

Concepts and Prototypes

James A. Hampton
City University London

Multiple Review of *Concepts: Where Cognitive Science Went Wrong* by Jerry A. Fodor - Concepts and Prototypes.
Mind & Language
Volume 15 Page 299 - April/June 2000- Volume 15 Issue 2&3

1. Introduction

The cover of Fodor's book proudly claims that this is his most irritating book in years, guaranteed to exasperate all those who read it. The book lives up to this promise. Although leavened by moments of wit and humour, Fodor misses no opportunity for the one-liner put-down, be it about lexical semantics, empiricism, cognitive neuropsychology or the psychology of cognitive development. He even writes a whole chapter on Prototypes without referring to Rosch except in a footnote, where her name is consistently mis-spelled for good measure. Those looking to see how his ideas relate to current work in psychology and linguistics on lexical concepts will be disappointed. Indeed the short reference list mentions no recent psychological work at all on the subject.¹

What we have instead is a (frequently charming) philosophical presentation of a radical position concerning the nature of concepts - namely Informational Atomism (IA). This thesis comes down to two major claims:

- a) informational semantics -- that the content of concepts is constituted by some sort of nomic mind-world relation, and in particular the class of entities in the world that the concepts are 'locked to', and
 - b) conceptual atomism -- that concepts have no internal structure at a linguistic or psychological level. They are in effect atoms.
-

Fodor notes that while his negative arguments about concepts are strong, the exposition of IA is still exploring new territory. Intriguingly, his version of IA comes tantalisingly close to the view that he is determined to reject! He spends five chapters explaining how prototypes (or stereotypes as, following Putnam, he also calls them) are not the right kind of thing to be concepts. Yet his account of just how concepts become locked to the appropriate set of objects in the world makes a direct appeal to them. To paraphrase:

A concept X is that property that minds like ours lock to in the context of instances of the stereotype of X.

Now Fodor comes down very hard on other theories that allow the term X to appear in the definition of Xness, on grounds of circularity. He claims to have circumvented the problem by making the concept depend on the prototype. The concept after all is an unanalysable internal mental particular, whereas the prototype is 'patently specifiable without reference to the [concept] viz. by enumerating the shapes, colours, functions, and the like that [Xs] typically have.' But he

The author would like to thank members of the London Concepts Group and in particular Neville Austin, Nick Braisby, Bradley Franks, and Andy Wells for many stimulating discussions of the ideas in Fodor's book. I also drew on the work of Jesse Prinz and Lance Rips in composing this commentary.

Address for correspondence: Professor J.A.Hampton, Department of Psychology, City University, Northampton Square, London EC1V OHB

¹ Fodor refers those interested in a review of the widespread effects of typicality to the seminal book by Smith and Medin published in 1981, thus neatly side-stepping nearly two decades of subsequent empirical and theoretical work.

Email: J.A.Hampton@city.ac.uk

offers us no suggestions at all of how these prototypes come to be formed, or what ontological status they should take, and so his claims of non-circularity appear empty. Are prototypes psychological representations or externally defined classes? If psychological (as seems plausible, since they are composed of sets of intensional propositional attitudes like beliefs and inferential roles) then what are the conditions for their acquisition, representational format and utilisation? The proposal that atomistic concepts are somehow created in our minds through exposure to instances, identified as such by some other part of our mind which has already created prototypes for them is truly bizarre. Why should we have evolved such a dual representation system? More importantly, isn't all the interesting action going to be in the study of the prototypes rather than in these unstructured abstractions? Alternatively if Fodor intends prototypes to have a mind-independent external existence then the situation becomes even more absurd. Not only must the stereotype of DOORKNOBS (and its associated similarity metric) be around in the external world, ready to lead our minds into getting locked to the corresponding concept DOORKNOB, but so, presumably must all our other concepts - KITSCH, FRUMPINESS, GRUMPINESS and the rest. It is not at all clear what has been gained by this move.²

Given that even Fodor accepts that prototype effects are truly widespread in cognitive psychology, and given that his proposal for concept locking presupposes the existence of prototypes, it is worth looking more closely at the arguments for conceptual atomism. What leads Fodor to assert that 'it's as certain as anything ever gets in cognitive science that concepts can't be prototypes' (p94.)? I will consider two of his chief arguments -- the lack of consistent analyses of internal structure, and the failure of prototype compositionality.

2. Internal Structure

Conceptual complexity effects are generally notable by their absence. KILL may have a definition of CAUSE TO DIE, but there are no studies showing that KILL is psychologically any more complex than DIE (or that BACHELOR is more complex than UNMARRIED). Similarly Fodor has much fun with lexical semanticist's vain attempts to explain how all the different senses of a word like KEEP are related to some internal conceptual structure. They are not. The senses that a word can take are sanctioned as much by arbitrary conventional usage as by rules of semantics.

But picking on particular failed analyses does not really advance the argument. Saying that what all instances of use of 'keep' have in common is that they all mean KEEP is as empty as it is trite. Fodor fails to explain why everywhere in his book the level of atomic concepts is equated with the lexicon of American English. There is a footnote (p.42) where he admits that if there should turn out to be many cases where a concept is lexicalised in one language but not in another then he has problems. But there are surely indefinitely many such cases (most particularly for words like KEEP). Nor is it just that we have (for example) a lexicalised concept in English for UP but the French do not. The locking theory that he advances has to suppose that where each language adopts a different way of dividing up a semantic domain into labelled prototype-based categories, so all of these potential concepts are out there, and which one your mind locks to will magically depend on the language community and culture that you are a part of.

In effect, there is nowhere in Fodor's account an explication of the role of language in the whole process, and yet coming to understand one's culture (and hence acquiring its concepts) obviously involves understanding the proper usage and reference of the terms of one's language. French and German have a range of emotion terms that have no simple one-to-one translation between each other or with English. It cannot be that the emotional lives of speakers of these languages are so different that they just happen to lock onto different concepts, which in turn just happen to correspond to lexical concepts in their own languages.

Note too the problems that Fodor's proposal is likely to encounter in meeting the Publicity constraint (number 5 of his not-negotiable conditions on a theory of concepts). How is it possible to show that two people have become locked to the same concept X? It cannot be because they use the label 'X' to describe or categorise the world in the same fashion, nor can it be because

² He seems also to be forced into accepting Descartes' proof of the existence of God, on the grounds that we could not have the concept of God unless there was such an entity for our minds to lock to.

they concur with the same set of beliefs about Xs. Neither of these are constitutive of conceptual content. We can only rely on blind faith or stipulation to support the hypothesis that we are locked to the same concept rather than two different ones. When we also happen to speak different languages, (or one of us is 4 years old), the problem is even more apparent.

Identifying the lexicon as the atomic vocabulary of one's conceptual system also rules out the possibility of accounting for the productivity of sub-lexical morphology. Obvious examples are sub-, super-, re-, de-, un-, -able, -ful, -ness, -ivity, and so forth. We can readily form neologisms with these morphemes, and they carry standard conceptual content in these novel forms. Thus my roses may be prunable, my dentist may detartar my teeth, my field may be replowable. There is no dictionary entry for these lexical forms, and their productivity strongly suggests other more familiar words also have internal structure. For if PLOWABLE means CAN BE PLOWED, then why should not a more familiar lexicalised form such as ENJOYABLE also have this analysis? Of course following this move leads back into the murky regions of lexical semantics where the complex residue of centuries of lexical evolution is revealed (VIABLE does not mean CAN BE VIED, etc.). Language is full of generative rules and fossilised forms that break those rules, and we had best get used to it. Denying there are any rules does not get us very far.

Words in the lexicon often act as if their meanings are 'encapsulated' and not analysed psychologically into component parts when encountered in some text or conversation. A number of different experiments (e.g. Barsalou, 1982) show that the contribution of a lexical concept to the representation of a thought or belief is strongly context-dependent. The lexical item acts as a retrieval cue or key to enable the construction of a meaning for the complex thought that is consistent with the constraints of general knowledge of the world. The meaning of lexical items can be considered atomic in the sense that there may be no finite way to list all the subtly different senses that the item can express (or for that matter the number of different sets of entities that it may pick out through reference). Words may be radically polysemous (Nunberg, 1979). If concepts are to be tied fairly closely to substantive words (as just about everyone, including Fodor, would have them be) then concepts too must be amenable to many and varied contributions as components of thoughts. How this is possible without some kind of internal structure is problematic.

3. Prototype compositionality

Fodor's second argument concerns the lack of compositionality of prototypes. For example the complex concept BROWN COW takes all of its meaning from its components, and the meaning that it derives from them is just the semantically evaluable features IS BROWN and IS A COW which are not prototypes but simply atomic predicates. Furthermore there are indefinitely many Boolean functions on simple concepts that will have not prototypes, such as NOT A DOG. The class of things that are not dogs is not categorised on the basis of similarity, so the concept clearly has no prototype.³ We therefore have to 'abandon the project of using prototype structure to account for the productivity (/systematicity) of complex Boolean predicates.'

If this is taken literally it seems no major abandonment, since no Prototype theorist denies that we are able to think in terms of sets and interpret the logical form of sentences in terms of intersections, disjunctions or complements. It's just that this is not a skill that comes easily, and it appears to bear little relation to our conceptualisation of the world. (See the literature on belief bias, and syllogistic versus heuristic reasoning). Furthermore, once we try to account for how people actually understand the natural language equivalents of Boolean operators (i.e. relative clause modifier constructions, 'not' and 'or'), we find that we will need those prototypes after all.

I hope I will not be considered immodest if I point to my own empirical work on the issue of how prototypes compose. (For reviews see Hampton, 1996a, 1996b). Fodor's only familiarity with the field is Smith and Osherson's (1984) model which accounts for typicality effects in simple

³ To possess the concept NOT A DOG requires the ability to judge dissimilarity from, as opposed to similarity to, the DOG prototype. We can possess such concepts, and so we must have this ability. Some negative concepts are even lexicalised, like UNBELIEVER. But note that the prototype is itself still a key constituent of the complex concept, which explains why a tractor is a better example of the concept than a cat, and why a wolf or fox may be considered borderline cases.

intersective adjective–noun combinations such as BROWN APPLE. Fodor criticises the model on the grounds that one could not raise the weight of the COLOUR feature sufficiently high to ensure that no non-brown objects could be included in the class of brown apples. Smith and Osherson were well aware of this problem and explicitly denied trying to account for class membership in the conjunction. It also happens that Fodor’s criticism is incorrect. There are several technical solutions to the problem, including raising the weight of the colour feature to a level that exceeds the sum of the weights of the other features and setting the threshold at that level (Hampton, 1995), or alternatively using logistic combination of feature weights.⁴ Smith and Osherson (1984, see also Osherson and Smith, 1997) are in fact good Fodorians. They both eschew definitions, and deny any role for prototypes in determining categorization.

Turning to the empirical evidence on how people interpret quasi-Boolean operators suggests however that prototype effects are in fact still at work. Yes, PET FISH should logically all be PETS and FISH, but people often fail to respect these logical constraints. First there is a consistent tendency to overextend the conjunctions -- that is people typically deny (for example) that chess is a sport, but they will happily agree that it is a sport that is a game. Most will deny that a tent is a dwelling, but everyone will agree that a tent is a dwelling that is not a building. Few people are willing to classify a mushroom as a vegetable, and none classify it as a fruit, yet a considerable majority classify it in the class ‘fruit or vegetable’. I have studied conjunctions (A that are B), negated conjunctions (A that are not B) and disjunctions, and in every case there is ample evidence that the way that we use these compound phrases is not accurately modelled by Boolean set logic (Hampton, 1997).

I have similarly looked at how the features that constitute concept prototypes are inherited by complex concepts. In brief, most features were inherited by conjunctions as a function of their importance for the constituent concepts. In addition, there were constraints such that attributes necessary for a constituent were also necessary for the conjunction, and attributes impossible for either constituent were also impossible for the conjunction even if important for the other constituent (these constraints were a notable concession by the subjects in the direction of classical logic). Of particular interest for Fodor’s argument was the finding of emergent features -- attributes considered unimportant for either constituent but which were nonetheless key features of the complex class (for more details see Hampton, 1996b). For example, pet birds live in cages and even talk -- something that no-one considers true of either pets or birds when considered alone.

How could prototypes show compositionality if (a) important stereotypic features of a concept are not inherited by a complex concept of which it is a part (pet birds do not build nests in trees), and (b) important features of the complex concept emerge which are not considered true of either constituent set alone (pet birds talk)? I would argue that these facts alone do not rule out the possibility of compositionality for prototype concepts. They merely rule out the simplest type of classical compositionality in which all concepts have a set of necessary features, which are then inherited by all conjunctions of those concepts.⁵ Instead concepts follow a default inheritance hierarchy in which features are inherited by default, and subject to constraints such as the impossibility of two features being exhibited by one and the same object.

Emergent features can be accounted for in one of two ways. First there is a principle which I termed extensional feedback. When combining prototype classes, people often identify individuals or classes that clearly fall into both categories. Noticing that creatures that are both pets and birds typically live in cages, they will add this to the list of attributes that they can generate about the conjunctive category, thus elaborating the complex prototype. Second, when faced with apparently incompatible attributes inherited from the two classes, people may also elaborate the complex prototype in creative ways to reduce the incompatibility and render the whole more coherent. For example, if asked to describe a piece of Fruit that is also a piece of Furniture, subjects are able to invent interesting hybrids of huge pineapples carved into chairs, or

⁴ For example, define a set of feature weights $w_i = [0, 1]$ and the degree of match of an instance to a feature $m_i = [0, 1]$. Categorization is then a binary decision based on placing a threshold criterion on a similarity scale defined as $\prod_i (1-w_i)^{(1-m_i)}$.

⁵ IA would come under this class of inheritance rules, with the proviso that the “necessary feature” for being an X is simply “Is an X”.

watermelons used like bean bags (Hampton, 1996b, see also Kunda, Miller & Clare, 1990). In doing so they consider many of the design problems involved (fruit rots and furniture shouldn't) and describe different ways to resolve them (disposable furniture versus genetically modified fruit). These studies clearly show people in the act of composing complex concepts from prototypes, although the process is neither simple nor respectful of logical form except in an approximate fashion.

There is insufficient space here to go into the wider study of conceptual combination when nouns are concatenated to form complex noun phrases (see for example Wisniewski, 1996 for an excellent account). Systematic compositional principles found include the selection or construction of some mediating relation (a LADDER BOX = a box to keep a ladder in), attribution of a salient property of the first noun to the second (a LADDER CHAIR = a chair with a back with rungs) or the formation of hybrid concept (a LADDER FIRE ESCAPE = a ladder and a fire escape). Suffice it to say that if we have only atomic concepts, we will never begin to understand the generativity and complexity of this aspect of our conceptual ability.⁶

Consider atomism in the familiar world of physics/chemistry. A Fodorian chemist might argue thus. Elements (primitive atomic concepts) enter into many different compounds (complex concepts) but there is nothing about the chemical properties of these compounds that can be derived directly from the known properties of the elements except for the fact that the one contains the other. What oxygen contributes to water just is the property [CONTAINS OXYGEN], *tout court*. Apart from the simple rule of composition, there is no way that the properties of water can be related to those of oxygen and hydrogen, any more than the properties of pet fish can be related to those of pets and fish. But this chemist would be quite wrong. Atomism in chemistry proved to be incorrect. Although we still talk of individual atoms of oxygen or hydrogen⁷, these are not the smallest indivisible bits of matter, and they do have complex internal structure from which the properties of their compounds can be predicted. Cognitive science has yet to find its Plancks, Heisenbergs or Schrodingers but it is surely premature to give up the hope of understanding the way in which prototypes compose.

4. Conclusion

Prototypes can be concepts, because prototypes do combine in systematic (although complex) ways. If they can be concepts, then there is no need, I would argue, for the Fodorian atomic concepts that we come to possess through locking to the properties that instances of the concept prototype possess. We simply come to be locked to classes in the world through possessing representations of those classes in the form of prototypes, familiar instances, background knowledge, naive theories and beliefs and the rest of it.

These prototype concepts may act in simple encapsulated ways as unanalysed atom-like tokens when (for example) following the logic of simple arguments such as $p \ \& \ q \rightarrow p$ (or for that matter any complex Boolean function you like, for those with a training in logic, a pencil and plenty of time and paper). But this is not their primary function.

James A. Hampton
Department of Psychology
City University London

⁶ Take PET FISH for example. What account could be given of PET HOBBY or PET THEORY? Who would be satisfied with an account that simply says "a pet theory just is a pet and a theory"? And what kind of prototypical examples lead our minds to lock to the appropriate primitive concept of PET that can cover all these generative cases.

⁷ An interesting case of how words can change meaning. In this case "atom" ceased to have its previously central meaning of "indivisible" and became the name attached to a particular level of divisibility.

References

- Barsalou, L.W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition*, 10, 82-93.
- Hampton, J.A. (1995). Testing Prototype Theory of Concepts. *Journal of Memory and Language*, 34, 686-708.
- Hampton, J.A. (1996a). Conceptual combination. In K. Lamberts and D. Shanks (Eds.) *Knowledge, Concepts and Categories*. pp 133-159. London: UCL Press.
- Hampton, J.A. (1996b). Emergent attributes in conceptual combinations. In T.B.Ward, S.M.Smith, & J.Viad, (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes*. pp 83-110. Washington DC: American Psychological Association Press.
- Hampton, J.A. (1997). Conceptual combination: conjunction and negation of natural concepts. *Memory & Cognition*, 25, 888-909.
- Kunda, Z., Miller, D.T., & Clare, T. (1990). Combining social concepts: the role of causal reasoning. *Cognitive Science*, 14, 551-578.
- Nunberg, G. (1979). The non-uniqueness of semantic solutions. Polysemy. *Linguistics & Philosophy*, 3, 143-184.
- Osherson, D., & Smith, E.E. (1997). On typicality and vagueness. *Cognition*, 64, 189-206.
- Smith, E.E., & Osherson, D.N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337-361.
- Wisniewski, E.J. (1996). Construal and Similarity in Conceptual Combination. In T.B.Ward, S.M.Smith, & J.Viad, (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes*. pp 51-81. Washington DC: American Psychological Association Press.