Available online at www.sciencedirect.com
ScienceDirect
Cognition 103 (2007) 1–22
COGNITION
www.elsevier.com/locate/COGNIT

Why stereotypes don't even make good defaults ☆

Andrew C. Connolly [a,*,1], Jerry A. Fodor [b], Lila R. Gleitman [a,b], Henry Gleitman [b]

[a] *Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA*
[b] *Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA*

Received 30 September 2005; revised 1 February 2006; accepted 15 February 2006

Abstract

Many concepts have stereotypes. This leaves open the question of whether concepts *are* stereotypes. It has been argued elsewhere that theories that identify concepts with their stereotypes or with stereotypical properties of their instances (e.g., Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Ed.), *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates; Smith, E. E., Medin, D. L. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.) fail to provide an adequate account of the compositionality of concepts (Fodor, J., Lepore, E. (1996). The red herring and the pet fish: Why concepts still cannot be prototypes. *Cognition*, 58, 253–270.; Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. New York, NY: Oxford University Press.). This paper extends this argument and reports an experiment suggesting that participants do not assume, even as a default strategy, that complex concepts inherit the stereotypes of their constituents. Thus propositions such as "Baby ducks have webbed feet" were judged to be less likely to be true than propositions like "Ducks have webbed feet." Moreover, manipulation of the type and number of noun phrase modifiers revealed a systematic departure from the unmodified noun's stereotype both with the addition of stereotypical modifiers ("Quacking ducks have webbed feet" versus "Ducks have webbed feet") and with the addition of a second modifier ("Baby Peruvian ducks have webbed feet" versus "Baby ducks have webbed feet"). Thus, in the general case the stereotypical properties of a head noun are systematically

☆ This manuscript was accepted under the editorship of Jacques Mehler.
* Corresponding author. Tel.: +1 215 898 1712.
E-mail address: andrew73@psych.upenn.edu (A.C. Connolly).
[1] Authors are listed in reverse order of superannuation.

doi:10.1016/j.cognition.2006.02.005

discounted when that head noun combines with modifiers. This effect represents a general principle of conceptual combination that argues against the inheritance of stereotypical features of concepts as a default strategy. Instead, we advocate a model of conceptual combination where concepts remain inert under combination, supported by a separate machinery that introduces pragmatic and knowledge-dependent inferences.

Keywords: Concepts; Conceptual combination; Stereotypes; Prototypes; Compositionality

1. Introduction

Because concepts matter so much to theories of cognition, discussions about what they are have been recurrent in the literature in cognitive psychology and in the philosophy of mind. People used to think that a concept is an icon (something like a mental image, e.g., Hume, 1739). A more recent view was 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). Though these traditional views are still maintained by some commentators, it is fair to say that in psychology today the dominant view is that a concept is a stereotype (something like a set of properties which things that fall under the concept typically have or are believed to have; Prinz, 2002; Rosch, 1978; Rosch & Mervis, 1975; Smith & Medin, 1981).

We do not think that concepts are images or definitions but we do not think they could be stereotypes either. Elsewhere, we have presented some experimental evidence suggesting that concepts might not *be* stereotypes even though many concepts indubitably *have* stereotypes (Armstrong, Gleitman, & Gleitman, 1983; see also Rey, 1983; Fodor, 1998).[2] The present paper continues to insist upon this *have/be* distinction concerning the relationship between stereotypes and concepts. Here we present further results that bear on this claim, this time centering on the issue of compositionality. Our argument can be made briefly as follows: (1) concepts must be compositional, (2) stereotypes are not compositional, (3) therefore, concepts are not stereotypes.

2. Compositionality and context dependence

We will assume that the class of concepts belonging to the human system of mental representations is productive and systematic.[3] Simply put, humans are in

[2] In particular, we showed there that such formal concepts as "odd number" have stereotypes (e.g., *three* and *thirty-one* are asserted by experimental participants to be better instances of odd numbers than are *eleven* or *twenty-one)*. Yet the same participants denied that some odd numbers are odder than other odd numbers. This response pattern exemplifies the difference between having a stereotype and being a stereotype.

[3] Since there is a substantial cognitive science literature on productivity and systematicity, and their relation to compositionality, the present treatment is cursory. For a recent book-length survey of the issues, see Aizawa, 2003.

possession of some finite number of primitive concepts, and some (recursive) principles for their combination. This view is shared by nearly everyone who has thought about these problems, whatever their theoretical disagreements in other regards (e.g., Chomsky, 1959; Hume, 1739). This is because the only way to understand infinite (unlimited) thoughts in a resource-limited brain is by acknowledging that the big thoughts are compositional products of the small ones:

(i) *the productivity and systematicity of human concepts is explained by their compositionality.*

There are several more or less equivalent ways to describe what compositionality amounts to. Very roughly, a system of representations is compositional iff: (a) it contains both syntactically primitive and syntactically complex symbols; and (b) the syntax and content of the complex symbols is exhaustively determined by the syntax and content of their primitive constituents. So, for example, the English expression *red hair* is compositional because: (a′) 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; and (b′) 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 (c) that AN representations denote the intersection of the As and the Ns. Thus *red hair* means 'HAIR that is RED.'[4] The crucial assumption is that the syntactic and semantic properties of the constituents of complex compositional concepts are context independent, viz., independent of the syntactic and semantic properties of their linguistic environment. RED does not change in its extension as a consequence of being combined with some other (primitive or complex) element.

A moment's reflection reveals, however, that there is more than this to the story of understanding complex expressions. For one thing, it is well known that the intersective combinatorics of AN expressions is not true over the whole range of adjective/noun pairs. *Decoy ducks* are not ducks, *putative examples* often are not examples, *former senators* are not senators, and so forth (see Partee, 1995; for a good discussion). And not all phrases that have intersective interpretations are AN: many, like *pet fish, gentleman farmer*, and *boy genius* are NP's with the constituent structure NN (thus you can be, grammatically speaking, a very brown cow but not a very pet fish). But the problems for compositionality under the classical approaches are much deeper, and are not limited to such exotica. Rather, context independence appears to fail very often when concepts combine. Compare, for example, such everyday phrases as *red hair*, *red cheeks*, and *red apples*. The particular hue of red first appearing in the mind's eye when hearing the phrase *red hair* is a different hue than that evoked by the phrase *red cheeks*. That is, the plausible or immediate interpretation of *red* is not consistent over the entire range of nouns that it can modify.

[4] Notationally, we use *italics* for the mention of a word or phrase, "quotes" for its utterance, and ALL CAPS for concepts.

The classical theory of conceptual combination relegates these plausibility effects to a separate inferential stage of language understanding. According to this position, all you get from your concepts and combinatorics is output denoting relations among sets, properties, or individuals (depending on the ontology assumed). Thus*red hair* designates the set of instances of hair whose colors are instances of red. A virtue of this kind of theory is that it affords the flexibility needed to account for all interpretive possibilities, regardless of considerations of plausibility. For example, you would not be likely to guess, upon learning that someone you have not 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 for example, that *red cheeks* not be interpretable as CHEEKS *that are* GREEN or *ELBOWS* that are RED.

Summarizing, a virtue of the classical theory is that it seems to be true and, as far as it goes, psychologically valid. The phrase *red hair* covers all and only the cases, plausible and implausible, that would be correctly considered as actual or possible members of the extension of the phrase. A potential problem of the theory is that precisely because it has this virtue it fails to account for the fact that some instances seem more plausible than others, as has been shown again and again in the conceptual combination literature (e.g., Kamp & Partee, 1995; Osherson & Smith, 1981; Rips, 1995). However, this drawback is easily fixed (so say classical theorists) by assuming that the recovery of meaning from the 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.[5]

The prototype approaches to conceptual combination are quite different. One of their aims is to reduce the machinery of understanding to a one-step process. This is accomplished in part by building some of the facts about plausibility and informal inference into the semantics itself rather than relegating them to a separate inferential process. Thus while the classical analysis of a concept like, say, 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 of hair could be specified along a color dimension within the representation for HAIR, the redness range for cheeks within the representation of CHEEK, and so forth (Kamp & Partee, 1995; Katz, 1964; Osherson & Smith, 1981).

There are known limitations as to how much work can be accomplished by a one-step theory, however. This is demonstrated by the phenomenon of unexpected

[5] When we say "easily fixed," we do not mean to imply – contrary to fact and reason – that the pragmatic inference procedures and properties will be easy to discover or describe. Only that their solution (better, lack of solution) is a feature common to all known semantic theories, both classical and prototypical.

probabilistic features arising after combination which cannot plausibly be attributed to the constituent concepts, even when these are prototypically represented. An example is an emergent feature of the combined concept BEACH BICYCLE; namely that such vehicles have extra fat or squishy tires so as to facilitate traveling over sand (Hampton, 1987). Presumably *having fat tires* can be inferred only after combination and is not entailed under the prototype for bicycle. As a consequence, the classical theory and prototype theory agree that there must be both a combination process and an interpretation process in order to account for the full range of possibilities.

The difference between the two theories lies in the commitments about the primitive representations themselves, and the consequences for what happens during the combination process. Under prototype theory, a primitive consists of a set of stereotypical inferences that can be brought to bear on the ultimate interpretation of the combined concept of which it will be a part. Thus at least some of the work of building inferences is accomplished during the combination stage as prototype structures interact with each other to form a composite prototype structure that can then be subjected to further inferential processing and interpretation. In contrast, all such inferences lie outside the combinatorics of the classical theory.

3. Default to the stereotype

The question that next arises, and that is the subject of the present paper, is whether the prototype structure account of the concepts can support an account of conceptual combination. As several authors have recognized, the classical supposition that concepts are inert under combination (‘‘context independent’’) fails if the representation of the primitives is taken to be prototypical. The iconic counter example to the claim that prototype concepts can account for the compositionality of concepts is the complex concept PET FISH (Fodor & Lepore, 1996). By the assumption that concepts are prototypes, this concept is a specification of the properties that are typical of pet fish as such; i.e., the properties the possession of which would make a pet fish a good or typical instance of a pet fish. Compositionality requires, in turn, that this specification should be constructed from the stereotype for PET (*aka*, the concept PET) and the stereotype for FISH (*aka*, the concept FISH).

The following line of thought might now suggest itself: something is a pet fish iff it’s a pet and a fish. So by analogy something is a stereotypic pet fish iff it’s a stereotypic pet and a stereotypic fish. But on second thought this proposal does not work. A good example of a pet fish (a guppy, as it might be) is neither a good example of a pet (compare poodles) nor a good example of a fish (compare trout). So there is a *prima facie* problem about how to reconcile the stereotype theory of concepts with the compositionality constraint.

Various solutions to this crux have been offered in the recent cognitive science literature. Their underlying idea is to avoid the inheritance of each and every prototypical feature of the modifier (in our example, *pet*) and the head (*fish*) by the prototype combination (*pet fish*). Rather, certain features will be modified under combination, *but the* rest *will not*. That is, the rest retain (default to)*their stereotypical values*.

For the pet fish example, presumably we would want the cuteness and relative smallness features of the prototypical pet to remain as is under combination with *fish* (yielding a guppy rather than a hammerhead shark) but not the stereotypical coats of common pets (lest we derive a furry guppy). If this selection process among prototypical values of inherited features can be made good in general, using prototypical representations, then the prototype theory will have accomplished just what is wanted: complex prototype representations that satisfy the compositionality constraint (i).

Summarizing: any successful prototype model that explains conceptual combination must assure that component properties not specified by or semantically incompatible with the particular combination are inherited *as is*. They *default to the stereotype* (henceforth, DS) of the component elements. That way *pet fish* can be as cute and compact as all stereotypical pets should be, but at the same time not inherit the mammalian features of stereotypical pets. The crucial intuition here (just a restatement of DS) is that *the less familiar a complex concept, the more one would assume the prototypical values of its constituents.*

The present paper aims to test whether the crucial DS prediction of prototype models is empirically satisfied, that is, accords with listener judgments about newly heard phrases. But before turning to the experiment, we want to give a bit more detail about how such models have been said to work. Several versions of DS have appeared in the psychological literature in various guises including *selective modification* (Smith, Osherson, Rips, & Keane, 1988), *feature mapping* (Wisniewski, 1997), *feature pooling* (Hampton, 1991), among others (e.g., Cohen & Murphy, 1984; Kamp & Partee, 1995; Murphy, 1988, 1990). The commonality among these theories of prototype combination and the reason why they represent versions of DS is that they all postulate the inheritance of (at least parts of) the feature matrices of the constituents of a combined concept. The *selective modification* model (henceforth SM) of Smith et al. (1988) will serve here as an example to illustrate the predictions of this class of models.

According to SM, prototype representations are explicitly construed as structured matrices of variously weighted features. In order to derive the prototype structure of the combined concept PURPLE APPLE on this account (scanting some of the details), you start with the prototype structure of APPLE (see the left hand side of Fig. 1), which contains several feature dimensions including *color*, *shape*, *texture*, etc. Each dimension contains a range of possible values that are associated with a number of ‘votes’ that reflect the default values for the stereotype of APPLE. So, before modification, the *color* dimension for APPLE reflects the stereotype for apple (i.e., apples are typically red). After modification by PURPLE, the APPLE matrix is altered by shifting the votes from the other color values to PURPLE resulting in a new prototype matrix reflecting the features of a typical purple apple.

As we noted, a critical feature of this and related prototype models is that, except for features that are explicitly affected by the modification, the default values of the prototype structures are simply inherited; in the example, the shape and texture dimensions of APPLE remain unchanged after modification of the color dimension by PURPLE. Generalizing, a central prediction is that for any dimension x of prototype structure N that is not explicitly affected or contradicted by the requirements

	APPLE		**PURPLE APPLE**
	RED 25		RED 0
1 *color*	GREEN 5	2 *color*	GREEN 0
	BROWN 0		BROWN 0
	PURPLE 0		PURPLE 30
	…		…
	ROUND 15		ROUND 15
0.5 *shape*	SQUARE 0	0.5 *shape*	SQUARE 0
	CYLINDRICAL 0		CYLINDRICAL 0
	…		…
	SMOOTH 20		SMOOTH 20
0.5 *texture*	BUMPY 0	0.5 *texture*	BUMPY 0
	…		…

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.

of combinations that N enters into, psychological measures that reflect the value of x should register no reliable differences across different combinatorial conditions. In this way, prototypes resemble the stereotypical members of their reference sets in the 'unmarked' case. As a result, the concept HAIR by this account contains a color dimension which when modified by RED produces the ruddy brownish red hue typically associated with hair. So far, so good. Now what happens to the uninfluenced properties? Suppose that hair has a range of textures as well as colors (e.g., fine, medium, coarse). In constructing the complex category RED HAIR, the weights for texture will not change, i.e., the model does not expect red hair to deviate from stereotypical texture, and *mutatis mutandis* coarse hair is not expected to be more or less red than is the stereotypical case for hair. As already implied, certain conceptual combinations should not perturb the stereotypical representations of their components at all. Consider LITHUANIAN HAIR. Assuming that the Lithuanians can have hair but do not have any particular non-stereotypical hair types (or even assuming that they do, but that the speaker-listener does not know this), all properties of the component representations will be inherited.

Before ending this section, we should note that the instance of *pet fish* need not be an embarrassment for prototype theories. Rather, it can be argued that the stereotype in this particular case is derived from experience and not by composition, so that *pet fish* is something like an idiom that is interpreted more or less holistically. That is not (a prototype theorist might reasonably argue) because stereotypes do not compose; rather, it's because the assignment of a stereotype to *pet fish* is unrepresentative of the standard case. In the standard AN (or NN) case, one does not have proprietary extralinguistic knowledge about which instances are stereotypic; that being so, one resorts to the (compositional) assumption that the stereotypic ANs are the intersection of the stereotypic As with the stereotypic Ns. So, *except that* one has indeed hung around fish and/or pet stores, one *would* assume (in effect, by default), that the stereotypic pet fish is both a stereotypic pet and a stereotypic fish.[6] That is, the *default to the* (compositional) *stereotype* strategy (=DS) applies only where composition is not superceded by prior experience and knowledge.

3.1. Testing the DS strategy

We claim that DS does not reconcile the stereotype theory with the compositionality constraint. Our argument has two prongs: first (Section 3.1.1) we will argue that *a priori* DS is a bad strategy. Thereafter (Section 3.1.2) we introduce an experimental approach designed to discover whether DS is the strategy that experimental participants use to interpret complex concepts.

3.1.1. Defaulting to the stereotype is a bad strategy

DS says:

(ii) *Barring information to the contrary, assume that the typical AN satisfies (inter alia) the N stereotype.*

But this is likely to be a losing strategy because, ceteris paribus, the probability that an arbitrary AN is a good instance of N is smaller than the probability that an arbitrary N is a good instance of N. Likewise, the probability that an arbitrary AAN is a good AN (or a good A or a good N) is smaller than either the probability that an arbitrary AN is a good AN or the probability that an arbitrary AN is a good N or a good A. And so forth.[7] Persistence in DS thus invites indefinitely many bad bets.

This is because there is, in general, no reason why the modal members of overlapping distributions should themselves be in the overlap. Fig. 2 shows a situation where they're not and there are, of course, indefinitely many other such distributions.

3.1.2. Human language users might not follow DS

It might be argued that, although the probability that an AAAN is a stereotypic N is likely to be pretty low, still that is one's best bet if one has literally *no* real-world information about AAANs, i.e., if one is going *solely* on one's conceptual competence, in this case by hypothesis a competence that is built out of stereotypic concept representations and their combinatorics.[8] But notice that, if that's the strategy, then

[6] On the other hand, it *is* something of an embarrassment for this view that the intersection of the stereotypic As with the stereotypic Ns is often empty. In fact, it is so in the case of pet fish since *nothing* is both a stereotypic pet and a stereotypic fish. Likewise for miniature collies; as a matter of fact, there are no miniature collies that are collies at all, stereotypic or otherwise; there are only regular collies and Shelties.

[7] If p(AN) is estimated to be higher than p(N), we have what Kahneman and Tversky (1972) call a 'conjunction error'. They suggest, plausibly in our view, that conjunction errors occur when an AN has a proprietary (i.e., non-compositional) stereotype. (Thus, for example, p(pet fish live in bowls) > p(fish live in bowls)). Such inferences are, of course, perfectly reasonable if their conclusions are understood to mean that the probability that pet fish typically live in bowls is greater than the probability that fish in general typically do.

[8] There is, however, an alternative (and more reasonable) course that one might pursue: Opt out. If you do not know the odds, do not bet. It strikes us as peculiar to take the forced choice case as the paradigm for evaluating a strategy for making decisions; one is, after all, not usually presented with forced choices.

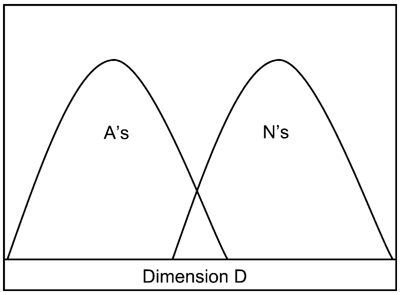

Fig. 2. A hypothetical distribution of the populations of As and the population of Ns with respect to values of the dimension D. Note that the intersection of the distributions (i.e., the A&Ns) includes neither the modal As nor the modal Ns. When a third population is considered (e.g., the AANs), the probability that the modal members of all three populations are in their intersection (viz. that they are in the set {AAN & AN & A & N}) would then be smaller still.

one's confidence that an arbitrary AN is a stereotypic N should be no less than one's confidence that an arbitrary N is a stereotypic N. Likewise, one's confidence that an arbitrary AAN is a stereotypic AN should be no less than one's confidence that an arbitrary AN is a stereotypic AN. And so forth. Piling up more As should not affect one's willingness to bet on the inference from a premise of the form A^nN to the conclusion that N is stereotypic. If anything, the less one knows the more one should fall back on the stereotype. That this is not the case is demonstrated by the data we're about to present. Contrarily, from the stance that DS is generally a bad strategy, subjects should be less certain about the applicability of the typical features since, as Fig. 2 illustrates, the modal members (and therefore also the features associated with those members) of overlapping distributions are less likely to be in the overlap.

4. Design of the experiment

4.1. Participants

The 40 participants were undergraduates (20 female, 20 male) at the University of Pennsylvania who received in return class credit in their introductory psychology course. All were native speakers of English.

4.2. Procedure

These subjects were randomly assigned to one of four groups. Each received a nine-page packet. The first page of each packet contained instructions. The remaining eight pages contained 175 sentences. The numbers 1 through 10 appeared to the right of each sentence. Subjects were instructed to circle a number for each sentence indicating the likelihood, in their judgment, that the sentence was true (1 = very unlikely; 10 = very likely). Subjects were allotted one hour to complete the packet, but most finished in between twenty and forty minutes.

4.3. Materials

The target stimuli consisted of simple declarative sentences of four types exemplified by the following:

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

The head noun ('ducks') of the subject noun phrase and the predicate ('have webbed feet') of all four sentence types were held constant while the number and character of the modifiers were altered according to four conditions (A–D). Each head noun was taken from a database of feature norms compiled by Cree and McRae (2003) who asked subjects to produce feature lists for a variety of common nouns. The predicates ('have webbed feet') and the 'prototypical' modifiers of condition B ('quacking') were chosen because they appeared with high frequency on this list of feature norms for the associated head noun. 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 for typical instances of the head noun. This condition was included to control for the effects of modification on its own. While 'prototypical' modifiers might be assumed to implicate the stereotype of the head noun, there is something pragmatically deviant about saying 'quacking ducks': it seems to indicate either ducks that are actively quacking at this moment, or perhaps ducks that quack more often than usual, etc. Such locutions might implicate something other than the stereotype of the head noun, and therefore, can produce judgments that are somewhat lower than the baseline condition (A). We call this condition the 'Gricean' condition because the very act of modification implicates an interpretation of the head noun that deviates from the stereotype of the head noun, which is reminiscent of Grice's cooperative principle (Grice, 1975) that a speaker cooperates with his interlocutor by providing just enough information (neither too much nor too little) to be maximally informative.

Condition C replaces the prototypical modifier with a non-prototypical modifier, and condition D adds an additional modifier to the one in condition C. The non-prototypical modifiers in conditions C and D were chosen using the strict criterion that they did not appear on the list of feature norms and on the more subjective basis that they should increase uncertainty as to the derivation of the noun phrase. That is, our aim was that the interpretation of the non-prototypically modified noun phrase not be something that has a prototype independent of the effects of modification. As is the case with 'pet fish', some combinations have prototypes because people have experience with instances of the combination's extension (i.e., hanging around pet stores). We wanted to avoid such cases so that the interpretations of our combinations could *only* be viewed as arising from compositional processes. (For we are testing compositional processes, asking whether these operate across stereotypic representations of their conceptual conjuncts.) In addition, the modifiers were chosen to be commensurable with the predicates, i.e., the modifier and the predicate should be orthogonal and independent, so that the introduction of the modifier does not necessitate a change in the applicability of the predicate. For example, "Legless ducks have webbed feet" would not have been a candidate because (except under very odd circumstances) being legless implies also being footless. According to DS, conditions C and D are precisely the kind where subjects should default to the stereotype of the head noun because (1) presumably subjects should not have a pre-stored stereotype that might preempt a compositional stereotype (re PET FISH), and (2) there should be nothing about the combination that prevents the stereotypic features from being inherited. Accordingly, DS predicts probability judgments for these conditions that do not differ from the control conditions, or from each other. The inclusion of both conditions C and D were necessary here to test the predictions of DS in light of the counter argument (offered in Section 3.1.1) that persistence in defaulting to the stereotype is an illogical strategy; in particular, that the probability that an instance of AAN is a typical N is lower than the probability that an AN is a typical N, and so forth. According to this logic, one would expect subjects' confidence levels in D to be lower than those in C.

Forty quadruples as in (1) were constructed based on forty different head nouns. Four lists were compiled such that forty target sentences (one from each head-noun quadruple and ten of each sentence type, A-D) appeared on each list. The target sentences were embedded in 120 filler sentences and were randomly ordered except that two target sentences could not appear adjacent in serial order. An additional 15 filler sentences appeared at the front of each list to prevent warm-up effects. The filler sentences were constructed to be similar but not identical to the target sentences in structure. In addition, the filler sentences were constructed so that there were equal numbers of sentences with three different levels of plausibility (45 of each): 1 – not plausible (definitely false; e.g., "The Roman coliseum is in Moscow."); 5 – intermediate plausibility (possibly true; e.g., "Philadelphia is larger than Atlanta."); and 10 – highly plausible (definitely true; "Bicycles have two wheels."). The serial position of all filler sentences was the same for all four lists, and each target position was always filled by a target of the same head noun quadruple. So, 'ducks' always appeared in the same serial position across lists, but with a different modification. Each subject filled out only one list, and therefore, saw one sentence type for each head noun.

4.4. Results

Subjects' judgments for all target sentences were entered into an ANOVA with modifier condition as a four-level within-subject variable and list condition as a four-level between subject variable and subjects as the random variable. There was a main effect of modifier condition in the expected direction (see Fig. 3): unmodified condition mean (A) = 8.36; Gricean condition mean (B) = 7.71; single modifier condition mean (C) = 6.91; and the double modifier condition mean (D) = 6.48 ($F(3,108) = 67.69$; $p < .001$). There was no effect of list and no interactions. Pairwise t-tests were significant ($p < .05$) between all levels of the modifier condition. An item

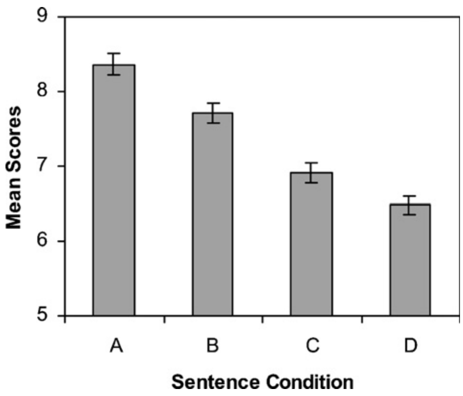

Fig. 3. Grand means for subjects' plausibility judgments on a scale from 1 (highly unlikely) to 10 (most likely), for four sentence types. (A) e.g., *Ducks have webbed feet.* (B) *Quacking ducks have webbed feet.* (C) *Baby ducks have webbed feet.* (D) *Baby Peruvian ducks have webbed feet.* Error bars show the standard error for the subject means.

analysis was also performed on the same data with one four-level variable (modifier condition) and with items as the random variable. Here, an item was taken to be a head noun with each item appearing in four different modifier conditions. Again, there was a main effect of modifier condition ($F(3, 117) = 43.55$; $p < .001$). Pairwise t-tests were significant ($p < 0.05$) for comparisons between each level.

5. Discussion

5.1. Summary and conclusions

Our ambition in this study was to evaluate the position that human concepts are represented as stereotypes; in particular, as matrices of dimensionalized weighted features. Common words, according to such a theory, express such feature matrices (i.e., the word "dog" is a linguistic label for a concept DOG, that concept being a congeries of sublexical features that vary in their centrality to the concept). Because we can speak and productively understand more than one concept/word at a time, these feature matrices must support compositionality. The mental representations of the concepts PET and FISH must yield the complex concept PET FISH. Complaints about prototype representations as sufficient to this task concern the absence of a compositional model that could operate on stereotypes systematically, as required (e.g., Armstrong et al., 1983; Fodor & Lepore, 2002; Osherson & Smith, 1981; inter alia). After all, how would one know which of the features with which of its range of weights is intended in some new modified phrase, e.g., if someone said, and someone else heard "I have a pet fish" or "I have a purple apple."

In response, interesting prototype models have been developed that assume that all dimensions and weights of the simple conceptual representations are held unchanged under modification unless they figure in it specifically (i.e., the feature matrix that we label with the word *apple* in English is unchanged except for its color specification when combined with a color term such as *red* or *purple*). This is the Default to Stereotype (DS) theory of (prototypical) conceptual combination that our experiment has tried to assess.

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. The "Gricean" 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. Furthermore, the introduction of one (C) and two (D) non-prototypical modifiers caused subjects to be progressively less certain as to the applicability of the predicates to the head nouns. The findings that the subjects' judgments in C and D were reliably lower than B show that there are effects of combination over and above those that can be attributed to the "Gricean" variables tested in condition B. This set of results is inconsistent with the claim of DS that conceptual combination entails the inheritance of stereotypical default values for features that do not figure explicitly in the combination.

Our belief is that our paper should end right here. However, we want to confront a frequent reaction to the component experiment and its interpretation. This reaction consists of three linked assertions (1) your (that is, the present authors') result of weakened confidence under conceptual combination is expectable and banal (we agree with this one); (2) the whole explanatory problem arises because of the assumed strategy of default inheritance of prototypical properties (we agree with that too); (3) so why not just keep the stereotypical representations but add on the increasing uncertainty (this requires a bit further discussion, we trust not too much).

5.2. Prototypes without tears: Can we keep the theory and scuttle its entailments?

It has been suggested to us by some commentators that it may be possible to account for our results by assuming the following general principle:

(iii) *Uncertainty under composition* (*UC*): *The more concepts enter into novel combinations, the less certain one is about the stereotypic properties of the instances of the combined concept.*

It has further been suggested that a prudent form of DS could incorporate this principle by automatically discounting collateral features by some fixed proportion, say 5%. This would allow for the inheritance of features as required by DS, while at the same time decreasing their weights, i.e., decreasing the likelihood that they are relevant to the combination. Furthermore, since this mechanism can be applied recursively, adding modifiers to an existing combination will have the result of further discounting inherited features, in line with the results from our condition D. Thus adding such a mechanism to DS could neatly predict our results.

Notice that this is simply a denial of the DS assumption. That is, a mechanism for systematically discounting inherited features would function solely to reduce the confidence with which S is prepared to assume that the new feature matrix inherits values from the old ones. By contrast, the Classical theory is not required to take AN constructions as an exception to the account of compositionality it proposes. The data are just as they should be if the Classical theory is assumed without emendation.[9]

To summarize, the new intuition represented in UC is counter to the idea that unmentioned values of the prototype representation are inert under combination. In contrast, UC says that one becomes less certain about these values with each modification – that is, the listener who hears "purple apple" is less sure than when hearing "apple" that this fruit is likely to keep the doctor away. And when one hears that the apple is not only purple but Lithuanian *and* purple, the thought arises that this piece of fruit, unlike a prototypical apple, might be detrimental to one's health. We agree (and have throughout) with something like this second intuition. In fact, we graphed it in Fig. 2 of this paper. Prototypical properties, if these were built into one's representation of unmodified concepts, would have to be discounted under conceptual combination. But this just means that prototypes do not compose.

It is commonly assumed that whatever combinatorial processes are proposed to handle conceptual combinations, the output of such processes becomes the input to a secondary knowledge-dependent stage in order to produce an ultimate interpretation (e.g., Murphy, 1990, 2002; Prinz, 2002; Smith et al., 1988; inter alia). This secondary stage has been proposed in the prototype literature as a way to deal with some of the problems associated with the inheritance of features, such as reconciling incompatible features (Hampton, 1991) and accounting for so-called emergent features (Hampton, 1987) consonant with the assumption that composition is defined over prototypes. But a more parsimonious solution is to avoid the problems associated with the inheritance of stereotypical features by abandoning that assumption. That would seem to be the tactic that the present results recommend.

The classical view of conceptual combination makes minimal assumptions about the knowledge deployed during conceptual combination, e.g., for the conjunctive concept XY, the interpretation is X AND Y. Accordingly, concepts remain inert under combination and *all* knowledge-based inferential processing, including which features get '*inherited*', occurs during a secondary stage.

The choice, then, is between two two-stage models. The classical view proposes minimal concepts and combinatorics followed by a knowledge-driven inferential stage. The prototype view proposes that knowledge relevant to deriving interpretations of combinations is built into the concepts themselves, providing an inferential

[9] We should point out as well that adopting any constant uncertainty proportion (e.g., 5%) per modifying element cannot really work as, e.g., adding *purple* to *apple* is sure to diminish one's confidence about its edibility more than adding *ripe* and less than adding *Martian.* But the idea of a variable deweighting procedure is a computational nightmare; namely, it is the ineffable all-purpose all-knowledge comprehension schema that we call "pragmatic inference." Another kind of disclaimer to our conclusions mentioned by an anonymous reviewer is to the effect that increasing "information load" is accounting for the growing uncertainty as to whether ducks remain ducky though Peruvian and immature. Because our subjects were under no time pressures in providing their confidence ratings, and since all of our stimuli were conceptually quite trivial, even laughably so, we do not think it likely that an account in terms of our task or item difficulty is *prima facie* plausible.

mechanism through the automatic inheritance of features. The combinatorics produce output that is systematically misleading, as long as it is uncorrected by the hypothetical UC machinery whose only purpose is to squelch the offending stereotypical values inherent in the prototype representation of unmodified concepts. Thus parsimony favors the classical approach given these considerations.

We might add that these results are not the only laboratory effects suggesting that the prototype view of conceptual combination is not the correct model. Another striking finding comes from Springer and Murphy (1992). As they note, the two-stage prototype model predicts that at some stage during the interpretation of a combined concept inherited default features should be activated before features that depend on knowledge-based inferences for the combined concept. Yet they found that emergent phrasal features are available immediately and before inherited features of phrase constituents. Subjects were faster to say yes in a truth verification task to sentences like "Boiled celery is soft" (expressing a true phrasal property, for unboiled celery is not soft at all) than that "Boiled celery is green" (expressing a true noun property for celery is green, boiled or not). Subjects were also not significantly slower to answer no to "Boiled celery is crunchy" (a false phrasal property), which might be expected if *crunchy*, a true noun property for celery singularly, needed to be activated and then subsequently suppressed as the two-staged prototype model predicts.

6. Conclusion

This study expanded on and presented experimental evidence in support of a well-known argument against prototype theories of concepts. Subjects do not default to the stereotypes of the conjuncts of a combined concept when interpreting a novel combination. This is hardly surprising since *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.

The one thing every English speaker knows (and knows that he knows) about purple apples is what compositionality tells him: purple apples are purple and apples. There is, if you like, a default interpretation of AN, but it is not the stereotype of A or N that one defaults to. Rather, it's the compositional reading *A and N.* Critics of the classical approach (including an anonymous reviewer of this paper) often protest that this theory manages to avoid being wrong only by abandoning hope of predicting anything at all. To the contrary, the classical combinatorics predict that increasing the string of modifiers will have no effect on such inferences as *purple apples are purple*, *purple apples are apples*, and so forth, both of which are warranted by the compositional structure that a Classical semantics assigns. But reiterated modification may well affect such prototypical inferences as *if it's an apple then it grew in the state of Washington and is sold in supermarkets*. These latter inferences are surely influenced by our typical past experience with apples, Washington, and supermarkets.

Appendix A. Target sentences

Candles	
A	Candles are made of wax.
B	Scented candles are made of wax.
C	Purple candles are made of wax.
D	Expensive purple candles are made of wax.
Caterpillars	
A	Caterpillars have many legs.
B	Furry caterpillars have many legs.
C	Poisonous caterpillars have many legs.
D	Canadian poisonous caterpillars have many legs.
Catfish	
A	Catfish have whiskers.
B	Edible catfish have whiskers.
C	Speckled catfish have whiskers.
D	Rare speckled catfish have whiskers.
Cellars	
A	Cellars are dark.
B	Damp cellars are dark.
C	Dry cellars are dark.
D	Dry northern cellars are dark.
Coins	
A	Coins are round.
B	Shiny coins are round.
C	Old coins are round.
D	Old Egyptian coins are round.
Coyotes	
A	Coyotes howl.
B	Carnivorous coyotes howl.
C	White coyotes howl.
D	Old white coyotes howl.
Crocodiles	
A	Crocodiles are dangerous.
B	Egg-laying crocodiles are dangerous.
C	African crocodiles are dangerous.
D	Albino African crocodiles are dangerous.
Doves	
A	Doves are white.
B	Feathered doves are white.

C	Brazilian doves are white.
D	Flightless Brazilian doves are white.
Ducks	
A	Ducks have webbed feet.
B	Quacking ducks have webbed feet.
C	Baby ducks have webbed feet.
D	Baby Peruvian ducks have webbed feet.
Hamsters	
A	Hamsters live in cages.
B	Pet hamsters live in cages.
C	Male hamsters live in cages.
D	Male hairless hamsters live in cages.
Kites	
A	Kites have strings.
B	Colorful kites have strings.
C	Silk kites have strings.
D	Silk weather kites have strings.
Lambs	
A	Lambs are white.
B	Fluffy lambs are white.
C	Norwegian lambs are white.
D	Long-haired Norwegian lambs are white.
Limousines	
A	Limousines are long.
B	Expensive limousines are long.
C	Inexpensive limousines are long.
D	Old inexpensive limousines are long.
Napkins	
A	Napkins are made of paper.
B	White napkins are made of paper.
C	Blue napkins are made of paper.
D	Blue checkered napkins are made of paper.
Nectarines	
A	Nectarines are juicy.
B	Sweet nectarines are juicy.
C	Bitter nectarines are juicy.
D	Bitter purple nectarines are juicy.
Ostriches	
A	Ostriches have long necks.
B	Ugly ostriches have long necks.

C	Paleolithic ostriches had long necks.
D	Paleolithic European ostriches had long necks.

Parkas

A	Parkas are warm.
B	Hooded parkas are warm.
C	Acrylic parkas are warm.
D	Machine-washable acrylic parkas are warm.

Pearls

A	Pearls are white.
B	Expensive pearls are white.
C	South Sea pearls are white.
D	Oval South Sea pearls are white.

Penguins

A	Penguins live in cold climates.
B	Flightless penguins live in cold climates.
C	Solitary penguins live in cold climates.
D	Solitary migrant penguins live in cold climates.

Pigeons

A	Pigeons live in parks.
B	Grey pigeons live in parks.
C	Tibetan pigeons live in parks.
D	Tibetan hunting pigeons live in parks.

Pigs

A	Pigs live on farms.
B	Dirty pigs live on farms.
C	Furry pigs live on farms.
D	Furry black pigs live on farms.

Ravens

A	Ravens are black.
B	Feathered ravens are black.
C	Jungle ravens are black.
D	Young jungle ravens are black.

Refrigerators

A	Refrigerators are used for storing food.
B	Kitchen refrigerators are used for storing food.
C	Commercial refrigerators are used for storing food.
D	Inexpensive commercial refrigerators are used for storing food.

Rhubarb

A	Rhubarb is used for pies
B	Sour rhubarb is used for pies.

C	Homegrown rhubarb is used for pies.
D	Homegrown Albanian rhubarb is used for pies.
Rifles	
A	Rifles are dangerous.
B	Hunting rifles are dangerous.
C	Used rifles are dangerous.
D	Used army rifles are dangerous.
Roosters	
A	Roosters live on farms.
B	Red roosters live on farms.
C	Yellow roosters live on farms.
D	Flying yellow roosters live on farms.
Saxophones	
A	Saxophones are made of brass.
B	Jazz saxophones are made of brass.
C	Handmade saxophones are made of brass.
D	Expensive hand-made saxophones are made of brass.
Scarves	
A	Scarves are worn for warmth.
B	Knitted scarves are worn for warmth.
C	Inexpensive scarves are worn for warmth.
D	Inexpensive synthetic scarves are worn for warmth.
Seaweed	
A	Seaweed is green.
B	Edible seaweed is green.
C	Baked seaweed is green.
D	Baked Indian seaweed is green.
Shacks	
A	Shacks are made of wood.
B	Storage shacks are made of wood.
C	New shacks are made of wood.
D	New Appalachian shacks are made of wood.
Shirts	
A	Shirts have buttons.
B	Cotton shirts have buttons.
C	Itchy shirts have buttons.
D	Itchy canvas shirts have buttons.
Sinks	
A	Sinks are found in kitchens
B	Metal sinks are found in kitchens.

C	Antique sinks are found in kitchens.
D	Round antique sinks are found in kitchens.
Sofas	
A	Sofas are found in living rooms.
B	Comfortable sofas are found in living rooms.
C	Uncomfortable sofas are found in living rooms.
D	Uncomfortable handmade sofas are found in living rooms.
Squirrels	
A	Squirrels eat nuts
B	Tree dwelling squirrels eat nuts.
C	Nicaraguan squirrels eat nuts.
D	Black Nicaraguan squirrels eat nuts.
Storks	
A	Storks have long legs.
B	White storks have long legs.
C	Domestic storks have long legs.
D	Domestic hybrid storks have long legs.
Strawberries	
A	Strawberries have seeds.
B	Red strawberries have seeds.
C	Lithuanian strawberries have seeds.
D	Sour Lithuanian strawberries have seeds.
Thimbles	
A	Thimbles are made of metal.
B	Sewing thimbles are made of metal.
C	Painted thimbles are made of metal.
D	Belgian painted thimbles are made of metal.
Tortoises	
A	Tortoises are slow.
B	Green tortoises are slow.
C	South American tortoises are slow.
D	South American fighting tortoises are slow.
Wagons	
A	Wagons are used by pulling them.
B	Cargo wagons are used by pulling them.
C	Futuristic wagons are used by pulling them.
D	Futuristic fruit wagons are used by pulling them.

Zebras

A	Zebras are fast.
B	Striped zebras are fast.
C	Namibian zebras are fast.
D	Giant Namibian zebras are fast.

References

Aizawa, K. (2003). *The systematicity arguments*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Armstrong, S., Gleitman, L., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13*, 293–308.

Chomsky, N. (1959). Review of skinner's verbal behavior. *Language, 35*, 26–58.

Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science, 8*, 27–58.

Cree, G. S., & 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. (1998). *Concepts: Where cognitive science went wrong*. New York, NY: Oxford University Press.

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

Fodor, J., & Lepore, E. (2002). *The compositionality papers*. Oxford: Clarendon Press.

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

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

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

Hampton, J. A. (1991). The combination of prototype concepts. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.

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.

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

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12*, 529–562.

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language, 29*, 259–288.

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. (Reprinted in E. Margolis & S. Laurence (Eds.). *Concepts: Core readings*, MIT Press, 1999.).

Partee, B. (1995). Lexical semantics and compositionality. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science*. Cambridge, MA: MIT Press.

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 research on conceptual combination. *Mind & Language, 10*, 72–104.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573–605.

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

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

Springer, K., & Murphy, G. L. (1992). Feature availability in conceptual combination. *Psychological Science, 3*, 111–117.

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review, 4*, 167–183.