Why We Should Do Without Concepts
BARBARA C. MALT

Abstract: Machery (2009) has proposed that the notion of ‘concept’ ought to be eliminated from the theoretical vocabulary of psychology. I raise three questions about his argument: (1) Is there a meaningful distinction between concepts and background knowledge? (2) Do we need to discard the hybrid view? (3) Are there really categories of things in the world that are the basis for concepts? Although I argue that the answer to all three is ‘no’, I agree with Machery’s conclusion that seeking a single characterization of concepts will not be fruitful for understanding cognitive representations and processes.

1. Introduction

For at least 100 years, psychologists have talked about concepts as fundamental units of thought. They have tried to characterize the nature of concepts and how concepts are acquired, stored, and used. Machery (2009) makes a bold claim that progress in the psychology (and neuropsychology) of concepts requires eliminating the notion of ‘concept’ from researchers’ vocabulary altogether.

The argument is driven by his Heterogeneity Hypothesis, which has 5 tenets:

1. The best available evidence suggests that for each category (for each substance, event, and so on), an individual typically has several concepts. There is not just a single concept of dogs (or chairs, or holidays, etc.) but several distinct ones.
2. These co-referential concepts are heterogeneous; they have very few properties in common. They store different types of knowledge and are used in different cognitive processes.
3. Prototypes, exemplars, and theories are among these different kinds of concepts. For example, people have a prototype of dogs, exemplar knowledge of dogs, and theoretical knowledge about dogs (where ‘theoretical’ encompasses knowledge that explains why things happen, such as causal relations between one property and another; e.g. dogs are friendly to humans because they have been domesticated through selective breeding from wild ancestors).
4. Prototypes, exemplars, and theories are typically used in distinct cognitive processes.

I thank Gregory Murphy for helpful comments on an earlier draft of this paper.

Address for correspondence: Barbara Malt, Department of Psychology, 17 Memorial Drive East, Lehigh University, Bethlehem, PA 18015, USA. Email: barbara.malt@lehigh.edu

5. Because of the preceding four points, the notion of a ‘concept’ ought to be eliminated from the theoretical vocabulary of psychology. The term concept implies that there is a single thing we are trying to understand, but the term does not actually pick out a natural kind. Using this term impedes correct characterization of knowledge in long-term memory (LTM) and its use in cognitive processes and leads to unproductive debates about which one is right. In doing so, it distracts attention from the real needs, which are to develop several distinct theories appropriate to the different kinds of knowledge and their acquisition and use, and to develop an overarching theory that explains how their use is coordinated by the larger cognitive system.

The book is careful and provocative. Machery provides an excellent review of major issues in the psychological literature on concepts and categorization and a very useful discussion of the contrasting goals of the philosophers who study concepts and the psychologists who do. He avoids muddying the waters by making explicit which set of goals (namely, the psychological ones) he is addressing, and he clearly defines what he counts as a concept. Doing so ensures that the book is explicit about how the claims advanced here contrast with other existing and possible positions.

I am sympathetic to key parts of the larger thesis that Machery advances. I agree that the field can only spin its wheels debating which view of concepts (prototype, exemplar, or theory) is right, because there is ample evidence for the relevance of all three notions for characterizing knowledge in LTM. I also agree that there is something fundamentally misguided about the way that psychologists have been approaching trying to figure out how knowledge stored in LTM mediates our interactions with substances, events, and other entities in the world. However, I take issue with some of the logic by which Machery reaches these conclusions. In the rest of my comments, I will question three of his arguments and then pose an alternative approach for consideration.

2. Is There a Meaningful Distinction between Concepts and Background Knowledge?

It is a given that people have much information about any given object, event, substance, etc., stored in LTM. Machery wants to separate the portion of knowledge that is in a concept from the portion that is general background knowledge. He proposes that concepts are the bodies of knowledge that are used by default in the processes underlying the higher cognitive competences. Default means what is normally retrieved, except when some specific additional information is provided. Thus, he says: ‘The knowledge that is stored in a concept of x is preferentially available when we think, reason, and so on, about x. So to speak, it spontaneously comes to mind. By contrast, the knowledge about x that is not stored in a concept of x is less available—it does not spontaneously come to mind. The knowledge that
is not stored in a concept of x is used only when the knowledge that is stored in this concept is insufficient or inadequate for the task at hand. In such cases, people access their long-term memory in order to retrieve some additional knowledge about x that helps them deal with the task at hand (i.e. some knowledge that is not stored in the concept of x)’ (pp. 11–12).

A first problem is how to specify the knowledge that will count as ‘normally’ and ‘spontaneously’ retrieved in a meaningful way. Consider the knowledge that is retrieved when hearing a word such as chair. What knowledge comes to mind depends on how long a person has to think about it. In the first 20 seconds, what comes to mind may be that chairs have 2 arms, 4 legs, and are for sitting. In the next 20 seconds, more information will come to mind: maybe that chairs can be made of wood or metal, can be upholstered, can be stuffed, and can be used at a table or in a living room. And with further thought, what comes to mind may also include that they can be made of plastic or aluminum, can have no arms, can be used at the beach, can have casters instead of feet, can have rockers, and so on. To define a default, there would need to be a cutoff for what counts as the default or spontaneously retrieved information. It is not clear if there is a non-arbitrary way to make that decision.

Suppose we decide that what comes to mind in the first, say, 20 seconds is what we’ll call spontaneously retrieved (and the rest requires more reflection or effort). What comes to mind in the first 20 seconds is no doubt heavily determined by frequency of association. Many things that are called chair have four legs and two arms, and so those properties are highly associated with the word. Maybe armless chairs are slightly less frequent in daily experience, so people are less likely to first think of armlessness when hearing the word chair. But in actuality, there are many armless chairs in common use—for instance, those at the kitchen table, used twice or more daily by many people. Does it make sense to say that having two arms counts as part of the concept of chair, and having no arms doesn’t? A second problem is that it does not seem well motivated to make a discrete cutoff point on a frequency of association continuum to determine what should count as belonging to a theoretically distinct type of representation.

A third problem is that people rarely, if ever, simply encounter someone saying chair and not saying anything else. Words are used in the interest of communication, and communications are generally in the form of phrases and sentences. Even sentences with little direct description of a chair bring to mind different properties of chairs, as in Pull up a chair versus She curled up in her favorite chair in front of the fire versus He picked up his towel, sunscreen, and chair and headed for the parking lot. Experimental evidence indicates that different contexts activate different properties associated with the same word (e.g. Barsalou, 1982; Roth and Shoben, 1983; as Machery discusses in the book). More generally, retrieval from memory is known to be highly context-dependent (e.g. Godden and Baddeley, 1975), so it seems inevitable that in ordinary discourse settings, what knowledge a word retrieves will differ from utterance to utterance. If we fall back on the idea that property retrieval in these contexts doesn’t reflect what happens by default, we are defining concepts
relative to a situation that rarely if ever occurs. Machery expressed concern in other parts of his discussion with the artificiality of some experimental paradigms. If people are given stimuli and tasks that are quite different from what happens in real life, researchers may find out something about how they function within that setting, but not what happens in real life. I think we run the same danger here. And if we look instead for some knowledge that might be retrieved regardless of context (which may exist, e.g. ‘brown’ for bears; Connell and Lynott, 2009; see also Barsalou, 1982) then, again, we run the risk of eliminating knowledge that intuition suggests should still count as part of the concept (e.g. that not all bears are brown, or that chairs can be armless as well as having two arms).

Now consider the knowledge that’s retrieved when encountering an object rather than a word. Someone sees a novel object, for instance, an unusual living room chair never seen before, and she interprets its properties by retrieving some stored knowledge about objects experienced in the past. Again, retrieval from memory is sensitive to the content of the probe. What will the chair in the left panel of Figure 1 retrieve? What about the one in the right panel?

Will they all retrieve the same default chair concept? It’s an empirical question, but I doubt it. Besides the basic facts of how memory retrieval works, it may not be useful to retrieve the same default knowledge of chairs to interpret each of these. It would be more useful to retrieve those aspects of stored knowledge that can best aid in interpreting the particular features of each. So a fourth problem is that it’s hard to imagine what retrieval of default knowledge could be in the context of encountering objects, as opposed to in the context of hearing a word.

Figure 1 Some novel chairs
Finally, consider exemplar representations, which Machery proposes as one of the three major types of concepts. What would it mean to have default knowledge that is retrieved for exemplar representations? Does it have to mean that whenever exemplar knowledge is accessed, exactly the same set of exemplars is retrieved? My last example suggests that this isn’t likely to be correct. If an object retrieves memories of related objects that have been experienced in the past, the retrieval is likely to vary as a function of the particular object that served as the memory probe.

For all these reasons, I suggest that it may be impossible in principle to segregate default knowledge in some domain from associated general background knowledge, and so impossible to define concepts as Machery has done with reference to what is retrieved by default.

3. Do We Really Need to Discard the Hybrid View?

Machery’s proposal that there are three distinct kinds of concepts must be distinguished from a related possibility, namely, that there is a single kind of concept that has different aspects to it. This second view has been proposed by various psychologists. For instance, Osherson and Smith (1981) suggested that concepts can have a core, which is a set of defining features, and characteristic features, which are those that are less essential but are commonly drawn on in identifying things in the world. The core of a bachelor concept might be knowledge of the properties ‘single’, ‘adult’, ‘male’, and ‘human’, but we can look for young men who party, drink, and drive sporty cars to help us find the bachelors. This sort of view has been labeled a hybrid view of concepts.

Machery suggests that to consider different types of knowledge to be part of a single integrated concept, two conditions must be satisfied. The first is that the parts are necessarily linked to one another. According to his definition of being linked, when one of the parts is used in service of one goal, it is possible ipso facto to use the other parts for other purposes. The Heterogeneity Hypothesis allows that different types of representations might be linked, but it also contends that there can be situations in which one can be used and another cannot. According to Machery, the hybrid view, in contrast, should say that all the linked parts are available when any one part is.

I appreciate Machery trying to make explicit what assumptions proponents of hybrid views have in mind, but I am not convinced that this condition is inherent in the hybrid view. I’ve already discussed the context-dependence of retrieval from memory, and the fact that depending on what the probe to memory is, what is retrieved may vary from occasion to occasion. This variation can range from subtle to substantial, depending on the circumstances of retrieval. For instance, when people reflect on their lives, they can retrieve many happy memories or many sad ones, depending on their current mood or recent experiences (e.g. Bower, 1981). Their assessment of the quality of their life experience may fluctuate depending
on what is retrieved (Schwarz and Clore, 1983). Another example is the common occurrence of recognizing a face but being unable to retrieve the person’s name, or recognizing the face but being unable to retrieve where you know the person from. Surely we have a single concept of the individual, Mary, or Mr Rogers, which has various distinct types of information contained within it. Principles of storage and retrieval of information in memory do not suggest that all linked information must be accessible and available for use when some of that information is retrieved.

Machery’s second condition for maintaining a hybrid view is that the parts should be coordinated. That is, the parts of a given concept do not produce inconsistent results, such as inconsistent categorization judgments. He discusses, among other things, examples of statements such as:

Tomatoes are vegetables,

which might be considered true in one sense but false in another. He suggests that these apparently contradictory judgments mean that people consult two different concepts of vegetable (presumably, one centered on a certain functional role in meals, and the other centered on an [incorrect]¹ belief about a botanical definition).

But the right interpretation of such examples is not so clear. To adopt an example of Lakoff’s (1987), suppose that a woman named Susan bears a child who is adopted and raised by Natalie. The adult child can later say either of the following two things, and he can say both at different times without being nonsensical:

Susan is my real mother, but I was raised by Natalie.
Natalie is my real mother, but Susan is my biological mother.

It seems likely that the prototype of a mother includes both the nurturing-and-raising element and the giving-birth element. These two elements can be selectively accessed or focused on and will produce contradictory intuitions. Yet they would be part of a single concept, by Machery’s definition, since they are both part of the prototype. Likewise, it could be that the prototype of vegetables includes information both about typical uses of vegetables in meals and about presumed typical botanical properties. Depending on which a person is focusing on, she will make different judgments about whether a tomato is a vegetable or is not (since its use in meals is usually in savory main courses or appetizers, but botanically, it is the fruit of the tomato plant). These contradictory judgments by themselves don’t imply the existence of different concepts.

Furthermore, if elements of a prototype can give rise to independent and contradictory judgments but still be part of the same representation, as I’ve just

---

¹ Despite a widespread folk belief that the term vegetable has a scientific or botanical definition, it is a purely functional term for foods that have certain characteristic qualities and play certain characteristic roles in meals (see e.g. Malt, 1991).
argued can be true, then it is not clear why it would be important to hold to this criterion more broadly and argue that prototype versus exemplar knowledge (or exemplar versus theoretical knowledge, etc.) must constitute separate concepts if they turn out to yield contradictory judgments. From this perspective, there is nothing in my (Malt, 1994) data that argues against a hybrid concept of water. People may have a single concept that includes knowledge of the scientific definition that water is H2O as well as that mineral water, swamp water, etc., have certain conspicuous non–H2O components in their composition. The former may be central to the prototype and the latter may be knowledge about exemplars, or perhaps each should count as theoretical knowledge, but even if they give rise to contradictory judgments at times, they can be part of the same concept.

Machery suggests that if proponents of a hybrid view reject his two tenets, they would have to explain what is meant by saying that two kinds of knowledge are part of the same concept. I think that is not so hard to provide an answer. One could simply say that elements of knowledge are part of the same concept if they are part of the knowledge about the same set of entities. The core definition of bachelorhood and the identification properties are all part of an understanding of the same set of people. Machery’s own criterion of linkage could also be added: These elements of knowledge are part of the same concept because they are linked to one another. As I’ve just argued, this criterion can be met even though the parts might not always be retrieved together or might result in contradictory judgments if consulted separately.

There are two other observations that argue against considering prototype, exemplar, and theory knowledge as constituting separate kinds of concepts. First, considering the hierarchical nature of knowledge about things, it seems that prototype and exemplar knowledge is inherently inseparable. Consider this hierarchy of abstraction in knowledge about chairs:

- Furniture
- Chairs
- Kitchen chairs
- My own individual kitchen chair.

It is often said that chairs are exemplars of furniture. But kitchen chairs are exemplars of chairs, and my own specific kitchen chair is an exemplar of kitchen chairs. What does the knowledge of kitchen chairs, or chairs in general, consist of? Presumably, it is prototypes abstracted over experiences with instances of the category. Even the representation of my own personal desk chair is abstracted over many experiences with it. Consistent with this last point, most exemplar models likewise assume that ‘exemplar’ representations entail abstraction over different encounters with a stimulus (and those representations will presumably end up containing the frequent or typical elements of the encounters; my desk chair is always red but only rarely tilted back, so the color but not the tilt may be part of my stored representation of it). That means that the same representation is at once an exemplar and an abstraction, and the line between the two is blurred. The only way around that conclusion is to take an extreme exemplar view that says
that all knowledge is an accumulation of memory traces of encounters with specific individuals (e.g. Bülthoff, Edelman and Tarr, 1995). In that case, the notion of having prototypes at all is discarded.

Second, it is not clear how theoretical knowledge could be independent of the kind of knowledge contained in prototypes and exemplars. As Machery explains, what is generally meant by theoretical knowledge in the concepts literature is knowledge that explains why things happen, such as causal and functional relations between one property and another. But this theoretical knowledge is, as the preceding sentence implies, not simply a set of abstractions about relations between variables, as is the case in the laws of physics. Much of it is tied to more concrete knowledge. For instance, theoretical knowledge can include knowledge that birds have wings in order to fly, and that because they can fly, they can nest in trees, or that seed-eating birds have stout bills for cracking open shells, and so on (see, e.g. Murphy and Medin, 1985). In fact, within psychology, the ‘theory theory’ is not generally considered an alternative to prototype or exemplar approaches. It is considered an enrichment of them. It proposes that prototype or exemplar knowledge is not just knowledge of isolated properties but includes this theoretical knowledge about how properties are related (Murphy and Medin, 1985). It would be redundant at best for LTM to contain prototypes and exemplars consisting of isolated sets of properties plus theoretical representations of the same sets of entities containing those properties connected by theoretical knowledge.

In sum, I am not convinced that evidence of how people store, retrieve, and use their knowledge about things in the world implies that we should consider the three types of information to constitute separate concepts. It seems more likely that they are integrated in some shared form of representation.

4. The Notion of Concept is Predicated on the Notion that There Are Categories of Things in the World, But Are There Really Such Categories?

According to Machery, concepts are about categories: Sam’s concept of dogs is a body of knowledge about a specific category, the category of dogs. I agree that for dogs, it is not so hard to find something in the world that could constitute the set of things that are dogs and that can be called a category. In general, for natural kinds, the situation may not be too tricky if there tend to be discontinuities in the distribution of properties across members of domains, creating natural groupings of things (such as birds versus rodents versus fish) that can be considered categories given by the world (e.g. Berlin, 1992; Hunn, 1977; see Malt, 1995 for a review).

What about other domains? The English language tells us that there are sets of things that are chairs versus those that are sofas, cups versus glasses, events that are about carrying versus holding, spatial relations that are about in versus on, and so on. But when we look across languages, we find that these same categories are by no means obvious to speakers of other languages. In fact, they may divide up the same
domain by name in a dramatically different way. For instance, English distinguishes between chairs to seat one person and sofas to seat multiple people. Chinese draws a different lexical distinction, using one word for plain unpadded seating and another for large stuffed seating (whether for one or more people), thereby cross-cutting the English distinction (Gao, personal communication). English distinguishes glasses, generally made of glass, from cups, coming in a range of shapes, sizes, materials (including tea cups, plastic cups, and paper cups). The Russian distinction between stakan and chashka separates tall, straight-sided vessels (whether of glass, plastic, or other materials, thereby cross-cutting English glass and cup) from smaller, round-bottomed vessels with handles for hot drinks (Pavlenko and Malt, in press). These are not just isolated examples; many more can be given (e.g. Malt et al., 1999; Malt and Sloman, 2003; see Malt and Wolff, 2010). And they do not occur just for concrete objects. For instance, English distinguishes between the actions of holding (in which an object’s location is stationary) and carrying (in which the object is moved), whereas Chinese lexically distinguishes similar actions based on how the object is being grasped, not on whether it is moved (Saji et al., 2008). English in and on distinguish spatial relations based on whether they entail containment or support, whereas Korean distinguishes based on whether the contact involves a tight or loose fit (Choi et al., 1999). It seems impossible to know what would constitute objectively defined categories in many domains.

One could say that the concepts each person has are revealed by the words of his or her language, reflecting the categories picked out by their language. This constitutes committing to a strong version of the Whorfian hypothesis, namely, that the language a person speaks substantially determines what her units of thought are. Aside from empirical evidence against such as strong, deterministic version of the Whorfian hypothesis, it makes a person’s concepts something that cannot be identified or defined independent of language (at least, the bulk of them cannot, allowing that there might be some concepts formed that haven’t been lexicalized). Generally, theories of concepts are not intended to have this nature. In fact, Machery refers to concepts that are and aren’t lexicalized, suggesting that he does not want to assume concepts to be formed and revealed primarily through language. In addition, his definition of concept as units of knowledge that are about categories indicates that he expects there to be a language-independent identification of the categories pointing to where the concepts lie.

In sum, saying that concepts are about categories poses problems given these considerations about whether the categories even exist in any language-independent sense in many domains, unless one is prepared to adopt a strong Whorfian position on the relation of language to concepts. On the other hand, if we take away the concepts–are-identified-by-categories explanation of how to find the concepts, then there needs to be some non-circular way of identifying concepts (Sloman and Malt, 2003). For Machery, in fact, it will be necessary to identify instances of three different kinds of concepts, and it is not clear how this will be accomplished. A better solution may be to stop looking for fixed categories that can reveal fixed concepts, and consider that portions of the information stored in LTM can be
Why We Should Do Without Concepts

retrieved flexibly in different combinations depending on the task and goals (e.g. Barsalou, 1987; Malt and Sloman, 2007).

5. Conclusions

In contrast to Machery, I have argued that it may not be possible to separate general background knowledge about some entity from knowledge that can usefully be labeled a concept of that entity. I have also argued that prototype, exemplar, and theoretical knowledge are not as separable as Machery suggests. Finally, I have argued that it is not possible to identify concepts by looking for categories of things in the world because there are often many serviceable ways to group the entities of a domain, reflected in the varied ways of dividing them up that different languages adopt.

Given the concerns about identifying concepts and types of concepts I just listed, how should the field move forward? Although we arrive at the ultimate recommendations through different routes, I agree with Machery’s perspective. Asking which single characterization of concepts—as prototypes, exemplars, or theories—is right will not be a fruitful approach to understanding cognitive representations and processes. The important underlying questions include: What types of knowledge do people have objects, substances, events, etc., in the world? What is the content and character (form) of the information stored in LTM? What pieces of the information stored in LTM are retrieved in the service of navigating the world under different circumstances, and how are these pieces of information used in doing so? What types of knowledge do people have about the words they can use to talk about objects, substances, events, etc., in the world? What pieces of the information stored in LTM are retrieved in the service of producing and understanding language? All of these questions can be addressed by posing them as just stated. There is no need to invoke the notion of ‘concept’ in asking any of them. But it will be possible to pursue these questions fruitfully, discarding the chase after what a concept is, without committing to all the elements of the Heterogeneity Hypothesis.

Department of Psychology
Lehigh University.

References


© 2010 Blackwell Publishing Ltd
B. C. Malt


© 2010 Blackwell Publishing Ltd


Sloman, S. A. and Malt, B. C. 2003: Artifacts are not ascribed essences, nor are they treated as belonging to kinds. *Language and Cognitive Processes [Special Issue: Conceptual Representation]*, 18, 563–82.