

Conceptual representations of exceptions and atypical exemplars: They're not the same thing.

Sandeep Prasada
Dartmouth College

Introduction

Counterexamples play a central role in psychological and philosophical research on concepts. The main argument against the classical theory of concepts is that it is not possible to come up with a set of necessary and sufficient properties that are true of all exemplars of a given kind of thing. Thus, the fact that a dog with three legs is still a dog is taken to be a crucial counterexample to any proposal that states that our concept of dogs is such that having 4 legs is considered to be a necessary property of dogs. Similarly, the fact that a dog that doesn't bark (because he has laryngitis or defective vocal chords) is nevertheless conceived of as being a dog, is taken to be a crucial counterexample to any proposal that states that our concept of dogs is such that the ability to bark is considered to be a necessary property of dogs. As is well known, we can continue in this manner to provide counterexamples against just about any property that is proposed to be conceived of as being a necessary property of a given kind of thing. This type of reasoning constitutes the central argument against the classical theory of concepts, and an important motivation for considering concepts to be prototypes. Counterexamples of the sort considered above have been interpreted as demonstrating that our concept of the kind dog could not represent having four legs as a necessary property of being a dog. Instead, it is posited that four-leggedness is represented as merely a typical property of dogs.

There is an alternative to this conclusion, however. The alternative is that our conceptual systems do, in fact, represent a necessary relationship as holding between being a dog and being four-legged, and that our conceptual systems treat instances of three-legged dogs as exceptions. It is an *empirical* question whether our conceptual systems treat three-legged dogs as evidence that having four legs is merely a typical property of dogs and thus do not represent a necessary connection as holding between the property of being a dog and having four legs, or if our conceptual systems represent a necessary connection between being a dog and having four legs and treat three-legged dogs as exceptions. In this paper, I present evidence that counter to the assumption implicit in all previous research on conceptual representation, our conceptual systems do, in fact, treat certain entities as exceptional exemplars. The notion of an exception requires, (i) that we represent a rule-like or necessary connections between being a certain kind of thing and some property, and (ii) that an exemplar can be considered to be that kind of thing, but nevertheless lack the property in question.

Evidence for necessary connections

In this section, I provide three types of evidence that show that we represent necessary connections between the property of being a dog and having four legs, and the property of being a chair and for being for sitting on, despite the fact that we acknowledge that there are dogs that have three legs and that there are chairs that are not for sitting on. As will be obvious, this type of evidence is available for indefinitely many types of things and properties and I've chosen these examples simply for the purpose of illustration.

Generic statements

Some evidence that our common sense conception of things like dogs and chairs contain necessary connections to certain properties comes from generic sentences such as (1) and (2).¹ Sentences such as these can be used to express thoughts that have contents such as those in (1') and (2'), which clearly display the necessary connection we conceive of as holding between being a dog and having four legs, or being a chair and being for sitting on. These properties are seen as being present simply in virtue of the things being the kinds of things they are. This contrasts sharply with cases in which the relationship between being an instance of a certain kind and a property is conceived of as being contingent, even if the property is typical of the kind of thing in question. Thus, we cannot express the latter type of relationship in a parallel form (3)-(4), and insofar as we may be able to use such sentences, they cannot be understood as expressing the thoughts given in (3') and (4'), which we regard to be false.

- (1) Chairs are for sitting on.
- (1') Chairs, in virtue of being chairs, are for sitting on.
- (2) Dogs are four-legged.
- (2') Dogs, in virtue of being dogs, have four legs.
- (3) *Chairs are wooden.
- (3') *Chairs, in virtue of being chairs, are wooden.
- (4) *Dogs are brown.
- (4') *Dogs, in virtue of being dogs, are brown.

Thus, generic sentences such as (1) and (2) provide evidence that our conceptual systems represent a rule-like or

¹ It is important to note that not all sentences with a surface form of this sort can express the type of relationship discussed here. It is not important for our purposes that there is not a one-to-one correspondence between surface forms of this sort and the types of meanings they express.

necessary relation between being a certain kind of thing and having certain properties.

Explanation of one property in terms of the other

A second type of evidence that shows that we conceive of there being necessary relations between being a certain kind of thing and having certain properties is that we can explain the presence of certain properties in an entity by citing the kind of thing the entity is. For example, if someone points at an entity and asks a question such as (5) or (6), we find that an answer such as (7) or (8) is appropriate and explanatory. Note that though the questions may strike us as unlikely, or a little odd, given that they are asked, there are certain answers that we find appropriate and explanatory, and others not. Only those answers that cite properties or kinds which are conceived of as being in a necessary relationship to the explananda are judged to be appropriate or explanatory, whereas those answers that cite properties or kinds that are in a contingent relation to the explananda are not judged to be appropriate or explanatory (7a-b) (8a-b), even if they happen to mention a property that is typically true of the kind of thing in question.

- (5) Why can you sit on that ? (pointing to a chair)
- (6) Why does that have four legs? (pointing to a dog)
- (7) Because it is a chair.
- (7a) Because it is brown.
- (7b) Because it is furniture.
- (8) Because it is a dog.
- (8a) Because it is brown.
- (8b) Because it is an animal.

These facts show that we conceive of there being necessary relations between being a certain kind of thing and having certain properties and thus can explain the presence of one property by citing the other property. In cases in which we don't understand there to be a necessary relation between properties, the presence of one property cannot be accounted for by reference to the other property.

Redundancy effects in prenominal modification

A third type of evidence that shows that our conceptual systems represent certain necessary relations between being a certain kind of thing and having certain properties comes from redundancy effects in prenominal modification. In cases in which we conceive of there being a necessary connection between a certain kind of thing and a certain property, prenominal modification of the kind by the property engenders a feeling of redundancy (9). In contrast, if the modification is by a property that is conceived of as being only contingently related to the kind, then no redundancy is engendered, even if the property is typical of the kind in question (10).

- (9) A four-legged dog is barking.
- (10) A red apple is on the table.

There is a sense in which our conceptual systems assume that dogs have four legs, and therefore, when *dog* is

modified by *four-legged*, there is a feeling of redundancy. In contrast, our conceptual systems do not assume that being an apple implies being red, even if most apples are red, and therefore, there is no feeling of redundancy when *apple* is modified by *red*. These facts provide further evidence that our conceptual systems distinguish properties that are conceived of as being necessary consequences of being a certain kind of thing from properties which are conceived of as merely being typical of that kind of thing.

The need for exceptions

We seem to have arrived at a paradox. On the one hand, there is evidence that we conceive of there being necessary connections between things like being a dog and being four-legged, while on the other hand, we find no problem in conceiving of the existence of three-legged dogs. There are only three ways out of the paradox. We could reject the validity of the first conclusion, the validity of the second conclusion, or find some way in which we can hold onto both conclusions. There does not seem to be any compelling reason to question the validity of either conclusion. They both rest on strong intuitions. Thus, what is called for is a way of thinking about the data that would allow us to hold onto both conclusions while avoiding the contradiction that seems inherent in them.

One way in which this may be possible is to regard cases in which a putatively necessary property is lacking as being *exceptions* in some sense. Crucially, the exceptional status of these cases, and thus the lack of the putatively necessary property, must derive from something *other* than their being the kind of thing they are. If this were the case, one could maintain that we do conceive of there being a necessary connection between being a dog and having four legs, but that we are willing to allow that dogs might not be four-legged for reasons other than their being dogs. As such, no contradiction would arise as we would not be predicating incompatible properties of dogs. Four-leggedness would be predicated of dogs in virtue of their being dogs, whereas three-leggedness would be predicated of dogs in virtue of something other than their being dogs.

This is, in fact, the solution that our cognitive systems seem to embody. Evidence for this solution comes from sentences such as (11) and (12).

- (11) Dogs are four legged, but/however/though these dogs have only three legs.
- (11') # Dogs are four legged, and these dogs have only three legs.
- (12) Chairs are for sitting on, but/however/though you can't sit on these chairs.
- (12') # Chairs are for sitting on, and you can't sit on these chairs.

These sentences transparently express the fact that we conceive of certain properties as following as a consequence of being a certain kind of thing while recognizing that there could be some instances of the kind in question that are exceptional in that they fail to have the property in question.

The fact that the instances that lack the necessary property are conceived of as being exceptional is highlighted by the fact that the use of a conjunction such as *and*, which does not confer exceptional status to that which follows, engenders a sense of oddness and contradictoriness to the sentences (11') (12'). Notice further that the existence of the exceptional exemplars cries out for explanation. We feel that sentences such as (11) and (12) should be followed by an explanation which gives the reason why the exceptional exemplars lack the property they should have in virtue of being the kinds of things they are. Importantly, neither three-leggedness, nor lack of sitability, unlike the four-leggedness of a dog or sitability of a chair, can be explained by citing the kind of thing the entity is. Thus, whereas (7) and (8) are appropriate answers to (5) and (6), (7) and (8) are not appropriate answers to (13) and (14). Instead, the lack of sitability in a chair or lack of four-leggedness in a dog must be accounted for by citing a reason other than the entity being a chair or a dog.

(13) Why can't you sit on that? (pointing to a chair you can't sit on)

(14) Why does that have three legs? (pointing to a dog with three legs)

Because the lack of four-leggedness of a dog, or the lack of sitability of a chair, are not conceived of as being due to the entity being a dog or a chair, but are conceived of as being due to something other than the entity being a dog or a chair, the three-leggedness of some dogs, or the lack of sitability of some chairs is not treated by our conceptual systems as providing evidence against the conception that dogs, in virtue of being dogs, have four legs, or that chairs, in virtue of being chairs, are for sitting on.

These facts show that the apparent paradox that we started with is exactly that. Our conceptual systems do, in fact, represent necessary connections between being a certain kind of thing and having some property, however, they are also capable of treating instances which lack the necessary property in question as exceptions because the lack of the necessary property in these exemplars is conceived of as being due to factors other than the exemplars being the kinds of things they are.

Implications

Methodological lesson

The foregoing suggests that the common method of trying to determine the necessary and contingent properties of our conceptual representations on the basis of the conceivability of some property being true and not true of exemplars in all possible worlds is flawed because it assumes, wrongly, that there are not other factors that might be responsible for the lack of a property that is conceived of, and represented as, being a necessary consequence of being an instance of that kind of thing. What such intuitions show is that there are relatively few limits on what is conceivable, however, such intuitions do not display the manner in which things becomes conceivable. The four-

leggedness of dogs is conceived of as simply being true in virtue of our conception of dogs, whereas a property such as the three-leggedness of dogs is conceivable only by imagining some factor or set of factors other than the entity being a dog in virtue of which the dog is three-legged. Conceivability, as such, is too weak a tool to investigate the structure of our conceptual representations because it does not distinguish what is conceived *simpliciter* from that which is conceivable but instead involves combination. Conceivability does, of course, put important constraints on theories of conceptual combination.

Necessary connections are not established on purely empirical grounds

If, as has been argued, we do establish necessary connections between concepts, it raises the important question of how it is that we do this. Clearly such connections cannot be established on empirical or statistical grounds because, as is well known, empirical and statistical evidence can never underwrite conclusions of necessity. Before speculating about how we do establish necessary connections, I want to present explicit evidence that the necessary connections that are established between concepts are not established upon the basis of solely statistical or empirical means.

The contrast between sentences such as (2) (reproduced here as (15)) and (16) helps make this point. As discussed earlier, (15) expresses what we conceive of as being a true generic fact about dogs. Because the statement expresses what we conceive of as a necessary connection between the property of being a dog and being four-legged, the fact that there exist some dogs that are three-legged for reasons other than being a dog does not render the sentence false. In contrast, (16) is an extensional statement that expresses a particular fact about dogs which is false because not all dogs are four legged, some are three-legged. Furthermore, even if there did not exist any three-legged dogs, it would be impossible to make a statement such as (16) without qualifying it in some way, for example, as in (16'). This is because (16), unlike, (15) is a statement about the extension of the concept DOG, and one could never make a statement about every instance of a given kind as there are, in principle, indefinitely many instances of any kind.

(15) Dogs are four-legged.

(16) All dogs are four-legged.

(16') All the dogs, observed up to this point in history, have been four-legged.

(17) Ninety percent of dogs are four-legged.

Similarly, all explicitly statistical quantified statements such as (17), can only be made with qualification, as they can only be truthfully asserted of some particular set of exemplars which exist at some particular sets of times and places. In contrast, (15) can be said without qualification, because it expresses a fact that we conceive of being true of dogs in simply in virtue of being dogs, rather than a fact about some particular set of exemplars which we conceive

of as having the properties they do in virtue of being dogs as well as all kinds of contingent facts that are particular to the instances in question. This contrast between our understanding of generic and extensional sentences provides explicit evidence that the basis upon which we establish the truth of the necessary connections expressed in generic sentences cannot be derived solely by statistical and empirical means.

A formal system of knowledge for common sense conception?

So, how is it that necessary connections between concepts are established? It seems that there are two options left open. The first is that the content of all of these connections (e.g. between dogs and four-leggedness, chairs and sitability, etc....) are innately specified and somehow triggered given appropriate experience. This, however, strikes me as extremely implausible. Alternatively, what may be innate is a *formal* system of knowledge in terms of which we form common sense concepts by analyzing the things around us in terms of this formal system of knowledge. For the purposes of illustrating how such a system might work, let us consider how the formal system of knowledge that constitutes knowledge of Euclidean geometry is put to use in reasoning about actual things in the world. If we are standing at the corner of a rectangular park, we know that it will *necessarily* take us less time to traverse the diagonal to get to the opposite corner than walking along two sides of the park (assuming that we walk at a constant speed and the park is relatively even). Here we have an example of how formal knowledge, which is not about parks or walking, can nevertheless be applied to establish a fact about parks and walking. Furthermore, the necessity that we know to hold concerning the relative sizes of the sides of geometrical figures is transferred to our reasonings about actual distances of paths in parks as long as we are justified in applying the formal system of knowledge to these aspects of the real world. Thus we see that it is possible for us to make statements about what is *necessarily* the case, either when we are making statements within a purely formal system of knowledge, or when we are using that system to reason about certain aspects of the world. This suggests that just as there is a formal system of knowledge that underlies our intuitions and reasonings about space, there may also be a formal system of knowledge that underlies our abilities to form common sense conceptions of kinds of things and allows us to come to conclusions about what kinds of connections between concepts are necessarily rather than contingently true. Furthermore, if it is the case that there is a formal system of knowledge that underlies our common sense conception of things and our conceptions of necessary connections between concepts, then we also have a way of accounting for the normative aspect of the necessary connections -- (e.g. dogs, in virtue of being dogs, *should* have four legs). Formal systems of knowledge such as those that constitute our knowledge of logic or language specify how we *should* reason or structure our sentences, not how we actually do, as many other factors affects performance. For a sketch of

what this formal system of knowledge might look like see Prasada (1998a; 1998b).

Representing necessary connections/exceptions.

The facts discussed above suggest that we represent a rule-like necessary relation between being a dog and having four legs, but allow for there to be exceptions in particular cases. This suggests that the four-leggedness of dogs should be represented by a *default* process which can be overridden if knowledge about a particular dog dictates otherwise. Defaults are commonly used in psychological research, however, they are usually used to represent probabilistic information. Thus, the default value for the colour of an apple might be "red" based on the fact that most apples are red. While this is a perfectly legitimate thing to do, there are important qualitative differences between representing probabilistic defaults, and rule-based defaults that involve necessary connections. First, as discussed above, probabilistic defaults are updatable on the basis of further experience, however, necessary connections are not revised in the face of evidence that suggests that they the property in question is not displayed by some instances of the kind in question. Second, probabilistic defaults need to be the most frequent value of the property in question, however, this need not be the case for rule-based defaults. It is even possible to have the majority of instances not display the default value. For example, (18) which has the meaning in (18') is considered to be true even though we know that the vast majority of chickens never live to the age of x because they are raised and killed on poultry farms.

(18) Chickens live to be x years of age.

(18') Chickens, in virtue of being chickens, live to be x years of age.

Third, as the difference in redundancy effects of (9) and (10) show, rule-based defaults can be, and are, assumed to be true with perfect certainty unless evidence for a different value is available, whereas, probabilistic default can never be assumed with perfect certainty. Fourth, it seems that rule-based defaults can only be established through the use of a formal system of knowledge, however, probabilistic defaults can be acquired on the basis of statistical information in the absence of a formal system of knowledge.

There are close parallels between the organization of conceptual knowledge and the organization of aspects of linguistic knowledge. For example, inflectional morphology also makes use of default rules that are formed through the use of a formal system of knowledge, and exceptions which are particular cases which are not formed through the use of a formal system of knowledge but are stored in memory and can block the application of the default rule (Pinker & Prince, 1988; Kim, Pinker, Prince, & Prasada, 1991; Marcus, Brinkmann, Clahsen, Wiest, & Pinker, 1995). Thus, there is independent evidence that the mind has the representational resources being proposed for

the representation of conceptual knowledge (see also, Pinker & Prince, 1996).

Conclusion: Typicality, atypicality, necessary connections, and exceptions.

We are now in a position to explicate the title of this paper. Typicality and atypicality are extensional notions. What is typical or atypical can only be established by observing the frequencies with which various properties of a given kind of thing actually appear in its extension. In contrast, exceptions are instances of a given kind that lack a property that is conceived of as being a necessary consequence of being that kind of thing. As such, typicality and exceptionhood are orthogonal to one another -- exceptions need not be atypical, and atypical instances need not be exceptional. However, since an exceptional property such as the three-legged of dogs, is not due simply to the entity being a dog, but in virtue of its being a dog *and* certain other particular properties that happen to be true of the instance, it is usually the case that the rule-based default property is most frequent because there is usually no reason why other circumstances that prevent expression of this property should be systematically present for instances of the kind in question. That is why we usually assume generics such as (18) to be true of the majority of instances unless we are explicitly made aware of a systematic reason why the majority of instances would not display the rule-based default property (Prasada, 1999).

References

- Kim, J.J., Pinker, S., Prince, A., & Prasada, S. (1991) Why no mere mortal has flown out to center field. *Cognitive Science*, 15, 173-218.
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiest, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, 29, 189-256.
- Osherson, D. N., Stern, J., Wilkie, O., Stob, M., & Smith, E.E. (1991). Default probability. *Cognitive Science*, 15, 251-269.
- Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193.
- Pinker, S., & Prince, A. (1996). The nature of human concepts/evidence from an unusual source. *Communication & Cognition*, 29, 307-362.
- Prasada, S. (1998a). Formal aspects of common sense conception. <http://mitpress.mit.edu/celebration>
- Prasada, S. (1998b). How to form stable concepts: The formal aspect of common sense conception. Unpublished manuscript, Dartmouth College, Hanover, NH.
- Prasada, S. (1999). Names for things and stuff: An Aristotelian perspective. In R. Jackendoff, K.Wynn, and P. Bloom (Eds.), *Language, Logic, and Conceptual Representation: Essays in honor of John Macnamara*. Cambridge, MA: MIT Press.