer is represented by an idealised notion of what most typical hammers are like, without a clear rule for deciding when something is or is not a hammer.

The initial idea of a prototype as just a list of features or properties clearly lacked sufficient representational power. There was also considerable confusion about whether prototypes were intended to be primarily *visual* (e.g. Osherson & Smith, 1981; Rips, 1989). Some theorists assumed that the notion of "similarity" appealed to by prototype theory was akin to "surface similarity" or "similarity in appearance". They therefore argued that prototype theory could not account for categorization on the basis of deeper or more abstract properties such as function or historical origin. It can be argued for example that the (presumably sincere) intention of a designer to create a particular kind of artifact is the single necessary defining property of artifact kinds. Their appearance and their actual ability to perform the intended function are irrelevant (see Bloom, 1998). Hampton et al. (2009) provided evidence that while originally intended function does influence categorization, it is treated along with current function and appearance as one feature among many affecting the likelihood of being included in the concept class. Similarly Malt and Johnson (1992) showed that the

function of an artifact is neither necessary nor sufficient for determining its kind. Prototype concepts should therefore be considered as integrating multiple sources of information, including perceptual, functional and historical features as required.

Another source of confusion comes from the claim that prototype concepts lack necessary and sufficient conditions for membership. On the one hand, such conditions can easily be formulated using similarity – to be a category member it is both necessary and sufficient to be more similar to the prototype than some threshold level (Hampton, 1995). On the other hand, concepts that do have necessary and sufficient defining features may still be represented as prototypes. When two or three features are individually given maximal weight such that each on its own outweighs the sum of all other features, then a threshold can be set which renders the features singly necessary, and jointly sufficient for category membership.

In any case, the notion of prototypes was largely supplanted in psychology in the mid-1980s with the notion of causal schemas (Medin and Murphy, 1985). When more is known about a domain, then knowing about the relations among the features becomes increasingly important. Perceptual similarity is only a starting point for forming conceptual categories. A child just observing how a word is used may form a simple prototype on the basis of associating the word with the appearance of a set of objects. Most creatures appear to have the ability to learn on the basis of similarity of appearance. Conditioned responses generalise to novel situations to the extent that the new situations are similar to those in which the responses were learned. Initial prototype formation on the basis of sensory features is therefore a very fundamental process. However even young children quickly learn that appearances may be misleading (Gelman et al. 1994), and adults understand that a range of different features and the right kind of relations among them are important for determining kinds (Rehder, 2003).

Current psychological accounts of concepts and categorization are consequently keen to emphasize the role of concepts in understanding and causal explanation. Even so, the key insight described above remains critically important. Our causal schemas still provide a descriptive account of the central or typical cases of a conceptual category. Only rarely will they provide a clear rule for deciding category membership at the boundary. In this sense the causal schemas are still prototypes.

2.2 Possible problems with the psychological approach

As an account of concepts in general, there are several well known problems with the path that psychology has followed. These problems relate to the difficulty of using the contents of an individual's mental representation as a means of pinning down a concept's meaning.

The first difficulty is the problem of error. People's concepts may in fact contain incorrect information about the world. For example, many people think that snakes are slimy to the touch. Never having felt one, and seeing the shininess of their skins, it is a reasonable inference to draw, but happens to be false. The problem is that if we take the concept SNAKE to be whatever people understand by that term, then we would have to agree that snakes are slimy. The proposition "snakes are slimy" would be evaluated as true or false, not by looking at actual snakes, but by looking at people's beliefs about snakes. This problem led Rey (1983) to propose, along with other philosophers, that the psychological notion needs a different name – say "conception" – to differentiate it from the correct concept which allows truth to be determined in relation to the

actual world. Otherwise, if it were the case that everyone in a given community believed that snakes were slimy, then their concept SNAKE would in fact refer to nothing in the actual world, thus remaining in the realm of fantasy. It is more plausible to suppose that their concept SNAKE still refers to the actual set of snakes, but that what the concept actually denotes is not just determined by the contents of their beliefs about snakes.

A second difficulty arises from the impossibility of having a debate about a topic that doesn't just descend into a question of terminology. Suppose that Joanna is a zoo keeper and believes that snakes are dry to the touch. Then along comes Katy who thinks that they are slimy. When they find that they disagree, all that we should be able to say using the psychological sense of concept is that they have different concepts of snake. They don't disagree about the facts, because the way Joanna uses the word "snake" is different from the way Katy does. Their different beliefs lead to different concepts, so all factual disagreement turns into terminological disagreement.

The third difficulty is closely related to this one. To be useful, the meaning of a concept needs to have some means of differentiating the information that constitutes its meaning from information that is otherwise true of it (Miller & Johnson-Laird, 1976). Compare the information that you would find about snakes in a dictionary as opposed to an encyclopaedia. The average person may be able to tell you all kinds of things about snakes, but most of it may be incidental to the question of what type of thing a snake actually is – its conceptual definition. Being cold-blooded and reptilian and lacking limbs seem to be a core part of the meaning of the concept, whereas being used for making handbags or being a Christian symbol of the Temptation of Eve do not. In spite of models that try to make the distinction between defining and characteristic features of concepts (e.g. Smith et al. 1984) no good way of differentiating the two psychologically has been found. For example, while all species of birds have two legs, but only some species fly, when describing birds these two features are equally likely to be produced, and are rated equally important to the concept.

3. External definitions – a solution to the problem

To resolve the problem of error and of disagreement, a good solution is to fix a concept meaning in relation to reference. If the meaning of a concept is tied to the class of THINGS to which it refers, then these difficulties do not arise. Conceptual classes according to this proposal are the real classes of things that actually exist in the world. Deciding whether snakes are slimy is not a psychological question but one about biology. People's mental representations of the world in their conceptual store are not concepts themselves, but are representations OF concepts –what Rey called conceptions. (For Fodor, 1998, they are atomic unanalysed symbols in an internal language of thought). Thus we can say that Joanna and Katy disagree about a fact, rather than just having different concepts, since the question of sliminess vs. dryness is to be resolved by examining the class of things being talked about (which for both of them is just the actual class of snakes).

To handle the problem of the dictionary versus the encyclopaedia, and assuming that the differentiation cannot be achieved by psychological methods (i.e. through behavioural measures), then there are two possible solutions. One possibility is to follow Melville's Ishmael and attempt a definition. A definition is an explicit verbal formula that will correctly classify cases that fall under the concept and discriminate them from those that do not. For example:

Snakes *are elongate, legless, carnivorous reptiles of the suborder Serpentes that can be distinguished from legless lizards by their lack of eyelids and external ears*

Wikipedia, 27-9-2012

This is a concept definition in a form of which Aristotle (and indeed Linnaeus) would have approved. A set of properties are described which are all true of all snakes, and all true only of snakes. Any beliefs not included in the definition (e.g. relating to the texture of the skin) would be part of the encyclopaedia and not the dictionary. As we saw with Ishmael's definition of a whale, however, definitions can change over time. The evolution of science and culture frequently require a change in concept definitions, and in this case the question arises of when they should be treated as the same concept –with an evolving definition—rather than as two different concepts.

The other way to resolve the dictionary/encyclopaedia problem is to follow Fodor's (1998) informational atomism. For Fodor there are no definitions, so that concepts are simply abstract unitary symbols that stand for the appropriate classes or properties in the world. Thoughts then are "just" structured sets of symbols, much as this text is represented on the computer's hard disk just as a series of 1s and 0s on a magnetic medium. They derive their meaning through their relation to the external world. For Fodor, all the properties of snakes are encyclopaedic, with the exception of "is a snake". Effectively the problem of discovering the meaning of a concept is passed over to other disciplines – in particular to the sciences, while the problem of how humans come to possess concepts is a part of evolutionary psychology.

4.1 Meaning as use

Let us turn next to the Words corner of the triangle. Wittgenstein famously wrote

> "For a large class of cases – though not for all – in which we employ the word meaning it can be defined thus: the meaning of a word is its use in the language game."

> Wittgenstein, *Philosophical Investigations* (1953), §43.

Perhaps then, language behaviour, or how people use words when talking to each other, might provide a way around the problem of deciding the content of a concept.

Consider the case of adjectives describing simple properties. What is the meaning of a colour term like "beige" or of a height term like "tallish"? It is not obvious how we can use the external route here (but see Williamson, 1994). We believe that whales and snakes have a real existence as kinds, because of our scientific and cultural knowledge of what it is to be a natural kind, what determines species boundaries, and so forth (I leave aside the interesting question of the difficulty of defining the concept of species in biological theory, Mayr, 1982). But we cannot turn to any equivalent set of validated knowledge to tell us that beige or tallish things constitute a class independent of our own understanding of them. There is no theory of beigeness or tallness.

How do we fix the content of these concepts? We could take a psychological view, and say that each person has their own concept of beige and their own concept of tallish. But that will again raise the problem of error. If Jed likes to call anything that is deep purple (to us) "beige" we will be justified in calling him to account – he is just wrong about that word and its meaning. On the other hand a

difference of opinion about just where on the spectrum beige turns into a light tan or a pink would not be grounds for denying that someone possesses the concept.

In cases like these, the most sensible course is to define concepts by the consensus that exists among a group of speakers. To have the "correct" concept of beige cannot be a matter of being in possession of a symbol or definition that correctly picks out the real class of beige things in the world, since there is no independent way of saying whether that class really exists. To have the correct concept is rather to use the word in a way that would be considered acceptable or normal by most other speakers of one's dialect. Partly these "others" will be one's peers with whom one interacts, but they will also include the full range of social and cultural influences to which one is exposed.

This proposal makes good sense for a whole range of concepts. For example it provides a very reasonable account of how children learn the meaning of many words. They hear and observe how words are used in different contexts and then gradually learn to use them in the same way. Keil and Batterman (1984) showed how young children first attach the meaning of words to the most obvious characteristic features of the named class. A bowl of cornflakes is breakfast regardless of the time of day at which it is eaten. A robber is a man with a gun who takes your TV away, even if heyou're your permission to do so. Similarly Satlow and Newcombe (1998) found that children's classification of geometric shapes as triangles or circles were based on broad similarity rather than defining rules. If children form hypotheses of word meaning on the basis of what they find interesting to attend to in the context in which they hear the word used, and a general notion of similarity, then one would expect just this kind of result. Interestingly, Lupyan (2013) has evidence that adults also may categorize triangles by similarity rather than by their strict definition. Triangles that were judged as atypical were also more likely to be rejected as being triangles.

The proposal also provides a way of rescuing the validity of psychological research tools for uncovering people's prototypes and conceptual schemas. In an unpublished study I ran with Danièle Dubois in 1995, we explored whether people would see any distinction between how words are used and conceptual classification. We selected eight common categories like Tools, Sciences or Fruits, and created a list of 24 items for each category. The list included a full range of degrees of membership in the category. For Sciences, for example, it included typical sciences (Chemistry), atypical sciences (Dentistry), and a range of less and less scientific disciplines down to a clear non-member (Literature). Two groups of participants then made Yes/No decision about each category list. The first were asked to decide for each item whether it belonged in the category. Is Chemistry a Science? Is Literature a Science? The second group were asked to make a judgment about word use. They were asked "As a speaker of your language, do you think that you would spontaneously use the word "Science" to refer to each of these items?" The study was conducted on two groups of students in France, and in the French language.

The results were very clear, as can be seen in Figure 2, which plots the probability of people saying yes to each item, either in the categorization task (the vertical axis) or the linguistic judgment task (the horizontal axis).

Figure 2

Linguistic judgment of word use
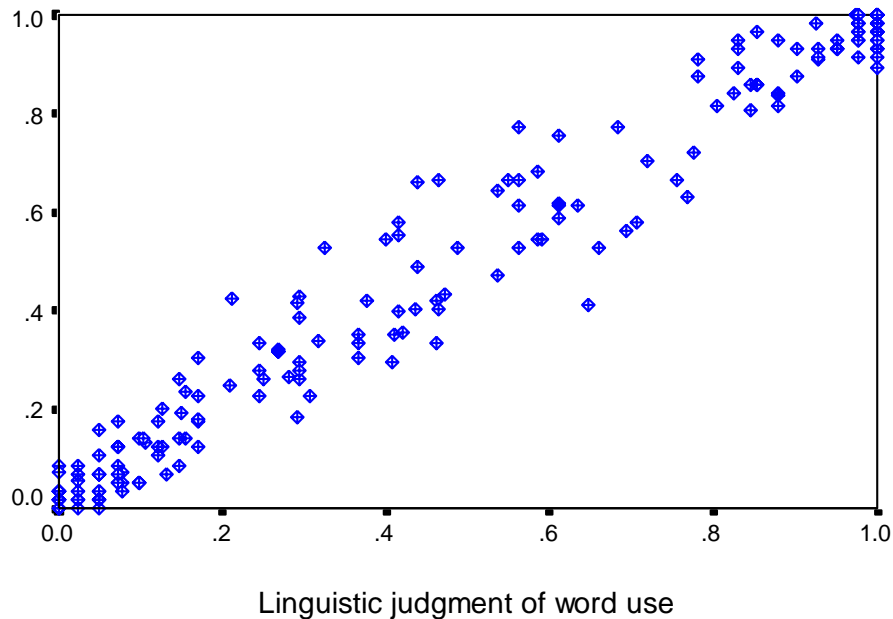
The correlation between the two groups was near perfect, with no systematic variation beyond that expected by the error of measurement of the probabilities. In addition these probabilities of categorization or agreement with word use were closely correlated with the data from a third group who were asked to judge how typical each item was of the category. In sum, there is just one underlying representation of the concept here. The closer an item comes to matching the expected prototype of the category, the more typical it is judged to be, the more likely people are to endorse it as a category member, and the more likely they are to imagine themselves using the category label to refer to the item. According to the Meaning as Use account, this prototype is the result of generalising across many contexts of use in which the word has been heard or successfully used in communication. The mental lexicon evolves through coordination with other speakers.

The evidence then suggests that asking people about the extensions of concepts (as in whether bananas are in the extension of fruits) is equivalent to asking them about the use of words (as in whether they would refer to a banana using the word fruit). Traditional psychological tasks such as typicality ratings, attribute listing and category membership judgments are therefore reasonably valid ways of getting at the social consensus about concepts.

4.2 Problems with the meaning as use proposal

Recent work by Barbara Malt and colleagues has thrown some doubt on the utility of the "meaning as use" proposal for fixing conceptual content. The difficulty arises from data that she has been collecting on the use of different terms in a language to refer to subclasses of artifacts or functional objects around the home. Malt et al. (1999) took photographs of a large number of simple containers such as cartons, boxes, bottles and jars, and showed them to different groups of speakers of English, Chinese and Spanish. Some groups judged similarity of the objects, while others had to name them in their own language. Remarkably, while there was a good correlation (r > .9) in overall

similarity across the different language groups, the naming data showed only weak correlations (.35 < r < .55). There was clearly a disjunction between the fact that the different cultures saw the similarity between the objects in much the same way, but had very different name categories. Note that the disparity was not just because different languages divided up the similarity space in different ways (as for example when Russian divides the English colour range BLUE into a dark and a light category). The conceptual categories derived from similarity were just a poor basis on which to base the classes with a common name. Objects that were similar could have different names, and objects that were dissimilar could have the same name, and different languages mapped onto the space in very different ways.

This result and subsequent research by Malt and colleagues throws some doubt on treating language use as a basis for exploring underlying concepts – especially for artifacts. The traditional "survey" method of asking people whether some object is (for example) a BOTTLE may not provide a clear picture of how objects are classed into conceptual types in people's minds. The difficulty, as identified by Malt et al., is that naming is affected by the pragmatics of communication and also reflects historical change. A pair of similar artifacts may start out with the same name, but over time they may evolve in different ways so that while they still keep the same name, their similarity can be lost.

So while the usage of the language community certainly provides a way of coordinating one's own concepts with those of others, the relation between knowing how to use words and possessing particular concepts is not as direct or straightforward as one might hope. In fact how things are named may not provide a reliable or valid basis for describing the underlying conceptual structure in people's minds.

Malt and Sloman (2003) go so far as to argue that many artifact concepts are not treated as kinds. They argue that classifying something as a weapon, or furniture or a hammer is highly context dependent. These concepts have strong functional properties. A weapon is something used to harm another. But there is no stable category of weapons, since not only objects created with this intention, but almost any other object becomes a weapon in the context of a person intending to use it to harm another. Thus the police may bring a charge against someone for being in possession of a baseball bat if they have reason to believe it is intended for use as a weapon.

5. How to ensure that we are talking about the same thing

When we communicate with others, the very definition of successful communication is clearly going to depend crucially on our ability to know when we are talking about the same thing. Otherwise we might not be speaking the same language, and disagreement and agreement would be meaningless. If by "fish" I mean what biologists mean by fish, while Ishmael means something like "fish-shaped creatures of the deep", unless we are aware that we have these different concepts, we will be open to all kinds of misunderstanding.

From the discussion so far, I have outlined two ways in which we might achieve this coordination of meanings. One is to pin meaning to external classes that exist in reality. This approach works well for concepts that contribute to scientific understanding. Snakes and whales exist independently of humans (so we have reason to believe), and so the most natural way to fix the meaning of concepts such as these will be in terms of the actual class. The second way is to pin meaning to implicit social

conventions about word use. Just as it is appropriate to say "thank you" when receiving something positive from another, so it is appropriate to say "red" when asked the colour of a particular shade. Both ways of coordinating meaning, either by the external world or by language convention, imply that it is possible for a person to represent a concept either correctly or incorrectly. As a consequence concepts and meanings are not "in the head" to use Putnam's phrase (Putnam, 1975). An individual who wishes both to understand the world, and to be understood by others, has a responsibility to represent the correct concepts.

Interestingly the problem of error works differently for Things and for Words. It is quite conceivable (indeed it has happened) that everyone in a cultural group has the wrong concept of (say) disease. They can all identify cases of individuals who are sick, but they believe that disease is caused by witchcraft and the malign attentions of invisible spirit forces, and that treatment involves invocation of spells and sacrifices. But it would be very odd to say that they all have the wrong *meaning* for the word "disease" in their language. When concepts are tied to language use, then either a majority of speakers or a powerful minority (for example the priesthood or an educational elite) must be the "keepers" of the correct use. While everyone's concepts can be incorrectly attached to the world, meanings cannot all be incorrectly attached to the language group.

There will be situations where the two principles may come into conflict. Obviously the question of what real classes actually exist is not always easy to answer. Most people take it as a given that the natural world exists and that modern science provides the best account of the kinds of things that are out there. But the scope of science is very narrow in relation to the full repertoire of concepts a person will have. Consider the realm of psychological disorders. To advance their science, psychiatrists and clinical psychologists need to determine what categories of psychological disorder are to be found in the population. (Their decisions are crucial, since deciding on the existence of these categories of disorder is also of vital importance for putative sufferers if they are to gain access to medical care funded by the state or other agencies).

Early theorists distinguished those conditions which were considered disorders of the nervous system (neuroses) from those that were disorders of thought (psychoses), but this distinction is far from easy to draw, and the term neurosis was dropped from the classification scheme currently in use by the American Psychological Association. The DSM IV-TR, a text revision of the fourth version of the Diagnostic and Statistical Manual of Mental Disorders (1994) was created by explicitly using prototypical descriptions of the behavioural symptoms associated with each condition. Classification based on theory (largely psychodynamic theory) had been earlier rejected in favour of categorization based on observable indicators (Robins & Guze, 1970). Thus, for example, the diagnosis of schizophrenia in DSM-IV includes a set of five characteristics symptoms any two of which need to be present over a period of a month for diagnosis to be triggered (unless one is very severe). It is therefore possible for two people in the same category to share no symptoms in common. The question of whether there is a single "real" condition or many different conditions with similar symptoms is left unanswered by the DSM. Only with a better theory of the underlying causes of the problem would such an account be possible.

Within the semantic triangle, we find that in the absence of an understanding of the real nature of the condition (the Things corner), we fall back on the common usage among experts – the Words corner of the triangle. The DSM attempts to codify this common usage, but acknowledges that the

definitions may need to be revised as knowledge advances. At the time of writing, the fifth edition of DSM is being prepared, with the aim of using neuroscience and genetics to provide a classification that will take more account of causal etiology and depend less on clusters of symptoms (Kupfer and Regier, 2011). Many patients are anxiously awaiting the decision of whether their condition will still exist in the new classification.

In medicine, because a cure is best achieved by identifying the cause of an illness (for example a virus, or other infection) "real" conditions are often understood as those with a single cause, in keeping with psychological evidence that people in categorization experiments in the laboratory will place greater weight on causal properties than on effects (Ahn et al., 2000). Thus malaria is an infection by the Plasmodium parasite. Having the symptoms without this infection is not a case of malaria. There is a single necessary and sufficient criterion.

However a medical category can also be defined by properties that are further along the causal chain. Primary hypertension is the condition of raised blood pressure and is attributable to a combination of different causal factors which converge on the same effect. Given a diagnosis of hypertension, then further effects and symptoms can be predicted, so the conceptual category still provides predictive value. Hypertension is not itself diagnosed on the basis of a single cause, but as a category it exists at an intermediate level, providing an explanation of further medical problems.

On the other hand a diagnosis of arthritis refers to a condition of inflammation in the joints, and can have over 100 different causes and related treatments. Diagnosis of arthritis therefore carries almost no information beyond its presenting symptom of persisting joint pain. It is at the end of the causal chain.

What can be learned about how concepts are defined from these examples? Medics have found a need for conceptual categories at all levels of causal depth. While real world categories with common primary causes provide the most useful source of concepts, there are also categories that are defined at intermediate levels in the causal path, or even just in terms of their effects. While it is possible to be wrong about a diagnosis of malaria, it seems highly implausible that a doctor could be wrong about a diagnosis of arthritis. Having the concept of arthritis is just a matter of knowing how to apply the word.

In sum, the concepts that evolve within a field such as psychiatry or medical sciences generally have to serve many purposes. They can be pinned to the cause or aetiology of a condition, they can be pinned to intermediate changes in physiology that have negative consequences, or they can refer to those consequences themselves. The selective pressures that lead to particular concepts becoming established must be a complex interaction of their "real world" status as providing explanatory and predictive potential and their communicative value in describing cases to expert and non-expert audiences. For concepts fixed to deep causal principles in the real world, we could all be wrong – for example schizophrenia may turn out to have no single cause or theory to explain it. But for concepts fixed by language use, we cannot all be wrong. It could not turn out that no doctors had really been using the word "arthritis" correctly.

6. Fixing conceptual contents by explicit definition

In what other ways are individual concept users constrained in the way they represent concepts? A third way of fixing meaning is by explicit definition. Concepts of this kind conform to what Smith and Medin (1981) called the Classical model of concepts. For example a bank note is a piece of paper that has a particular origin in a particular place of manufacture (e.g. a central bank or mint). One can specify the precise conditions of whether something belongs to this class. Indeed it is vitally important economically to be able to do so. The distinction between valid and invalid currency absolutely requires an explicit definition if people are to have faith in it. The construction of categorization schemes such as the DSM described above seeks to provide explicit operational definitions for psychiatric conditions.

To take another practical application, when lawmakers create a new law they have to provide definitions of concepts in a way that can be applied by the courts. When a juror in the UK has to judge the evidence concerning a charge of receiving stolen goods the judge instructs them that the offence requires three things to be established beyond reasonable doubt. The accused must have received the goods, those same goods must have been previously stolen, and the accused must have had knowledge that the goods were stolen at the time that they were received.

This is not to say that courts do not have to make difficult decisions about the meaning of concepts enshrined in the law. In fact Endicott (2000) argues that vagueness is a necessary feature of the law, given the need to allow judges discretion to take into account all the possible circumstances in any given case. Explicit definitions do not preclude difficult decisions, and lawyers spend many happy and lucrative hours trying to persuade the courts of the interpretation that best suits their clients. But one function of civil society has been to provide explicit definitions for concepts that were originally based on social consensus about typical exemplars. Once these definitions are in place, then we have a third means of ensuring that people are talking about the same thing. The individual who accepts the rule of law thereby accepts the way in which the law defines particular concepts.

As a final example, stipulation of definitions can also be found in the physical sciences. In 2006, following a period of consultation, the International Astronomical Union created by a vote of its General Assembly a new classification of large objects orbiting the sun, by which Pluto was "demoted" to the status of a dwarf planet. The types of argument and debate around this controversial move make for fascinating reading for those interested in concepts, and how the choice of particular definitions affects the practice of science (Messeri, 2010).

7. Fixing concepts by their mental representation

The discussion so far has focused on two sources of conceptual content. There is the real nature of the world and the kinds that are found there, and there is the existence of socially coordinated terms in language which require individuals to adapt their concepts according to word use including socially stipulated definitions of concepts provided by those in a position of authority to centralise and regulate this process of coordination.

What then of the last corner of the triangle – the idea of concepts as mental representations of kinds? We have seen how the problems of error and disagreement make it difficult to fix meaning in terms of Thoughts. However one can also see how it is only through their appearance in thoughts themselves that we know anything at all about concepts. So one notion of "concept" that can be defended is that one cannot have concepts without people to conceive of them.

The right way to understand this corner is as the place where the other two sources of constraint come together. I doubt that there are concepts that exist in our minds that are neither tied to real classes in the world, nor to the classes resulting from learning how others use words to label things. But the last corner of Thoughts is in a way the most interesting, both psychologically and philosophically. It is in how our minds come to represent concepts that we must look to explain how people think, communicate and understand the world.

As I discuss in Hampton (2012), we have learned a great deal about concepts from looking at how individuals think and reason. In everyday cognition people categorize the world according to prototype similarity and this leads to errors in reasoning. In that paper I review a range of examples to illustrate this point. Examples are the well-known conjunction fallacy reported by Tversky and Kahneman (1983). They showed that when judging the probability that some individual is in a given class, people do not use the logic of sets and class inclusion, but instead turn to judgments based on similarity and representativeness. Tversky and Kahneman called this type of reasoning "intuitive". Here I will briefly discuss an example of this intuitive mode of thought, based on recent research I have conducted with Martin Jönsson and Alessia Passanisi (Jönsson & Hampton, 2006; 2012; Hampton, Passanisi & Jönsson, 2011).

Our research was triggered by an effect discovered by Connolly et al. (2007). They asked people to judge how likely it was that certain sentences were true. Among their sentences were some simple generic statements such as "ravens are black", and others in which the same statements had modified subject nouns, as in "feathered ravens are black" or "jungle ravens are black". They observed that when the subject noun was modified, ratings of likelihood decreased relative to the unmodified forms. Moreover, the effect was larger if the modifier was atypical (as in jungle ravens) than typical (as in feathered ravens). Connolly et al. used this effect to argue that the view of concepts as prototypes is incorrect. We were naturally concerned therefore to explore the underlying basis of the phenomenon.

In Jönsson & Hampton (2012) we replicated the effect and explored people's explanations for choosing the unmodified sentence as more likely. Primary reasons were based on pragmatic considerations of trying to be maximally informative, and knowledge based reasons of imagining a scenario in which the modifier would affect the property (for example that jungle ravens might be camouflaged for jungle living). We concluded that, contrary to the argument advance by Connolly et al., most of the time people do in fact assume that a modified concept will inherit the properties of its unmodified parent concept.

Following on from this study, we ran a study considering what would happen to the likelihood of modified and unmodified statements when they were expressed as universally quantified – "all ravens are black" versus "all jungle ravens are black" (Jönsson & Hampton, 2006, which actually appeared before the Connolly et al. paper because of publication lag). In this situation, a rational answer should not allow that it is more likely for a property to be true of a whole class if it is not also true of a subclass. If all jungle ravens are not black, then clearly it cannot be true that all ravens are black. However our participants continued to rate the unmodified version of the statement as more likely to be true – just as if the universal quantifier had not been present. Similar results showing that people ignore the logical implications of universal quantification have been reported by Leslie, Khemlani and Glucksberg (2011) in what they term the Generic Overgeneralization Effect. In their

studies, participants judged sentences like "All ducks lay eggs" to be true in spite of the experimenters' best efforts to remind them that male ducks do not lay eggs.

From these two examples, we can see that mental logic works in a different fashion from the logic that one would use when defining actual classes in the world. If conceptual contents are fixed by reference to real world classes, then the logic of real world classes should equally apply to conceptual contents as we represent them. Logics (including fuzzy logics – see Hampton, 2011) provide a poor model of how people reason conceptually. For a psychological theory of concepts and conceptual thinking, we must be able to explain people's behaviour as well as people's idealised competence. The parallel with other forms of reasoning is a good one. While the Words and Things corners of the triangle are concerned with ensuring correct or accurate conceptual frameworks, the Thoughts corner provides us with insight into the messy heuristics that underlie actual behaviour.

8. Conclusion

In this chapter I have considered three approaches to fixing conceptual meaning and it should be clear that there is no one right approach. The example of snakes and people's mistaken beliefs about them shows how people's mental representation of concepts can be in error. To explain how we can have incorrect conceptions we have to be willing to let the real world be the arbiter of the meaning of our concepts. On the other hand there are equally clearly cases where it is not the physical world but rather the social world that fixes concepts – either through informal coordination among speakers of a language, or through the social power structures that permit particular bodies to stipulate how words and concepts should be defined. It is hard to escape the view that the fixation of conceptual meaning has more than one source. Rather than seeing this as an indication of heterogeneity or plurality of concepts, I would argue that the different sources actually come together and are integrated in our mental representation of the concepts. Possessing a concept, in the Thoughts corner of the triangle, is a matter of building a schematic or prototypic representation of a concept that will be tied to the real world, to the use of language, and to the social constructions of one's society. Many of the most important social and moral issues of the day, be it abortion, drugs legislation, same-sex marriage, mental disorders or human rights are debates about concepts. Is an embryo a human being? Can one define a category of dangerous substances that includes cannabis but excludes alcohol? Is it a defining feature of the concept of marriage that it should be between a man and a woman? Is sexual fetishism a mental disorder? The list of such debates is extensive, and an understanding of both the social influences on how concepts are negotiated and the psychological barriers to conceptual change are central to understanding the debates as they evolve.

Future directions

Concepts provide an exciting area in which all the cognitive sciences -- particularly philosophy, psychology, social anthropology, and linguistics -- can develop useful and mutually enlightening collaborations. It is important that researchers recognize that there is no single answer to the question of what concepts are, or where their meanings reside. There is a variety of concepts, some clearly grounded in physical reality, while others are grounded in culture and language. The mind itself has been shown to have different systems for categorization, involving both rules and similarity and associations (Ashby et al., 1998. All interested parties need to expand the range and variety of

the concepts that they use in their thought experiments, their linguistic analyses or their behavioural experiments. Often the literatures have failed to engage with each other because of an insufficiently broad view of the range of different concepts that humans possess. For this reason the development of Experimental Philosophy as a discipline is to be welcomed. Not only is it valuable to test and explore the intuitions underlying different philosophical positions, but the very act of conducting empirical tests leads to interesting reflection and debate on the value of the evidence. A philosophical argument that is based on the philosopher's clear intuition that in circumstances Y people would think X is clearly open to empirical test by placing people in circumstances Y and measuring what they think. However this suggestion can be countered with the argument that we should not be doing philosophy by conducting opinion surveys. The correct answer is not necessarily the one that attracts the largest endorsement. So some important work needs to be done to develop better guidelines on just which questions can be settled by behavioural experiment, and which should be answered by rational argument.

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105,* 442-481.

Bloom, P. (1998). Theories of artefact categorization. *Cognition, 66,* 87-93.

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. Oxford: Oxford University Press.

Hampton, J.A. (2012). Thinking intuitively: The rich (and at times illogical) world of concepts. *Current Directions in Psychological Science, (in press)*

Hampton, J. A., Storms, G., Simmons, C. L., & Heussen, D. (2009). Feature Integration in Natural Language Concepts. *Memory & Cognition, 37,*

Hampton, J.A., Passanisi, A., & Jönsson, M.L. (2011). The Modifier Effect and Property Mutability. *Journal of Memory and Language, 64*, 233-248.

Jönsson, M.L. & Hampton, J.A. (2006). The Inverse Conjunction Fallacy. *Journal of Memory and Language, 55*, 317-334.

Jönsson, M.L. & Hampton, J.A. (2012). The modifier effect in within-category induction. Default inheritance in complex noun phrases. *Language and Cognitive Processes, 27, 90-116.*

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Lupyan, G. (2013). The difficulties in executing simple algorithms: why brains make mistakes computers don't. Manuscript under review.

Machery, E. (2009). *Doing without concepts*. New York: Oxford University Press.

Malt, B. C. & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory and Language, 31,* 195-217.

Messeri, L.R. (2010). The Problem with Pluto**:** Conflicting Cosmologies and the Classification of Planets**.** *Social Studies of Science, 40,* 187-214.

Rehder, B. (2003). Categorization as Causal Reasoning. *Cognitive Science, 27,* 709-748.

Weiskopf, D. (2009). The plurality of concepts. *Synthese, 169,* 145-173.

Wikforss, A.M. (2005). Naming natural kinds. *Synthese, 145,* 65-87.