Can prototype representations support composition and decomposition?

Lila R. Gleitman
University of Pennsylvania, Rutgers University

Andrew C. Connolly
Dartmouth College

Sharon Lee Armstrong
LaSalle University, Rutgers University

ACKNOWLEDGMENTS

Acknowledgements: We thank Jerry A. Fodor and Henry Gleitman, who were co-authors on the research reviewed here, and Dan Osherson for helpful conversation. We also thank James Hampton and an anonymous reviewer for penetrating comments that led to significant revisions of the manuscript. Writing of this chapter was partially funded by NIH grant R01-HD37507 to Lila R. Gleitman and John C. Trueswell and by NIH grant R01-MH67008 to Sharon L. Thompson-Schill. Authorship is in reverse alphabetical order.

> Tho' a particular colour, taste, and smell, are qualities all united together in this apple, 'tis easy to perceive they are not the same, but are at least distinguishable from each other.
>
> [David Hume, Treatise, Book I]

> **Red** is any of a number of similar colors evoked by light consisting predominantly of the longest wavelengths of light discernible by the human eye, in the wavelength range of roughly 625–740 nm.
>
> [Wikipedia]

> **Red hair** (also referred to as auburn, ginger or titian) varies from a deep orange-red through burnt orange to bright copper.
>
> [Wikipedia]

Concepts have always been central to theories of cognition, and yet even in the modern era there is considerable disagreement about just what they are. Early theories often treated concepts as icons (something like a mental image, e.g., Hume, 1739). A more recent view has been that a concept is a definition, something like a set of necessary and sufficient conditions for a thing to fall under the concept, e.g., Frege, 1892; Miller & Johnson-Laird, 1976). So, for example, to satisfy the APPLE concept ("to be an apple…"), a thing must exhibit *roundness*, *edibility*, *redness*, *fruitiness*, and so on. Such theories have been touted partly on the grounds that they are said to yield a classification of the objects, properties, events, and relations in which human cognition traffics, and to explain the resemblances that hold across the set of concepts. For instance, apples resemble fire engines by sharing the characteristic of *redness*, they resemble balls by sharing *roundness*, and peaches by being *edible*. However, this theory of concept structure has lost much of its popularity, largely because apples, though most often red, are not necessarily so (there are Granny Smith apples and spoilt brown apples); they are usually but not necessarily round (there are oval and squashed apples); and they are occasionally inedible or at least indigestible (remember Eve and wicked step-mothers). In response to these manifest differences among category members, the dominant positions in psychology and cognitive science today hold that concepts are prototypes (something like a set of weighted properties which things that fall under the concept typically have or are believed to have; Rosch, 1978; Rosch & Mervis, 1975; Smith & Medin, 1981; Prinz, 2002).[1]

In the present paper, we review two kinds of experimental evidence from our laboratories that challenge the adequacy of prototypes for representing human concepts. First, we will review experiments suggesting that prototype theory does not distinguish adequately among concepts of maximally variant types, such as formal (e.g., ODD NUMBER) versus natural kind and artifact (e.g., APPLE and HOCKEY) concepts (Armstrong, Gleitman, and Gleitman, 1983). Second, we will review a more recent experimental line demonstrating how theories of conceptual combination with lexical prototypes fail to predict actual phrasal interpretations, such as language users' doubts as to whether Lithuanian apples are likely to be as edible as apples (Connolly, Fodor, Gleitman, & Gleitman, 2007). We should emphasize at the outset that these studies in no way challenge the view that many concepts have prototypes. Rather, we take our findings to support a distinction between concepts' *having* prototypes and *being* prototypes. Before introducing the experiments, we want to sketch the relation, as we see it, between the varying views of what concepts are and the problem of compositionality that is the more general topic of the present volume.

[1] Throughout we will follow the notational convention of referring to properties or features using *italics*, to mentioned words and phrases using 'single quotes', and to concepts using all CAPS. For example, the concept RED represents or denotes the property of *redness*, for which the English word is 'red'. For utterances, we use "double quotes."

Features, prototypes, and the problem of compositionality

Most people who have thought about the meanings of common words assume that the vast majority of them are complex, composed by conjoining several simpler (primitive) concepts, sometimes called features or attributes. We regard this view as the Humean status quo, though we do not necessarily endorse it. After all, as there are only finitely many words in a speaker-listener's repertoire, it is possible to believe that our elementary concepts are at about the same level and grain as the word/morpheme (e.g., Fodor, 1981), though to be sure in a minority of cases languages draw their lexical boundaries in different places. Nevertheless the idea that lexical-level concepts are compositions out of simpler formatives has been attractive for several reasons, among them the possibility of reducing the number and types of hypothesized mental primitives, and – as we remarked earlier – explaining the resemblances among concepts as a matter of feature overlap (why and how apples and peaches are more alike than either of these is to, say, pencils). [2]

While there is debate as to whether lexical-level concepts are compositional, it is a truism that understanding language requires the compositionality of word meanings (Frege, 1914). Whatever the concepts expressed by the words 'apple', 'red', etc., the standard view is that these must be the ultimate constituents of compositional mental representations for the meanings of phrases such as 'red apple', 'purple apple', and 'Chinese apple'. It follows that these complex (phrasal) representations resemble each other in meaning by component overlap: They have something in common, namely that the concept APPLE is a constituent of each. According to conventional wisdom, then, we can understand each of these phrases because we have the concept referred to by 'apple', which in turn is composed of *roundness, edibility,* and so forth. In sum, the compositional properties of phrasal concepts would seem to be necessary so as to explain the productivity and systematicity of thought, and thus by-and-by, how we are able to produce and understand more than one word at a time. Thus, a theory of concepts must satisfy the compositionality constraint:

---

[2]. To the extent that something like Fodor's view is correct, APPLE is an elementary concept and composition applies to it *tout court* as a constituent of complex concepts such as PURPLE APPLE and POISONED APPLE. This position in no way denies our knowledge about characteristics shared (necessarily or probabilistically) by apples, denying only that this knowledge is constitutive of or (in some cases) even relevant to the concept itself.

The compositionality constraint (CC): The meaning of a complex expression is determined by the meanings of its constituents plus the syntactic rules used to combine them. [3]

Consider as a further example the descriptions of ‘red’ and ‘red hair’ in the prefatory quotes to this chapter. Classical theories of concepts hold that RED denotes a certain range of hues without regard to the frequency with which any one component hue might have been observed or to its position within the range of reds (whether at the center around wavelengths 700 nm or at the margin where red meets, say, brown). ‘Red hair’ is compositional because its being a noun phrase is entirely determined by the fact that ‘red’ is an adjective, ‘hair’ is a noun, and AN structures are NPs in English; its meaning is fully determined by the fact that ‘red’ expresses the property *red* and ‘hair’ denotes the range of follicular mammalian skin coverings, together with the principle that AN representations denote the intersection of the As and the Ns. Thus ‘red hair’ means *hair that is red.* The crucial assumptions are that

(i) a concept expresses the full range of variation allowed for instances that fall under it;

(ii) the syntactic and semantic properties of the constituents of complex concepts are context independent.

In virtue of these strictures, the classical concept descriptions are hard-edged (all or none): ‘red’ applies to all and only *redness*, ‘hair’ to all and only *hair*, and their combination covers all and only the cases that would be correctly considered as actual or possible members of the extension of the phrase, ‘red hair’. Such a theory affords the flexibility and abstractness needed to account for all interpretive possibilities. For example, you wouldn’t be likely to guess, upon learning that someone you haven’t met has red hair, that the particular hair-hue was vibrant fire-engine red. But as anyone who has been to Manhattan’s East Village can attest, this is certainly possible. Conceptualization has to allow for bizarre entities and events. For the same reasons, context independence properly bounds conceptual combination, in the sense of assuring

[3] . That the principle of compositionality is exceptional in several respects is well known, e.g., former senators are not senators, decoy ducks aren’t ducks and stone lions aren’t lions (see Kamp & Partee, 1995 for discussion). Another class of difficulties has to do with the context-dependence of indexicals and pronouns such as ‘here’ and ‘it’. Exactly what the principle of compositionality requires of meanings and concepts is, for these and related reasons, a matter of open debate. To be sure, the arguments made in this paper would be strengthened if a strong form of the compositionality constraint turned out to be true, especially one requiring that **reverse compositionality** holds as well (e.g., Fodor, 1998; Fodor & Lepore, 2002; Pagin, 2003; see Robbins, 2005, and Patterson, 2005, for counter arguments). However, the findings we report and our interpretations of them require only that over an important (and indefinitely large) subset of phrasal-level concepts, the standard formulation of CC holds.

for example, that ‘red cheeks’ not be interpretable as *cheeks that are green* or *elbows that are red.*

Nevertheless, the fact that this theory does not address the question of plausibilities may be taken as a defect. Consider again the concept RED. The colors of cherries and apples seem “better,” more typical, or more central instances of reds than do the colors of human hair, but the classical theory fails to account for such effects. In fact, under combination the typicality facts in this case reverse such that the expected hue for a human redhead (or a red fox) is not prototypical red at all.[4] This appears to be context dependence *par excellence,* an ominous potential violation of (ii), itself a non-negotiable property of the classical theory. Insofar as traditional concept theories attend to such issues, it is by assuming that the recovery of meaning from the concept descriptions and their combinatorics is only a first step in the real business of everyday understanding. A second, and also crucial, step is the application of a further set of pragmatic-inferential processes that draw on general knowledge of the world. These latter processes supply the plausibility facts.

An alternative set of views, collectively termed **prototype** approaches, were introduced in the mid 1970’s by Eleanor Rosch and her colleagues, and today these and related probabilistic perspectives dominate theories of concepts in the psychological literature (Rosch, 1973, 1978; Rosch & Mervis, 1975; Smith & Medin, 1981; Lakoff, 1987; Hampton, 1993; Barsalou, 1999; Murphy, 2002; and Prinz, 2002). The general properties of prototype theories are, as the saying goes, too well known to require much introduction, but we will mention a few highlights that are important for our further discussion.

Rosch and later theorists allied with the prototype tradition assert that concepts are internally organized roughly as a set of weighted components or features, held together in a family-resemblance structure (following earlier suggestions from Wittgenstein, 1953). What this means is that the set of features that comprise a concept and establish membership in it is not all-or-none, but graded. The conditions for membership in a concept are satisfied for some item

---

[4] It is an embarrassment to any theory of concepts we know of (except the one that says ‘red hair’ is an idiom) that several typical hues of human red hair (auburn, ginger, titian) as referred to by Wikipedia turn out to be hues that the same source describes as among the browns rather than among the reds. (See the prefatory quotes to this chapter) Another embarrassing instance is that the dog often called a Miniature Collie (because it looks just like a miniature collie) turns out to be another breed altogether (the Sheltie) and no Collie at all. Such cases are common. This argues either that the compositionality constraint is too strong (see again ftnt. 3) or that there are more phrasal idioms than you might have thought.

when it exhibits a number of these features, but by no means necessarily all of them. As well, the features themselves may bear different weights within the category such that an item's partaking of the heavily weighted features also counts toward its position within the category. Centrality or typicality in the category, for any potential item, is computed as some composite of the number and weighting of the concept features the item exhibits, with high scores being the measure of typicality of an instance.

Consider for example the concept BIRD. Typical properties of birds presumably are that they *fly*, *have wings*, *feathers*, *claws*, *lay eggs*, and have certain *body proportions*. Ostriches lack a heavily weighted property of the bird category (they do not fly) which relegates them to marginal status. Pelicans lose some typicality points too owing to their ungainly proportions but at least they can fly (though awkwardly, it is said). A robin, embodying many heavily weighted bird-properties – a good flier of graceful birdy proportions – is a central or prototypical member. In sum, the all-or-none property of the classical theory is relaxed under prototype theory with category membership a matter of degree. On the matter of conceptual combination, the prototype views further part company with tradition by allowing some degree of context dependence. Thus while the classical analysis of a concept such as HAIR makes no distinction as to color, a HAIR-prototype may very well incorporate this very distinction, representing and weighting the typical range of mammalian or human hair colors. Such an analysis provides the first components of an explanation for the differential hue-expectations for hair, cheeks, and apples: the *redness* range for hair could be specified along a color dimension within the representation of HAIR, the *redness* range for cheeks within the representation of CHEEK, and so forth (Katz, 1964; Kamp & Partee, 1995; Osherson & Smith, 1981).

Summary and prospectus

As we see it, the fundamental idea behind prototype theory is, as much as possible, to build the facts about typicality and plausibility directly into the representations of the concepts themselves and, consequently, into the combinatorics for complex concepts and the phrases that express them. Though most prototype theorists acknowledge that a partly separate inferential-pragmatic theory incorporating our general knowledge of the world is a crucial part of human understanding, the idea behind the prototype approach is to develop a theory in which the concept representations themselves will bear as much of the interpretive burden as possible. In the experiments that we now discuss, we will explore the adequacy of prototype theory in terms of two questions: (1) Are lexical concepts represented as prototypes? and (2) to the extent that they are, could such prototypicality representations compose lawfully into phrases?

Part I. Are concepts prototypes?

An extensive body of empirical research seems to provide evidence for the psychological validity of the prototype position. For example, the left hand column of Table 1 shows two everyday superordinate categories—FRUIT and VEHICLE—and some exemplars of each (e.g., apple, fig, for FRUIT). In an influential study, Rosch (1975) asked subjects to indicate how good an example each exemplar was of its category by use of an appropriate rating scale. It is worth quoting part of the instructions that were used in this experiment:

> "This study has to do with what we have in mind when we use words which refer to categories... Think of dogs. You all have some notion of what a 'real dog', a 'doggy dog' is. To me a retriever or a German Shepherd is a very doggy dog while a Pekinese is a less doggy dog. Notice that this kind of judgment has nothing to do with how well you like the thing... You may prefer to own a Pekinese without thinking that it is the breed that best represents what people mean by dogginess. On this form you are asked to judge how good an example of a category various instances of the category are..." (Rosch, 1975, p. 198).

Notice then that the instructions assent to the membership of both these animals in the category DOG but then equivocate ('doggy dog') about how the subject is to make distinctions among the class members, a matter to which we will return later. In any case, subjects now rated each instance of each category using response templates of this approximate sort:

FRUIT

apple

1 2 3 4 5 6 7

good poor

It turns out that given instructions of this kind and stimulus presentations of this form, people will say that apples are commendable examples of FRUIT, and deserve the lower-numbered ratings (that is, the 1's and 2's rather than 5's and 7's), while figs and olives are poor exemplars and deserve the higher-numbered ratings. Moreover, agreement among subjects is remarkably high with split-half correlations between subject group rankings of approximately .97.[5]

The left-hand column of means in the top half of Table 1 shows a successful replication of these effects by Armstrong, Gleitman, & Gleitman (1983, henceforth AGG). Notice that if these subjects were successfully ranking, e.g., the apples and the figs for their membership in the category FRUIT, their performance appears to be incompatible with the classical theory of concepts. This is because that theory holds that membership in a category depends on having vs. not having some specified necessary and sufficient set of features. An item lacking any one of these features would be out of the category altogether (off the bottom of the scale at 7+), while those having them – from apples to olives – would be equally FRUITY (rated uniformly as 1's). In short, a coherent ranking of concept membership should have been impossible if subjects thought that membership in the FRUIT and VEHICLE categories is all or none.

Rosch and her colleagues interpreted these findings as evidence that category membership is graded and thus inconsistent with the standard classical theory. Dozens of experiments in the concept literature during the subsequent 30 or so years have achieved the same kinds of results and thus seem to bolster this interpretation of concept structure in general. For instance, asked to name members of a category, subjects reliably list the more stereotypical ones first; subjects can name more attributes/features and agree on more of them for prototypical members than for marginal members (Cree & McRae, 2003); and subjects respond faster in a verification task to items with high exemplariness ratings (e.g., 'A robin is a bird') than to those with lower ones ('An ostrich is a bird') with appropriate controls for word frequency (Rips, Shoben, & Smith, 1973; Rosch, 1975).

Seemingly related typicality effects are found again and again in almost every domain of human existence and relation. There have been prototype-like analyses of cultures (Sinha,

[5] There has been some objection to the adequacy of the split-half statistical procedure used by Rosch and colleagues to assess cross-subject reliability in these ratings tasks, though this method was state of the art at the time these investigators used it (see Barsalou, 1987). Armstrong et al (1983) and the report thereof in the present paper continue to use this relatively weak assessment tool, for comparability with prior findings and reports.

2002), of social structures and groupings (Hess, Pullen, McGee, 1996), of the profiles of drinkers and smokers (Spijkerman & van den Eijnden, 2004), of love (Aron & Westbay, 1996), and of mobile telephone users (Walsh & White, 2007), to name just a few. In the light of the reliability and domain-generality of such findings, one might well conclude, as have many cognitive psychologists, that the psychological validity of the prototype descriptions of human concepts and categories has been demonstrated beyond reasonable doubt.

TABLE 1 ABOUT HERE

But perhaps these victories for prototype theory have been wrested too cheaply. The basis for claiming that certain categories have a prototypical, nondefinitional, feature structure has always been the finding of graded responses to their exemplars in various experimental paradigms. But this is only half of the required demonstration, for the truth of the contrapositive has been left implicit rather than being tested directly: If you believe certain concepts are nondefinitional because of graded responses to their exemplars, that must be because you also believe that if the categories were all-or-none in character, the graded responses would not have been achieved. Thus a necessary part of the proof requires finding some categories that do have definition-like, categorical, descriptions, showing as well that subjects patently know and assent to these definitions; and, finally, showing that these new category types do not yield the graded outcomes.

AGG attempted to carry out this further part of the required experimental program by repeating some of Rosch's original procedures, but adding putatively well-defined categories (EVEN NUMBER, ODD NUMBER, FEMALE, and PLANE GEOMETRY FIGURE) to those that previously had been studied (such as FRUIT, FURNITURE, VEHICLE, SPORT). One replication was of the exemplar-rating procedure. In addition to the previously studied categories, they added the formal ones, and presented all of them using the original instructions devised by Rosch (reproduced in the present article on p. 000), and the original response templates, e.g.:

ODD NUMBER

501

1 2 3 4 5 6 7

good poor

The findings for ODD NUMBER and FEMALE are shown in the bottom half of Table 1. The exemplars of the well-defined categories elicited differential ratings much as had categories such as FRUIT, and at high levels of reliability, (rank order split-half correlations were .94, .81, .92, and .92, for EVEN NUMBER, ODD NUMBER, FEMALE, and PLANE GEOMETRY FIGURE, respectively). Keep in mind that these subjects were being asked, for example, to distinguish among odd numbers for their oddity, and common sense asserts one cannot do so. But the subjects could and did: They judged 3 a better ODD NUMBER than 501 -- and Mother a better FEMALE than comedienne.

One trivializing response to these findings has been that the subjects were knowingly responding in different ways to the two types of stimuli, despite the task instructions. For the categories studied by Rosch, perhaps, they answered with their true assessments of the prototypical organization of, say, FRUIT; but for the formal categories, they responded as though answering silly questions with silly answers. To assess this interpretation, AGG also replicated earlier verification tasks (Rips, Shoben, and Smith, 1973; Rosch, 1975) which are not as susceptible to such a disclaimer because the requirement for speeded responses discourages self-conscious consideration of category types. Subjects were presented with sentences of the form *An A is a B* in which *B* was a category of which *A* was said to be an exemplar, for both formal and everyday categories. Half of the sentences were true (e.g., 'An orange is a fruit') and half were false ('An orange is a vehicle'). The subjects' task was to answer (by a key press) true or false to each such statement as rapidly as possible. Items with higher exemplariness ratings were verified more quickly than those with lower ratings. This was true for both the putative prototype categories studied by Rips et al. (1973) and Rosch (1975) and for the formal categories (e.g., 'A circle is a plane geometry figure' *vs*. 'A circle is an odd number').

Some responses to these demonstrations have been to the effect that perhaps concepts such as ODD NUMBER and FEMALE are prototype-like in the same manner as FRUIT and

VEHICLE (e.g., Lakoff, 1987). Exactly what would be implied by such a move is hard to fathom, for clearly notions like ODD NUMBER have definitions that are known to their users and pattern within a theory of arithmetic whose organization cannot be rendered in prototype theory.

AGG reasoned that the many demonstrations of prototype theory are relevant to the exemplariness of instances of a concept rather than to membership (see also Rey, 1983, for an important discussion). Good exemplars exhibit the surface features that are most frequently associated with a concept and thus they are easily recognizable as members, but this recognition function need not bear straightforwardly on the issue of category membership. Mammals that swim (such as whales) and albino tigers are atypical and thus easily misclassified, but in the end they are nevertheless whales and tigers respectively. And similarly, as Wanner (1979) showed, people's judgments of prototypical prime numbers are those that go through certain heuristic decision procedures easily. Indeed when we examine the instructions in the Rosch exemplar rating task, we see a number of confusing and perhaps contradictory phrases ("a very doggy dog," "the breed that best represents what people mean by dogginess," "how good an example of a category various instances of the category are") that sometimes allude to the category itself and sometimes to attributes of its members.

AGG explored this distinction between concept membership and concept exemplariness in a final experiment. Subjects were asked straight out "Does it make sense to rate items in this category for *degree of membership* in the category?" Subjects clearly distinguished formal categories from prototype categories in this paradigm: 100% judged that it was nonsensical to rate instances of EVEN NUMBER, ODD NUNBER, and PLANE GEOMETRY FIGURE as to "how good" they were as members of their respective classes, and a substantial percentage (86%) said the same of FEMALE. Percentages for "everyday" categories were much lower, with FRUIT, SPORT, VEGETABLE, and VEHICLE being judged all-or-none by 43, 71, 33, and 24 percent of subjects, respectively.

A crucial further task was then presented to these same subjects: They performed the Rosch exemplar ratings task for instances of all the categories including the formal ones. The results are shown in the bottom half of Table 1. Note particularly the response characteristics of those subjects (those in the right-hand columns of the table) who had previously averred that rating degree of membership in the specified category made no sense at all, e.g., those who said that being a FRUIT or being a FEMALE was all-or-none. Yet subsequently presented with the Rosch instructions to distinguish “doggy dogs” from less doggy ones, these subjects provided differential rankings, judging “really odd odd numbers” better than “less odd ones.” For instance they rated 3 better than 501, among the odd numbers. Indeed it is true as Table 1 shows that these subjects used less of the scale for the formal categories than they did for the everyday categories (rating no odd number as worse than a 2, on a scale of ODDity that ranged from 1 to 7). But even so! No person who knows and states that all odd numbers are equally odd should rate some of them more odd than any others, even by a smidgen.

Arguably the subjects in the two parts of this experiment did not contradict themselves at all, despite first appearances. Rather, their differential behavior reflects the fact that the instructions assigned them two different tasks. In part 1 of the experiment, they were asked to (and did) consider the issue of category membership which in the case of formal categories was judged to be all or none. But in part 2 they were asked to consider the issue of exemplariness of an item for its category, not at all the same thing. The implications of this disconnect present problems for the usual interpretation of the tasks asking subjects to rate items within categories. It cannot be assumed that the results of ratings tasks reveal concept membership or structure.

In sum, the psychological literature contains scores of demonstrations that people distinguish between typical and marginal members of a concept or class. Plausibly the prototypical instances are those that exhibit several properties (features) that are quite regularly observed in members of that category. These properties represent things we know are true of most, e.g., tigers we have seen, and therefore serve as rough and ready clues that some new creature we are viewing is probably also a tiger. Indeed it might be very difficult to recognize a tame, albino, three-legged, toothless tiger as a tiger, and easy to judge that it is a sorry example of a tiger. All the same, it is likely an error to conclude that these surface features are constitutive of the TIGER concept. AGG, in the work just reviewed, studied this difference between *having a prototype* and *being a prototype* with formal concepts where the distinction arises starkly: Nine is a prototypical odd number, probably because in addition to being an integer not divisible by two without remainder it is low in cardinality, familiar, and “looks primey” (cf., Wanner, 1979). All the same, to claim that 9 is odder than 99 is to be risibly ignorant of the facts about addition and subtraction, systems on which the concepts of NINE and NINETY-NINE are defined. These exclude factors such as low cardinality.

Part II: Do prototypes compose?

In the previous section, we raised some doubts as to whether everyday human concepts are represented prototypically as much theorizing in psychology has suggested. The argument was an indirect one for in fact all that was shown was that formal concepts with known internal structure exhibit typicality effects that are at odds with that structure in the psychological laboratory. This bears relevantly on the adequacy of typicality (or measures of typicality) for assessing the structure of concepts, but does not by itself force us to reject (or accept) some variant of a prototype theory. We now ask the same question about concept representation in a more direct way: Suppose that concepts really were prototype-style mental representations. Could these representations underlie our actual interpretation of the phrasal concepts, thus satisfying the compositionality constraint? For example, assuming the prototypes of RED and APPLE as the constituents of the complex concept RED APPLE, we must assume that these compose to a prototypical apple that is prototypically red. But is this the desired outcome, the one that comports with how people understand phrases expressing these complex concepts?

A pessimistic argument from Fodor (e.g., Fodor & Lepore, 1996) considers as a test case the phrase 'pet fish'. [6] Perhaps there is a widely shared image that comes to mind for 'pet fish', something like the guppies that typically inhabit home aquariums. This example, among countless others, shows us that complex concepts can have stereotypes. Prototype theory says that the concept PET is itself represented as the set of stereotypic properties of pets and FISH is represented as the stereotypic properties of fish. Compositionality under prototype theory thus entails that to understand the linguistic expression 'pet fish' we must compute the prototype as a

[6] Several commentators have claimed that the failure of the prototype theory for ‘pet fish’ may be dismissed on grounds that this phrase is idiomatic in the sense that the stereotype is set not through composition over stereotypes, but rather through direct experience of pet fish in the world (a.k.a., “extensional feedback”, see Hampton, 1988; Rips, 1995). But notice that the same arguments Fodor and Lepore (1996) made for pet fish can be made just as well for brown cows. For it could very well be the case that brown cows are not prototypically brown (rather, they might always or usually be a reddish brown) or prototypical cows (rather, they might be especially large, robust, or cantankerous). Furthermore, notice that the combinatorics of PET FISH are exactly as one would expect in the classical theory, i.e. something that is necessarily a PET and a FISH. Suppose that the king of South Dakota kept a Great White Shark in his aquarium for the amusement of his guests and himself. Then this creature is a pet fish, though hardly a stereotypical one. In contrast idiomatic phrases fail the test of compositionality which is why they are called idioms in the first place. For instance the idiomatic green thumb is neither green nor a thumb, but rather a person who disports in the garden. On the compositional interpretation, a green thumb is something that is both green and a thumb. To count as an idiom, the phrase ‘pet fish’ would similarly have to have two disjoint interpretations, and it does not.

function of the prototypes for 'pet' (i.e., something like a golden retriever) and 'fish' (i.e., something like a trout). Given these prototypes, the derivation of the prototype for 'pet fish', which is neither dog-like nor trout-like, appears on its face to be intractable (though Hampton, 1988, attempts such a derivation for 'pet fish' in his composite prototype model which we will discuss later in this section). If this problem generalizes, it presents a major challenge for prototype theory. We believe it does generalize and we next review experimental evidence in support of this position from Connolly, Fodor, Gleitman, & Gleitman, 2007, henceforth CFGG.

An obvious reason to believe that the pet fish problem is general is that we modify nouns when speaking just in those cases where we are talking about something other than the typical case. We say “green lime” and “orange carrot” very rarely just because these are the stereotypic colors of these edibles and therefore the modifier seems superfluous, hence odd to utter under Gricean conventions (Grice, 1975). It follows that the stereotypical properties and inferences assumed to be true of unmodified nouns are likely to apply less to modified nouns in general simply because the act of modification is ordinarily a mapping away from the typical case. This fact is problematic for using the prototypes of the simples to compose the complex concepts just because the prototype would seem less relevant within the context of combination. Nevertheless, there have been several attempts to reconcile prototypes with the requirements of compositionality.

One of the most well-specified and widely cited models of prototype-based composition is the **selective modification model** of Smith, Osherson, Rips, and Keane (1988, henceforth SORK). A diagram of how their model works to accomplish conceptual combination for Hume’s iconic APPLE instance is shown as Figure 1. Notice that the model avoids proliferating and uncontrolled context dependence by assuming that almost all features of the constituent elements retain their original (prototypical) values under combination, the only revision for RED APPLE being the feature dimension explicitly influenced by the combination; namely, its color. While the specification for APPLE gives some weight (‘*votes*’) to colors other than red, the rules of combination shift all the color weights to *red* and boost the valence on the color dimension. They leave the other feature specifications (*shape* and *texture*) unchanged. These simply retain the stereotypical weightings they had when unmodified: They default to their stereotype.

< FIGURE 1 ABOUT HERE >

In detail, the aspects of the model just stated are what allow it to work. A key aspect of selective modification, as SORK emphasized, is the selectivity itself. Thus ‘a purple apple’ is an atypical apple – in the combinatorics -- solely by virtue of its atypical color. But the compositionality constraint requires that the concept APPLE be a proper part of that combination; and further, under prototype theory, that it be represented as the APPLE-prototype. It’s the prototype, therefore, that gives the APPLE concept its identity, and thus it’s the prototype that confers APPLEness on the phrasal conjunction.[7] Preserving the structure of the prototype as much as one can is therefore necessary for this position to be internally consistent. Such preservation is possible if the prototype of the head noun of a NP is only minimally modified, where the dimension picked out by the modifier is altered selectively, preserving the values along other feature dimensions (i.e., *roundness, crunchiness*, and so forth). CFGG entitled this preservation of other dimensions, built in to the SORK apparatus, as *the **d**efault to the compositional **s**tereotype strategy* (henceforth DS).

DS says that barring information to the contrary, assume that the typical adjective-noun combination satisfies the noun stereotype. Thus, when pressed, one should judge a purple apple just as likely to be as crunchy or as sweet as any regular apple, ‘purple’ having selected only the color for modification. CFGG tested this prediction by having subjects judge sentences of four types, as exemplified below:

A. Ducks have webbed feet.
B. Quacking ducks have webbed feet.
C. Baby ducks have webbed feet.
D. Baby Peruvian ducks have webbed feet.

[7] Denying this claim is simply to deny that concepts should be equated with prototypes. Thus, for example, a hybrid theory wherein a concept consists of its prototype plus its denotation may be free to discount or disregard the prototype under combination while contributing its denotation to the complex. However, such a theory is not in our sights. We take it that the prototype bears the entire compositional burden according to the prototype purist (see also AGG, 1983, which discusses but discards such a “dual theory” for reasons related to the present ones).

The subjects' task was to judge how likely each sentence was to be true on a 10-point scale. The head noun ('ducks') of the subject noun phrase and the predicate ('have webbed feet') were held constant while the number and character of the modifiers were altered according to the four conditions (A – D). In the baseline condition A, an unmodified noun appears with a predicate that is true for typical instances of the noun. Condition B introduces a prototypical modifier – a modifier that is true of typical instances of the head noun. Condition C replaces the prototypical modifier with a non-prototypical (but not bizarre or contradictory) modifier, and condition D adds an additional modifier to the one in condition C. The predicates ('have webbed feet') and the prototypical modifiers of condition B ('quacking') were chosen because they appeared with high frequency on a list of feature norms for the associated head noun (Cree & McRae, 2003).

DS predicts that there should be no differences between these conditions in judged likelihood of the truth of these statements. That is to say, if we assume the inheritance of features from the head noun that are not directly implicated in the modification, 'baby Peruvian ducks' should be judged just as likely to have webbed feet as 'quacking ducks,' and so on. Listeners react with skepticism to this idea because it is implausible on the face of it, and indeed the results of this experiment show that experimental subjects did not react at all according to the prediction. Figure 1 shows the average subject ratings for the four conditions. While DS predicted that there would be no change from the baseline condition (A) across our experimental conditions, there was in fact a systematic deviation away from this baseline. Condition (B) produced judgments that were reliably lower than those of the baseline despite the fact that the modifiers in this condition belonged to the stereotypes of the head nouns. The introduction of one (C) and two (D) non-prototypical modifiers caused subjects to become progressively less certain as to the applicability of the predicates to the head nouns.

< INSERT FIGURE 2 ABOUT HERE >

These results (see Figure 2) show that our subjects did not use DS in judging the sentences, i.e., they do not conform to this crucial aspect of the model shown in Figure 1. Rather, they reflect the commonsense intuition that noun modification involves entertaining ideas other than what is typically assumed in the generic case. This is what one should expect because persistence in DS invites indefinitely many bad bets (for more argumentation along these lines see Fodor, 1998). There is no reason, in fact, to assume that a typical purple apple will be a typical apple that is typically purple. It could very well be (and so, apparently, our subjects reasoned) that purple apples are, in nature, some especially livid shade of purple and they might as a group be especially little and shriveled (or especially huge and bloated) apples. That is, the prototypical features, unless explicitly specified, do not carry over into the complex combination. This is exactly the pet fish problem, now examined quantitatively.

SORK in fact anticipated and discussed the failure of the model's DS predictions (and related issues) in their original paper and so cautioned that the selective modification model could not handle all types of conceptual combination. For example, the model deals only with simple feature dimensions such as *color* and *shape* that might plausibly be represented in the prototype, as opposed to dimensions picked out by more exotic possible modifiers as in 'Chinese apple,' or 'elephant apple.' Thus a large range of potential modifiers was left out of the equation, the major intent of SORK's demonstration being to expose some first principles of a successful model though temporarily sacrificing full coverage. It was also acknowledged at the time that emergent features resulting from some combinations, such as the largeness of wooden spoons, and the fatness of the tires on a beach bicycle, were problematic for the model, or indeed for any model that posits the inheritance of features of the head noun in forming composite prototypes (see Hampton, 1987). The principle of selectivity captures the primary assumption about the composition of prototypes that we have identified as DS. The contribution of the experimental findings (see again Figure 2) is to show that DS fails generally.

Recently Jönsson and Hampton (2008, henceforth JH) challenged the interpretations of these findings as presented in CFGG. They argue, on the one hand, that experimental subjects do in fact follow the DS strategy (most of the time) and cite their earlier findings that purport to support this claim (Jönsson & Hampton, 2006), and on the other hand, they argue that models of prototype composition, SORK included, need not entail DS anyhow. We reserve comment for now on their experimental findings, but we will address here their claim that DS is not assumed by models of prototype combination, including both SORK and Hampton (1987, 1988).

In the case of selective modification, JH's claim that DS is not assumed by selective modification is based more on a proviso provided by JH than by the original presentation of the model. Returning again to Figure 1, it depicts what happens to the APPLE prototype when it is modified by *red*. According to the model, only the color dimension is modified by switching all the weights (votes*)* to the *red* attribute (moving 5 from *green* to *red*) and boosting the diagnosticity of the color dimension from 1 to 2 (that is, the color dimension of the head noun concept becomes more prominent when it is modified with a color concept). The model embodies DS by keeping the weights on all unmentioned dimensions, e.g., *shape* and *texture*, unchanged. JH emphasize that relative to the color dimension, however, the other dimensions are in fact diminished.

Of course, *relative* to the color dimension, we agree that the value on the other dimensions is necessarily less. But this isn't the pertinent relation. Relative to the overall prototype representation, those values are in fact, unchanged. If the total *proportion* of weights was intended to matter, the sum of the weights should sum to a set total, but they don't: In the unmodified prototype the weights sum to 1.75 and in the modified representation, they sum to 2.75. Changing the model to reflect a set total in diagnosticity may be a reasonably minor change, but the only motivation for doing so appears to be to account for results similar to what we have reported in CFGG. For example, according to the revised model, a 'very red apple' will be expected to be less *crunchy* than a 'red apple,' and a 'very very red apple' will be less *crunchy* still. This is a counter-intuitive prediction, especially if one is ignorant of the outcome of CFGG. More importantly, adopting a reflexive demotion of corollary features as a modification to the SORK model amounts to a denial of DS, which the selective character of the model was supposed to preserve, thus accounting for the compositionality of prototypes. The new model would entail not only selective boosting of feature dimensions, but also non-selective squelching of corollary dimensions. The result would be a model that captures the intuition: *as more and more modifiers are heaped onto a complex phrase, the prototype of the head noun becomes less and less relevant to its meaning*. In our opinion, the intuition is correct, but it does not lend support for prototypes as the input to a compositional semantics. As we will next see, similar strategies to demote the properties built into prototype representations have been offered to accommodate to the fact that prototypes don't seem to compose.

In addition to the SORK model, JH point to another widely cited and empirically successful model of prototype composition, **the composite prototype model** (CPM**,** Hampton, 1987,1988) as another that does not entail DS. We disagree, and contend that it does in fact entail DS in its first step. The 6-step model (from JH, p. 917) is reproduced here:

(1) a composite prototype is formed by the union of the features of the conjuncts,

(2) all features with centrality so high that they are deemed necessary for either conjunct (e.g. fish have gills) will also be necessary for the conjunction,

(3) other features are assigned the average of their weights for each of the conjuncts, (a feature is given a weight of zero for a conjunct if it is not part of that prototype),

(4) features with low resulting weights are eliminated,

(5) a consistency checking procedure is run (informed by general knowledge), possibly resulting in the elimination and addition of further features in order to improve coherence,

(6) examples of the conjunction may also be retrieved from memory, and features of these may be added.

Thus, all features of the conjuncts are inherited by the complex concept as step 1. We take this to be an example of DS. In subsequent steps the model advocates pruning features, adding features, and adjusting weights, in effect undoing or fixing up what was done in step 1. This is analogous to the process advocated in JH's suggested revision of the SORK model; i.e., first inherit the features, then ratchet them down because they are likely to be misleading or irrelevant. Arguably, step (5) acknowledges that offending prototypical features have to be revised or eliminated under conceptual combination lest the output not be coherent. To return to the pet fish example, one wants to be sure to get rid of the *furriness* and *waggy-tailed* properties of prototypical pets in talking about fish, but preserving some other properties of convenient size and friendliness so as to exclude hammerhead sharks as good cases of pet fish. Just how to do this, however, remains obscure. While it is undeniable that models such as CPM can and do account for a good deal of data concerning subjects' intuitions about the features of combined concepts (e.g., Hampton, 1987, 1988), our concern is that this is not because it is a good model of lexical conceptual semantics, but, rather, because it is a good predictor of general pragmatic-inferential cognition. The model is under-constrained to the point that it blurs the line between conceptual and general knowledge. Indeed, it is also a familiar criticism of prototype theory in general that the criteria for what counts as a feature or property are similarly under-constrained.

To summarize, subjects do not appear to default to the stereotypes of the conjuncts of a combined concept when interpreting a novel combination. This is hardly surprising because *the more words/concepts combine, the less likely it becomes that they refer to things that satisfy their stereotypes.* We typically use adjectival modifiers in noun phrases when we are talking about something *other than* typical instances of the head noun. As this necessarily implies, any combinatorial scheme whose constituents are prototypes will therefore have to scramble to remove such typicality specifications as a condition for getting the interpretations of the complex concepts anywhere near the mark. Whether there is a general way of doing this is as much in doubt as it was when Hume (1739) and especially Locke (1690) discussed this very problem. In contrast, traditional theories of concept combination avoid this backtracking by not representing constituent concepts as stereotypes in the first place.

Summary and final thoughts

> …the necessity of communication by language brings men to an agreement in the signification of common words within some tolerable latitude that may serve for ordinary conversation and so a man cannot be supposed wholly ignorant of the ideas which are annexed to words by common use in a language familiar to him. But common use being but an uncertain rule, which reduces itself at last to the ideas of particular men, proves often but a very variable standard.
>
> John Locke, 1690, Book 3.XI.25

This discussion has focused on the question of how well the theory of prototypes can serve as the representational basis for human concept structure and understanding. The findings of AGG (1983), while never challenging the probabilistic feature-based views of concepts that have been ascendant in psychology and some schools of philosophy during the past 30 or so years, simply asked whether these representations were specific and nuanced enough to differentiate among central categories of human thought that are palpably different at their cores, say, between the concept SEVEN and such concepts as RHUBARB or SKATE-BOARD.

Experimental review of this question suggested that the experimental techniques widely taken to reveal prototypical concept structure failed even to render these fundamental distinctions among concept types. Such results cast doubt on the explanatory power of prototype and exemplar-based representations as those that feed conceptual combination, or at least on the empirical literature in psychology that purported to support this view. CFGG explored the same kind of question from the other way around, asking about the composition rather than the decomposition of concepts. Specifically, they asked whether the compositional rules operate over stereotypical representations of their constituents. The results suggest that prototypical properties associated with these constituents do not figure prominently under composition, but are systematically demoted. That is to say – and this is a tautology – compositionality must fail if there is context dependence: if the combinatorics alters the nature of the constituent elements.

The reasonable retort from prototype theory is that, after all, there *is* context dependence in the interpretation of complex concepts and it is manifest in our everyday understanding. Purple apples not only are purple (as the standard combinatorics tells us) but probably won't keep the doctor away, won't be appreciated by the teacher, and aren't good ingredients for American pies. Any theory of human conceptualization that does not answer to these facts is a failure on the face of it. On such grounds, it is a fair question whether the classical theory of compositionality avoids error only by abandoning hope of predicting almost anything at all about complex concepts. Our answer has been to the contrary. The classical combinatorics does a limited but absolutely required initial job in supporting concept combination and inference. It predicts what every English speaker knows and must know to understand words more than one at a time: *Purple apples are purple and they are apples*. It predicts as well that increasing the string of modifiers will have no effect on such inferences as *Large purple apples are purple, Large purple apples are apples*, and so forth, both of which are warranted by the compositional structure that a classical semantics assigns. But modification may well affect such prototypical inferences as *If it's an apple then it probably grew in the state of Washington and is sold in supermarkets*. These latter inferences derive not from the combinatorics but from our typical past experience with apples, Washington, and supermarkets.

References.

Armstrong, S. L., Gleitman, L. R. & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*, 263-308.

Aron, A. & Westbay, L. (1996) Dimensions of the prototype of love. *Journal of Personality and Social Psychology*, 70(3), 535-551.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.) *Concepts and Conceptual Development.* Cambridge University Press.

Barsalou, L. (1999) Perceptual Symbol Systems. *Behavioral and Brain Sciences*, 22, 577–660.

Connolly, A.C., Fodor, J. A., Gleitman, L. R. & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition, 103*(1), 1-22.

Cree, G. S., and McRae, K (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, 132, 163-201.

Fodor, J. A. (1981). The present status of the innateness controversy. In Representations: Philosophical essays on the foundations of cognitive science.

Fodor, J. A. (1998) 'There are no recognitional concepts — not even RED, Part 2: The plot thickens.' In In Critical Condition. Cambridge, MA: MIT Press, pp. 49–62.

Fodor, J. A., Garrett, M. F., Walker, E. C. T., and Parkes, C. H. (1980) Against definitions. *Cognition, 8*, 263-367.

Fodor, J. A., & Lepore, E. (1996). The red herring and the pet fish: why concepts still can't be prototypes *Cognition 58*(2), 253-270.

Fodor, J. A., & Lepore, E. (2002) The Compositionality Papers. Oxford: Oxford University Press.

Frege, G. (1892). Translations from the philosophical writings of Gottlob Frege. In P. Geach & M. Black (Eds.). Oxford: Oxford University Press.

.

Frege, G. (1914) ‘Logic in mathematics.’ In Hermes et al. eds., Posthumous Writings. Chicago: Chicago University Press, 1979: 201–250.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), Syntax and semantics, Vol. 3: Speech acts (pp. 41-58). New York: Academic Press.

Hampton, J.A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition, 15*, 55-71.

Hampton, J.A. (1988). Disjunction of natural concepts. Memory & Cognition, 16, 579-591.

Hampton, J. A. (1993). Prototype models of concept representation. In I. Van Mechelen, J. Hampton, R. Michalski, and P. Theuns (Eds.), *Categories and Concepts: Theoretical Views and Inductive Data Analysis* (pp. 67-95). New York: Academic Press.

Hess, T. M.; Pullen, S. M.; McGee, K. A. (1996). Acquisition of prototype-based information about social groups in adulthood. *Psychology and Aging 11*(1), 179-190.

Hume, D. (1739). A treatise of human nature: Being an attempt to introduce the experimental method of reasoning into moral subjects. In D. F. Norton & M. J. Norton (Eds.). Cambridge: Oxford Philosophical Texts.

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

Jönsson, M.C. & Hampton, J.A. (2008). On prototypes as defaults. (Comment on Connolly, Fodor, Gleitman and Gleitman, 2007). Cognition, 106, 913-923.

Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition, 57*(2), 129-191.

Katz, J. J. (1964). Semantic theory and the meaning of good. *Journal of Philosophy, 61*(23), 739-766.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University Of Chicago Press.

Locke, J. (1968). *An essay concerning human understanding.* Cleveland, Ohio: World Publishing Co. Original publication 1690.

Miller, G. & Johnson-Laird P. N. (1976). Language and perception. Cambridge, Mass: Belknap Press of Harvard University Press.

.
Murphy, G. L. (2002). The big book of concepts. Cambridge, MA: MIT Press.

Osherson, D., & Smith, E. (1981). On the adequacy of prototype theory as a theory of concepts. Cognition, 9, 35-58.

Pagin, P. (2003) ‘Communication and strong compositionality.’ Journal of Philosophical Logic 32: 287–322.

Patterson 2005 Learnability and CompositionalityMind & Language, Vol. 20 No. 3, pp. 326–352.

Prinz, J. (2002). Furnishing the mind: Concepts and their perceptual basis. Cambridge, MA: MIT Press.

Rey, G. (1983). Concepts and stereotypes. *Cognition, 15*, 237-262.

Rips, L. J. (1995). The Current Status of the Research on Concept Combination. *Mind and Language, 10,* 72-104.

Rips, L. J., Shoben, F. J., & Smith, E. F. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning & Verbal Behavior*, *12*, 1-20.

Robbins, P. (2005) The Myth of Reverse Compositionality Philosophical Studies 125: 251–275.

Rosch, E., (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.), *Cognitive Development and the Acquisition of Language*. New York, Academic Press.

Rosch, E. (1975) Cognitive representations of semantic categories. *J. exper. Psychol: Gen., 104,* 192-233.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 573--605.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B, B. Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ, Erlbaum.

Sinha, A. K. (2002) Prototype theory and the conceptualisation of culture. *Psychology and Developing Societies, 14*(1), 45-54.

Smith, E. E., and Medin, D. L. (1981) *Categories and concepts.* Cambridge, Harvard University Press.

Smith, E., Osherson, D., Rips, L., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, *12* (4), 485-527.

Spijkerman, R., & van den Eijnden, R. J. J. M. (2004). Addiction Research Institute, Rotterdam, Vitale, S., & Engels, R. C. M. E. Explaining adolescents' smoking and drinking behavior: The concept of smoker and drinker prototypes in relation to variables of the theory of planned behavior. *Addictive Behaviors*, *29*(8), 1615-1622.

Walsh, S. P., & White, K. M. (2007). Me, my mobile, and I: The role of self- and prototypical identity influences in the prediction of mobile phone behavior. *Journal of Applied Social Psychology*, *37*(10), 2405-2434.

Wanner, E. (1979). False identification of prime numbers. Paper presented at the 1979 meeting of *The Society for Philosophy and Psychology*, New York, NY.

Wittgenstein, L. (1953) *Philosophical investigations*. New York: MacMillan

Figure 1

APPLE		**RED APPLE**	
	RED 25		RED 30
1 *color*	GREEN 5	2 *color*	GREEN 0
	BROWN 0		BROWN 0
	…		…
	ROUND 15		ROUND 15
0.5 *shape*	SQUARE 0	0.5 *shape*	SQUARE 0
	CYLINDRICAL 5		CYLINDRICAL 5
	…		…
	SMOOTH 25		SMOOTH 25
0.25 *texture*	ROUGH 5	0.25 *texture*	ROUGH 5
	BUMPY 0		BUMPY 0
	…		…

Figure 2.

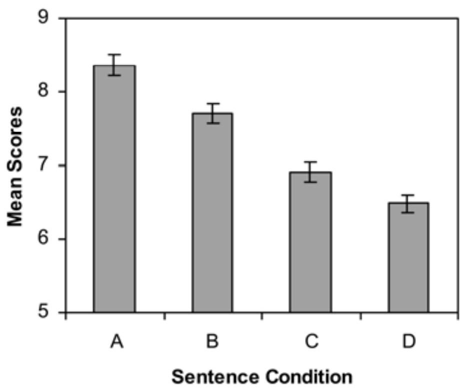

Table 1. **Exemplar ratings.** *Categories, category exemplars, and mean exemplariness ratings for prototype and well-defined categories for subjects who were asked only to give exemplariness ratings (Experiment I) compared to subjects who were asked first whether it made sense to rate items for degree of membership within the category and then to give exemplariness ratings (Experiment II). Only data for subjects who said NO to this question are included here.) Lower numbers correspond to ratings of comparative goodness of the exemplar, e.g., apples were judged as better fruits (mean rating 1.3) than olives (mean rating 6.4). (*adapted from Armstrong et al, 1983).

	Experiment I (all 31 subjects)		Experiment II (the subjects who said NO out of N = 21)	
	n	M	n	M
Prototype categories				
Fruit	31		9	
Apple		1.3		1.3
Strawberry		2.1		1.7
Plum		2.5		1.9
Pineapple		2.7		1.3
Fig		5.2		3.3
Olive		6.4		4.2
Vehicle	31		5	
Car		1.0		1.0
Boat		3.3		1.6
Scooter		4.5		3.8
Tricycle		4.7		2.6
Horse		5.2		2.8

	Skiis	5.6		5.2
Well-defined categories				
Odd number	31		21	
	3	1.6		1.0
	7	1.9		1.0
	23	2.4		1.3
	57	2.6		1.5
	501	3.5		1.8
	447	3.7		1.9
Female	31		18	
	Mother	1.7		1.1
	Housewife	2.4		1.8
	Princess	3.0		2.1
	Waitress	3.2		2.4
	Policewoman	3.9		2.9
	Comedienne	4.5		3.1

FIGURE CAPTIONS

Fig. 1. Illustration of Smith et al. (1988) selective modification model for deriving a prototype for the combined concept *purple apple* by modifying the color dimension of the *apple* prototype. Crucially, dimensions not directly affected by the modification process are inherited as defaults.

Fig. 2. Grand means for subjects' plausibility judgments on a scale from 1 (highly unlikely) to 10 (most likely), for four sentence types. (A) e.g., Ducks have webbed feet. (B) Quacking ducks have webbed feet.(C) Baby ducks have webbed feet. (D) Baby Peruvian ducks have webbed feet. Error bars show the standard
error for the subject means