RUNNING HEAD: Feature integration

Feature Integration in Natural Language Concepts

James A. Hampton
City University, London

Gert Storms
Catholic University of Leuven

Claire L. Simmons, Daniel Heussen
City University, London

Address for correspondence:
James A. Hampton
Psychology Department
City University
Northampton Square
London EC1V OHB, UK
hampton@city.ac.uk

Feature Integration in Natural Language Concepts

Abstract

Two experiments measured the joint influence of three key sets of semantic features on the frequency with which artifacts (Experiment 1), or plants or creatures (Experiment 2) were categorized in familiar categories. For artifacts, current function outweighed both originally intended function and current appearance. For biological kinds, appearance and behavior, an inner biological function, and appearance and behavior of offspring all had similarly strong effects on categorization. The data were analyzed to determine whether an independent cue model, or an interactive model best accounted for how the effects of the three feature sets combined. Feature integration was found to be additive for artifacts but interactive for biological kinds. In keeping with this, membership in contrasting artifact categories tended to be super-additive indicating overlapping categories, whereas for biological kinds it was sub-additive, indicating conceptual gaps between categories. The results underline a key domain difference between artifact and biological concepts.

Keywords: Concepts, Feature Integration, Domains, Categorization

Feature Integration in Natural Language Concepts

When an object is identified as being an instance of a conceptual class, what aspects of the object are involved in the decision, and how are those aspects combined? This question has been central to the field of concepts and categorization. Having rejected the classical notion that each concept has defining properties which provide a conjunctive classification rule (Smith & Medin, 1981), the prototype model of concepts (Rosch, 1975) held instead that concepts of general classes are represented in the mind by clusters of features constituting a prototype of the most characteristic or representative example of the class (for more recent accounts of prototype theory see Hampton, 1995; 2006). Categorization proceeds by determining how many of these features are possessed by any particular instance, or subtype, and giving a positive decision in the case that the weighted sum of positive features reaches some criterion value. Feature weights are assumed to reflect the "correlational structure" of the world, with greater weight given to features that have higher statistical association with category membership.

Critics of the prototype view (Murphy, 2002; Murphy & Medin, 1985; Rips, 1989) have rightly pointed out that there is much more to our conceptual knowledge than a simple clustering into kinds based on similarity. In particular, research has shown that statistical properties alone do not predict the degree to which people weight different features in conceptual categorization, especially when there are causal links between the features (Ahn & Dennis, 2001; Rehder, 2003; Sloman, Love & Ahn, 1998). Similarity is in any case a notoriously underspecified notion, since a similarity metric between objects can only be defined once the relevant aspects or features have first been identified, and the basis of similarity will thus change with each pair of concepts being considered, raising the specter of circularity or computational intractability (Hampton, 2001; Medin, Goldstone & Gentner, 1993; Smith & Medin, 1981). As an alternative to a model based on similarity, it has been argued that categorization in natural concepts is based primarily on deeper "theory-like" representations (Bloom, 1996; 1998; Murphy & Medin, 1985; Rips, 1989). For example an object may be a chair because it was the designer's intention when it was made that it would fulfill that particular role (Bloom, 1998; Gelman & Bloom, 2000; Jaswal, 2006; Matan & Carey, 2001; Rips, 1989). Or a creature may be a tiger because it has some essential property within its cells that causes it to look and behave like one

(Gelma, 2003; Medin & Ortony, 1989). While the evidence for deeper causal factors affecting categorization is strong, other studies have found that similarity of physical appearance is by no means ignored when categorizing either artifacts (Malt & Johnson, 1992) or biological kinds (Hampton, 1995; Hampton, Estes & S. Simmons, 2007; Hampton & Simmons, 2003). Thus if similarity-to-prototype were computed over both appearance features and deeper properties such as shared origins or inner structures, the simple notion that the probability of a categorization depends on the degree of match between the known characteristics of an individual object and a concept representation remains a strong theoretical contender. As argued in Hampton (2008), the theory-based and essentialist approaches still need an account of why categorization is fundamentally probabilistic and vague, and positing degrees of membership based on degree of match to conceptual contents is the best way to provide such an account.

The primary focus of the current research is on the probabilistic nature of categorization. Most natural categories are defined along multiple dimensions. A chair has a characteristic appearance, is made of certain materials, can be used in particular ways, and is created by certain processes. An apple has a visual appearance, taste and texture, internal biological processes and causal relations to the apple tree on which it grew, and the potential apple trees that may grow from its seeds. The issue to be addressed is how these different dimensions are combined in order to arrive at a categorization decision. Suppose that a fruit had the appearance of an apple, but was picked from a pear tree – how do people resolve this contradictory evidence in deciding whether it is an apple or whether it is a pear? Or consider the lighthouse/bell-tower in the small port of Collioure in South-West France (see Figure 1). Beginning in mediaeval times as a lighthouse to mark the entrance of the port, in the 18th Century with the demise of the port it was turned into a bell-tower for a chapel. How do people decide in this case what kind of thing it is? Is it still a lighthouse? Is it a bell-tower? When a Tuscan style cupola was added in the 19th Century did this change its status further? (See http://fr.wikipedia.org/wiki/Collioure).

Hampton (1995) conducted an investigation of this general question and found that people responded in a probabilistic fashion, taking account of all available information. The more apple properties that an object has, then the more people were likely to say that it is an apple. Our experiments aimed to investigate how this likelihood is determined. Does each feature have a constant additive effect, or does the effect of one feature vary as a function of the presence or absence of others?

The integration of information in the mind is a central issue that has been investigated in a wide range of situations, ranging from perceptual judgments (Ernst & Banks, 2002) to real-life decision making (Dhami & Harries, 2001). In relation to categorization a number of different models of feature integration have been proposed (e.g. Reed, 1972). Rosch (1975) suggested that similarity to prototype might be measured by counting the number of matching features – in itself a rather imprecise notion. Rosch and Mervis (1975) developed a method for calculating a “family resemblance score” as an operationalization of the notion of distance from prototype, and this score was shown to predict typicality within a category. Hampton (1979) suggested that a weighted sum of category features should be used to predict degree of membership, where the degree to which an object matches each category feature, multiplied by the feature’s weight for the category concept, is summed across the features of the concept. Feature weights are specific to each concept, but are constant across potential objects. These proposals are versions of the independent cue model for categorization (Reed, 1972) according to which each feature or cue provides a constant amount of weight to the decision, independently of the other features.

An alternative to the independent cue model is the set of models based on Medin & Shaffer’s context model (Medin & Shaffer, 1978; Nosofsky, 1988). In their model they proposed that degrees of match for each feature should be multiplied together to determine similarity. In the case of the context model, similarity was computed to a set of exemplars, but the same notion can readily be applied to calculating similarity to a prototype (Smith & Minda, 1998).

In order to determine whether features are being combined additively or non-additively, a simple test is to look at the influence of a given feature on categorization when other features are either present or absent. If the influence is the same, then the independent cue model fits the data best. If the influence is stronger in the presence of other features than in their absence, then a non-additive model, such as a multiplicative model is needed. In a preliminary test of this proposal, Hampton (1995) analyzed the data from four experiments in each of which six groups of participants categorized six different versions of concept instances. The six versions were constructed by using two sets of features. One set (based on appearance) was present or absent, while the second (based on function or biological essence) was either fully present, partially present or absent. When a feature was absent, it was replaced by a contrasting feature of a closely related category. Frequencies of "yes" responses were analyzed to determine whether the effect on categorization probability of changing each feature was independent of, or dependent on, the state of the other feature. For example, did changing the appearance of a creature have the same size of effect on categorization regardless of whether the essence feature was present, partially present, or absent? The results indicated that, contrary to the independent cue model, the effects of the two features were not additive. The effect of changing a feature was greater when the other feature was fully present than when it was partially or fully absent. However the test was not ideal, involving as it did pooling data from 4 experiments, with some items repeated between experiments. Nor was it possible to make any differentiation between different domains of concept. The goal of the current research was therefore to examine this phenomenon in a more systematic and controlled fashion. By using three binary features in a 2x2x2 rather than a 2x3 design, a more comprehensive test of feature independence was possible. The effect of removing any particular feature could be tested at three levels: with both the other features present, with just one other feature present, or with neither present.

We considered concepts in two broad domains: artifacts and biological kinds. There is already considerable evidence that these domains differ in many ways in terms of conceptual structure. As an early example, Keil (1986) found that transforming the appearance of an artifact would change its type, whereas transforming a biological kind would not. On the other hand making a discovery about a biological kind could affect its categorization, while similar discoveries did not affect artifact categorization. Further work by Kalish (1995) and Estes (2004) found that many people tend to consider membership in biological kind categories to be an objective fact so that a disagreement could be resolved by reference to an expert, whereas membership in artifact categories may be more a matter of subjective opinion. There is also interesting evidence of domain specific semantic aphasias which has been attributed to different patterns of correlation observed amongst the features of biological and artifact kinds (Tyler et al., 2000; Cree & McRae, 2003). Ruts, Storms & Hampton (2004) showed similarly that superordinate biological kinds have much tighter similarity clusters than artifact kinds, allowing the similarity space to be easily divided into linearly separable conceptual categories. We therefore predicted domain differences in how features combine. Biological kinds may be represented with a common-cause structure (Rehder, 2003) – an underlying genetic cause leads to inner functions and outward appearance, and to the appearance of offspring. With a common cause leading to three features, the absence of any one of those features may act strongly to reduce confidence in the cause still being present. Once this confidence has been reduced, the absence of a second feature will be less critical – the biggest drop in confidence occurs with the first sign of trouble. In contrast, artifact kinds may have a much looser causal structure. The artifact's appearance is related to its designer's intentions via a causal path, but we predicted that this influence would be much weaker. Because membership in artifact kinds is considered to be more a matter of subjective opinion (Estes, 2004; Kalish, 1995), we considered that each feature may have an independent effect on categorization, with people summing the evidence for an object being of one kind or another in a simple linear additive fashion.

Our first goal was to establish three different aspects of the concepts that would each affect categorization, so that their interaction could be tested. For artifacts there is evidence from both adult and developmental literature that the *function* of an artifact is the most important feature in determining its class (Bloom, 1998; Margolis & Laurence, 2007; Rips, 1989). A prototypical chair may broadly be defined as a movable object that is used for sitting on. In addition to current function, it has been argued that an even more crucial feature is the *function or use for which it was intended* (Bloom, 1996; 1998; Gelman & Bloom, 2000; Jaswal, 2006; Matan & Carey, 2001). Thus the fact that a craftsman constructed an object with the intention that it be a chair may over-ride the fact that for some reason the object cannot be sat upon, or that it is currently in use as a bed-side table. Evidence of the relative importance of current versus historically intended function is mixed. For example, Chaigneau, Barsalou & Sloman (2004) gave people different scenarios in which various features of an object such as a mop were independently manipulated. They found in their studies that a change in current function always had a greater effect on naming than did a change in original intended function. In addition to function, there is also evidence from Malt and Johnson (1992) that the *general appearance* of an object can affect its categorization. For example a chair is arguably differentiated from a stool more in terms of its appearance (having a back), than in terms of its function (enabling the action of leaning back). Malt and Johnson (1992) gave people descriptions of artifacts that either had unusual functions together with a normal appearance, or unusual appearance together with a normal function. In this study, participants actually placed more weight on appearance than on function. They confirmed the category membership of objects with normal appearance but unusual functions 58% of the time, but those with unusual appearance and normal functions only 25% of the time. This pattern obtained even when the story explicitly stated in the former case that the normal function was not served by the object.

For biological kinds, there are similarly three major aspects that may be considered crucial to determining type. *Prima facie*, the first aspect is the appearance of a plant or animal. Species may often be differentiated on the basis of their outward physical appearance and behavior. In addition it has been shown that people entertain essentialist beliefs about biological kinds (Gelman, 2003; Medin & Ortony, 1989). Two kinds of feature have been proposed that could figure in such beliefs (Hampton et al., 2007; Strevens, 2000). On the one hand, people may believe that the "innards" of the organism are crucially important (Gelman & Wellman, 1991). If, for example, some biochemical function is present that is characteristically only found within one species, this could be taken as strong evidence that the organism has the essence of that species. On the other hand people may believe that the genotype constitutes the essence of the organism, so that if the organism has offspring resembling a particular kind, that would constitute strong evidence of the true nature of the parent, and hence of how the organism should be categorized (Rips, 1989). Hampton et al. (2007) obtained evidence using Rips's (1989) transformation task that each of these views may be found in a student population.

In our experiments, we therefore aimed to construct materials in which three roughly equally weighted aspects of an object could be independently manipulated. For biological kinds in Experiment 1, they were (1) appearance and behavior of animals or appearance and taste/smell of plants (appearance for short), (2) internal biology (innards for short), and (3) appearance and behavior of the offspring (offspring for short). For artifacts in Experiment 2, these features were (1) appearance, (2) current function, and (3) originally intended function (or original function for short). By forming eight different descriptions of objects, plants or animals, corresponding to the presence or absence of each of the three features, we were able to measure the relative strength or importance of each type of feature as it affects categorization, and then to determine whether they combined in an additive or non-additive way.

It is worth noting that strong versions of psychological essentialism would predict that only one of the features of animals and plants should be critical. Although appearance features of biological kinds are indicators of essence (since genotype determines phenotype), they should be discounted when more is known about the innards or the offspring of the organism. People whose essentialist beliefs focus on the internal causal homeostasis of biological kinds (Boyd, 1999) should categorize according to what is known about the inner biological functions. Others who focus on genetic inheritance of essence should categorize on the basis of the appearance and behavior of offspring. Similarly a strong version of Bloom's thesis about artifact kinds would predict that originally intended function should outweigh any other features for artifact categorization (Bloom, 1998). Alternatively people may feel that the current use of an object determines its kind (see Keil, 1986). Finding that appearance affects categorization in both domains would therefore provide some additional evidence concerning the validity of these different views.

The experiments presented below were scaled-up versions of two studies presented in Hampton and Simmons (2003). To provide some background on the materials and method to be used, these earlier studies will be briefly reviewed. Each experiment used 8 pairs of biological kinds (4 plants and 4 animals) and 8 pairs of artifacts. Each concept pair consisted of two closely contrasting concepts – for example shark and whale, or tie and scarf. For each pair of concepts three sets of features were identified as above, each with two values – one for the first concept and the other for its contrast. The features were combined in all possible combinations to construct 8 possible exemplars, which were given to 128 participants to categorize. Participants were given a cover story about a nuclear accident on a large remote island (for biological kinds) or a secluded community in a remote area of Eastern Europe (for artifacts) and asked to classify each item. The two studies differed in whether participants gave a Yes/No judgement to each item with respect to one of the categories, or whether they chose the category in which the item was best placed.

To summarize the results, biological kinds showed strong and significant influences of all 3 features on categorization probability. Appearance and behavior of a creature, or appearance and smell or taste of a plant were considered important information for categorization over and above the biological innards and the offspring information. The materials were therefore well suited to the test of cue independence that was planned for Experiment 1 below, since a test for the interaction of features will require first that each feature has a reasonably strong individual effect on the probability of categorization. For artifacts, one feature, current function, dominated the rest, in keeping with Chaigneau et al. (2004). The appearance of an object had a very small effect on categorization and in neither study did it reach statistical significance. Original function had a minor influence and was only statistically significant in the second study. In order to provide materials for a test of cue independence, the artifact concepts for Experiment 2 were therefore adapted in an attempt to balance up the three types of feature. Current function was weakened by suggesting that the objects were now only rarely used – but that when they were employed it was solely with a particular function, and original function was boosted by stating that the object was both designed for that function and used to serve that function in the past.

Experiment 1

Method

Participants. Participants were 375 students at the Catholic University of Leuven, Belgium who each completed a booklet for course credit.

Design and Materials. The biological kinds used in Hampton & Simmons (2003) were extended and revised. There were 16 pairs of biological kind concepts, half plants and half animals (see Appendix). The pairs were chosen to be sufficiently similar for a hybrid possessing some features of each to be reasonably plausible (e.g. a crab versus a lobster). For each pair of concepts Appearance, Innards, and Offspring features[1] were created. Appearance was a set of features that included behavior for the creatures, and either smell or taste where appropriate for the plants. Innards referred to a biochemical property found in the creature that was specific to only one species. Offspring used the same set of appearance features but they were attributed to the offspring of the organism.

Thirty-two booklets were constructed, each containing a set of instructions to set the scenario and 16 different items to categorize. Half the booklets asked for categorization relative to concept A, and half relative to concept B of each pair. Illustrative examples may be seen in Appendix A, and the full list of pairs in Appendix B. Within each booklet there were 2 items for each of the 8 possible combinations of the 3 features. Items were rotated across feature conditions across booklets. Two orders of items in booklets were used. The first order had items randomly ordered within blocks for plants and for creatures, and the second was the reverse of the first. Materials were prepared in English and translated into Flemish Dutch by the second author. Instructions were as follows:

> "Many years ago, there was a nuclear accident near to a large remote island populated with a wide variety of animals. The accident resulted in its being contaminated by radiation. At some point in the future scientists are sent to investigate the long-term effects of this accident. They find and examine a number of individual creatures. Can you help them to decide what kind of creature each one is?"[2]

Results

Feature integration. Booklets were distributed to 384 students, 12 of each of the 32 different booklets in the design, and 375 of these were returned in usable form. In addition to the 9 missing booklets, there were 20 individual missing responses. Overall, missing data accounted for less than 3% of the data. Estimates of categorization probability used in the analyses were based on between 21 and 24 participants per cell. A further exclusion of data was unfortunately required because of a typographic error in the construction of the booklets, which meant that one of the 16 concept pairs (tiger-wolf) had to be dropped from the analysis. The results reported below are based on the remaining 30 concepts in 15 concept pairs.

Mean probabilities of categorization are shown in Table 1 together with standard deviation across items. The data for plants and animals were very similar, and so are reported together. (Animals had a slightly greater effect of appearance than plants – perhaps because behavior for animals is more salient than smell or taste for plants.) Probabilities were converted to z-scores for analysis (see Hampton, 1995). This transformation assumes that (1) people judge similarity of an item to the category prototype based on the descriptions given, (2) they say "Yes" if the similarity passes some threshold, and (3) the distance between the judged similarity of a given item and the threshold has a Gaussian distribution across individuals with constant variance. Probabilities of 1 were replaced by 23.5/24, and probabilities of 0 were replaced by 0.5/24 for this purpose. Mean and standard deviation for z-scores are also shown in Table 1. The three features were entered into a 3-way repeated measures ANOVA. All three main effects were significant. Mean (and standard deviation) effect sizes (in normal deviates *z)* for Appearance, Innards and Offspring were, respectively, 1.13 (0.28), 0.67 (0.22) and 1.42 (0.26). All main effects had $F(1,29)$ greater than 250, all $p < .001$. All four interactions were also significant, (Appearance x Innards, $F(1,29) = 4.4$, $p < .05$, Appearance x Offspring, $F(1,29) = 73.6$, $p < .001$, Innards x Offspring $F(1,29) = 22.9$, $p < .001$, and the 3-way interaction, $F(1,29) = 42.1$, $p < .001$). Breakdown analysis of the 3-way interaction showed that the interaction between Innards and Offspring features was only significant when Appearance was negative. Similarly Appearance interacted with Innards, but only when Offspring was negative. All of the interactions took the form of reduced effectiveness of a feature when the other was missing.

The pattern of significant interactions was consistent with a mode of combination of features in which a feature has more weight in the presence of other features. The independent cue model as adopted by early prototype models (Hampton, 1979; Reed, 1972) would predict that features contribute equally to similarity, independently of each other. The results suggest a dependence between the features, as would be found in a multiplicative model where the degrees of mismatch for each feature are multiplied together to determine dissimilarity (Medin & Shaffer, 1978; Smith & Minda, 1998).

A second analysis directly compared the effect of removing a single feature (measured in *z* transformed probability) when both other features were present, when only one was present, or when neither was present. The result is shown in the top panel of Figure 2. For all three features, the effect of removing the feature was greater when other features were present than when they were absent. There was no difference however between the case where just one other feature was present and the case where both were present. The pattern was confirmed with a significant effect of feature presence overall ($F(1.6, 47) = 49.3$, $p < .001$ with Greenhouse-Geisser correction), and for each individual feature (all $F > 20.2$, $p < .001$). In each case both linear and quadratic contrasts were significant.

A final check was run on whether the interaction effects could in part be the result of a high level of positive responses to the [– – –] stimulus (surprisingly there were 12% yes responses to this set of stimuli). Ten of the 30 concepts were selected with the constraint that the [+ + +] stimulus had at least 90% yes responses and the [– – –] stimulus had at least 90% no responses. The resulting data looked very similar, with a mean change in *z* of 1.43 when both features were present, 1.25 when just one was present and 0.72 when neither was present. The main effect of feature presence was significant ($F(2,18) = 14.9$, $p < .001$), again with a strong linear trend ($F(1,9) = 23.8$, $p < .001$).

Truth gaps and gluts. Having the same stimuli categorized by half the participants for one concept and by the other half for the other (contrasting) concept meant that it was possible to determine the degree to which the two probabilities sum to one. If features are combined multiplicatively, one should expect a stimulus that has some features of each concept to fall into a “conceptual gap” between the two concepts. For example a creature that had some squirrel features and some rabbit features may tend to be considered *neither* a squirrel *nor* a rabbit. To test for this, the observed probabilities of a creature being classified in either category were summed for each type of stimulus. Since the [+ – +] stimulus for squirrel was the [– + –] stimulus for rabbit (and so forth) the eight stimuli could be paired up into four possible conditions.

The interesting cases were those where a stimulus combined one feature of one concept with two features of the other. These creatures or plants were chimerical, having properties of more than one type. These cases were broken down into three kinds according to which one of the features was pitted against the other two. When either Appearance or Offspring was at odds with the other features, the sum of the two alternative categorizations was in each case significantly below 1 ($M = 0.85$, $SE = .03$, $t(29) = 5.1$, and $M = 0.81$, $SE = 0.03$, $t(29) = 6.6$, respectively, $p < .001$ in each case). There was a truth “gap”. Creatures or plants with inconsistent feature combinations were more likely to be rejected from both classes than included in both. (The pattern for plants and animals did not differ significantly).

Surprisingly the final set of cases – those in which Innards were opposed to Appearance and Offspring – did not show sub-additivity. The sum of probabilities was 0.999, ($SE = .02$) which was clearly not different from 1. When the hidden biological function was at odds with the observable facts about the creature and its offspring, then the degree to which the creature was categorized in one category was exactly matched by the degree to which it was *not* categorized in the other.

Discussion

Experiment 1 provided a powerful test of the way in which features are integrated in determining categorization for biological kind concepts. Thirty biological concepts were tested, with more than 20 participants categorizing each of the 8 stimuli for each concept. The results confirmed the significant part played by all 3 features in categorization of biological kinds. Appearance, innards and offspring all affected the likelihood of categorization. In addition the results confirmed that features of biological kinds are combined in an interactive way, consistent, for example, with a multiplicative rule. The effect of changing any of the features into that of its contrasting concept was greater when other features were present, and was much lower when both others were missing. As a result creatures or plants with inconsistent sets of features tended to fall between the two categories and were more likely to belong to neither kind than to belong to both.

Interestingly the drop in effect size with other features was not linear in Figure 2. There was little change between Both and One feature present, and then a large drop in effect size when both were absent. When both other features are absent, then a feature is on its own, and in a minority. While this attracts a certain number of “yes” responses, its influence on responding is quite small. When one other feature is present and the other is absent however, then the feature in question holds the “deciding vote”, turning the number of matching features from a minority to a majority, and hence the feature’s influence is much greater. Finally, when both other features are present, the feature in question is the first feature to show that the organism is odd in some way. It is here that the large effect size indicates that feature integration is non-additive. Even though the two other features still hold the majority vote, the effect of losing the first of three features is large, and is much greater than the effect of losing the last of three features.

Experiment 2

Experiment 2 followed the same design but used artifact concepts. Prior research by Hampton and Simmons (2003) failed to find convincing evidence that appearance was considered important in categorizing artifacts once original and current function were defined. In order to test the interaction between features it was therefore necessary to adapt the materials in order to balance up the features. Appearance was left as it was, but Current Function was downplayed by stating that the object was now only rarely used – but that when it was used it only had that function. At the same time, Original Function was strengthened by stating that the object was not only designed to serve a given function, but that it did originally serve that function. Of course if Appearance is truly irrelevant to artifact categorization we would not expect it to have an influence on categorization here. However given earlier results of Malt and Johnson (1992) there was reason to suppose that all three features would affect categorization, enabling additivity to be tested.

Method

Participants. Participants were 320 students at the Catholic University of Leuven who participated for course credit. No booklets were returned incomplete.

Design and Materials. The preparation of materials followed exactly the same design as in Experiment 1, except that 16 pairs of artifact concepts were used in place of 16 pairs of biological kinds, and the three features manipulated were Appearance, Original Function, and Current Function. Pairs of similar concepts were chosen so that the features could plausibly be swapped between them (see Appendix A for examples, and Appendix B for the full list of pairs). Materials were prepared in English and translated into Flemish Dutch by the second author. Note that Original Function and Current Function were deliberately the same in the case where the two features were both positive (or both negative). Most ties were originally made to be worn round the neck and continue to serve that function. In order to reduce the dominant strength of Current Function as a feature, the current function of the objects was made to sound occasional. For example the putative church was "occasionally used for Christian services, and has no other function", or the putative tie was described: "now, when used at all, it is only ever worn with shirts and suits by male members of the community as a part of formal dress." Instructions included a scenario to lend some degree of plausibility as follows:

> "Anthropologists visited a secluded community in a remote area of Eastern Europe, where they found and studied a number of cultural artifacts. The members of the community were very resourceful and had found ways of sometimes adapting things to new uses. The anthropologists were puzzled about how each item should be

classified. Can you help them to decide what kind of thing each one is?”

Results

Feature Integration. Probability of categorization for each stimulus in each of the 32 concepts was estimated from the frequency of “yes” responses, and results are tabled in Table 1, along with corresponding z-score data. Examination of individual pairs of concepts revealed an unanticipated effect in the case of the concept pair Theatre/Cinema. In Flemish (as in US English) a cinema may also be called a theatre, so that even when the object had only cinema features it was still categorized as a theatre by 70% of the participants. (In the UK the term “theatre” means primarily a place for live performances of plays.) The item THEATRE was therefore excluded from the analysis. The contrasting category (CINEMA) could however still be used, since theatres used for live plays are never called cinemas in Belgium.

Frequencies for the remaining 31 concepts were converted to z scores. Since there were 20 responses per probability estimate, probability values of 0 were taken as 0.5/20, and values of 1 as 19.5/20. The z scores were submitted to a 3-way repeated measures ANOVA across items with features as factors. All 3 features had strongly significant effects on categorization frequency. Main effects (and standard errors) in z score differences and their F ratios were: Appearance, $z = 0.45$ (.07), $F(1,30) = 44.2$, Original Function, $z = .91$ (.06), $F(1,30) = 238.8$, and Current Function, $z = 1.89$ (.07), $F(1,30) = 698.1$, all $p < .001$. In contrast to the biological kinds in Experiment 1 there was less evidence that the features interacted. Original Function did not interact significantly with Current Function ($F(1,30) = 3.56$, $p = .07$) or with Appearance ($F < 1$), but Current Function and Appearance did interact ($F(1,30) = 10.6$, $p < .005$). As in earlier experiments, the interaction showed that one feature had a greater effect when the other was present than when it was absent. The 3 way interaction was marginal but not significant ($F(1,30) = 3.42$, $p = .07$).

As in Experiment 1, the effect of removing one feature (i.e. replacing it with that of the contrasting concept) was measured when both, just one, or neither of the other features were present. The results are displayed in the lower panel of Figure 2. Unlike the biological kinds, there was no clear trend for the changing of a feature to have a greater effect when the other features were present. Degree of change in z was entered into a 2-way ANOVA with feature (3 levels) and presence/absence of the other features (3 levels) as repeated measures factors across the 31 concepts. There was no overall main effect of presence/absence of other features ($F(2,60) = 1.89$, $p = .16$), but there was an interaction of this factor with type of feature ($F(1.8, 53) = 7.8$, $p < .005$, with Greenhouse-Geisser correction for significant lack of sphericity). The interaction can be seen in Figure 2 (lower panel). Breakdown analysis of the interaction confirmed that Appearance was the only factor that had reduced effectiveness when the other factors were absent (linear trend $F(1,31) = 5.8$, $p < .05$). Neither of the other two features showed any effect of presence/absence of other features when considered alone (linear trend $F(1,30) = 2.1$, for Original Function and $F < 1$ for Current Function, both $p > .15$).

Finally, given the very different pattern of results from Experiments 1 and 2, a direct comparison was made between them (the design, language and subject populations were the same in each experiment). Since the features did not correspond between domains (with the exception of Appearance), the three features were collapsed within each domain and an ANOVA was run across the 61 items (30 biological kinds and 31 artifacts), with presence of other features as a repeated measures factor with 3 levels (both present, just one present, neither present), and Domain (Experiment) as a between items factor with 2 levels, artifacts and biological kinds. The effect of interest was the interaction between Feature Presence and Domain, and this interaction proved highly significant ($F(1.7,102) = 16.6$, $p < .001$, with Greenhouse-Geisser correction for sphericity).

Truth gaps and gluts. In Experiment 1 there was significant sub-additivity when the likelihood of an inconsistent stimulus being in one biological kind category was added to the likelihood of its being in the contrasting category. A similar analysis was conducted for the artifacts in Experiment 2. The observed probabilities of an object being classed in either category were summed for each pair of concepts to yield 4 summed probabilities.

Hybrid cases where a stimulus combined one feature of one concept with two features of the other were again broken down according to which one of the features was pitted against the other two. When Appearance was at odds with Current and Original Function, or when Current Function was at odds with the other two features, then categorization was still additive, with summed probabilities of 0.95 ($SE = .03$) and 0.99 ($SE = .03$) respectively not significantly less than 1 ($t(30) = 1.86$, $p = .07$, and $t(30) = 0.4$, $p > .5$). However when Original Function contradicted Appearance and Current Function, there was a significant tendency for categorization to be *super-additive,* with a summed probability of 1.09 (*SE* = .03, $t(30) = 3.37$, $p < .005$). This result was in stark contrast to the biological kinds, where the general trend was for categorization to be *sub-additive*.

Whereas the biological kinds in Experiment 1 had tended towards a truth *gap*, so that items falling between two concepts were likely to be considered not to belong to either, in Experiment 2 items lying between two artifact concepts showed no truth gaps. Likelihood of being in one category was well predicted by the likelihood of *not* being in its contrast, and where the data deviated from this pattern, items with mixed features were more likely to be categorized in *both* categories than to be placed in neither – a so-called truth *glut* (Bonini et al., 1999).

Discussion

The results of Experiment 2 contrasted strongly with those from the first experiment. With minor exceptions, the effect of changing one of the features of an object—be it the Original Function, the Current Function or the Appearance—was equivalent regardless of the other properties of the object. As a consequence there were no truth gaps between concepts. Indeed there was some evidence for objects falling into more than one class at the same time. As discussed above, there is evidence that artifact categorization is much less based on underlying causal schemas than is the case for biological kinds.

It is also interesting to note that all three types of feature had a role to play in categorization. Given the strong advocacy of functions as the basis of defining artifact types, it was interesting that an object with the wrong appearance was not as well accepted as one with the correct appearance, even when both original and current function were the same. Probability of categorization decreased from 95% to 84% when the appearance mismatched the category. So at least for some items and some participants appearance was enough to over-rule function. Of the 31 items, 22 had reduced categorization probability when appearance was the only mismatching feature.

General Discussion

The major question driving the research concerned the way in which features are integrated when judging category membership. Two possibilities were considered – that features contribute <u>independently</u> to the similarity of an instance to a category concept (and hence its probability of categorization), and that features <u>interact</u> in their effect. Our results demonstrated very clearly that different systems appear to be in operation for artifact and for biological kind categories. For artifact categories, the effect on categorization of altering a feature was the same, regardless of whether the other two features were present or absent. In line with this result, when an instance lay between two different categories (having some features of each), the probability of being in both categories (as measured independently) slightly exceeded 1. In contrast, biological kind categories showed an interactive pattern of feature integration. The effect of altering a feature was much greater when others were present than when both were absent. As a consequence, the likelihood of an intermediate instance falling in both of two contrasting categories tended to be less than 1.

The different pattern of integration for artifacts and biological kinds is perhaps the most important result from this research. Discussion of categorization models in the literature has tended to assume that one model for relating feature possession to categorization should fit all cases (Nosofsky, 1988; Smith & Minda, 1998). While there has been much discussion of important differences between artifact and biological kind domains (e.g. Estes, 2004; Gelman, 2003; Kalish, 1995; Keil, 1986) this is the first clear demonstration that the way in which information is integrated in the two general domains is different. With the benefit of hindsight, it is not difficult to provide some plausible accounts of the difference we have observed. Artifact kinds lack an underlying network of causally linked properties. In fact Sloman & Malt (2003) have argued that artifact categories are not true conceptual categories at all, but correspond more closely to "naming" categories – items that for one reason or another have happened to end up with the same name. The loose and overlapping landscape of artifact categories lends itself readily to the notion that instances that fall between two categories could be considered to be in both, rather than in neither.

On the other hand, our beliefs about biological kinds tend to include the notion that there is a strong set of causal principles within each organism that lead to the homogeneity of the class as a whole. Boyd (1999) referred to this notion as "causal homeostasis". At least in folk understanding of biological kinds the classes represent tight clusters of similar items with large gaps in between. Even relatively close categories like foxes, wolves and husky dogs are assumed to form easily distinguishable categories in terms of their appearance, and it is assumed that underlying the similarity of appearance is some deeper causal story involving innards and germ lines. In this domain, it makes sense that some individual creature which had the appearance of one type of animal, but the innards of another should be considered to belong to neither category, rather than to both. The interactive pattern of feature integration found for biological kinds reflects the integrated nature of the features. The first feature to be altered (be it appearance, innards or offspring) immediately casts doubt on whether the organism has the full set of interlocking features that characterize a "proper" member of the kind. The effect of an altered feature is therefore greatest when the others are present.

In order to conduct the test of additivity it was first necessary to find different sets of semantic features that determine the likelihood of an instance being placed in a class. Some comment is worthwhile on the relative weights observed. For artifacts current function or use was clearly the strongest influence, but there was still an effect of the other two types of feature. Direct comparison is of course difficult, given that the features may differ in diagnosticity and centrality (Sloman, Love & Ahn, 1998). In addition the appearance features referred to the whole set of perceptual properties of an object (visual, dynamic, taste and smell), whereas the inner biological function was a much more specific single feature. The finding that original function is not the only factor affecting artifact categorization may at first glance appear to contradict the position advocated by theorists such as Bloom (1996; 1998) who have argued that the kind of an artifact is determined exclusively by its creator's intention. If a designer had it in mind to create a chair, it should not matter what the object looks like, or whether it can be sat upon, it is still a chair. In contrast we found that the original intended function of our artifacts played a relatively minor role in determining categorization. A way to resolve the issue would be to amend Bloom's thesis. In our covering story we explained how objects had become adapted to new purposes, and it would therefore seem plausible that the kind of an object is not determined by its first designer, but by its most recent designer. When the people in our story took a church and started to use if for art exhibits, (so that it no longer was used as a church) they were thereby "re-baptizing" the object as a new kind. This interpretation would also fit with Keil's study of transformations of artifacts (Keil, 1986). When children were told of how a metal coffee pot was reshaped and hammered so it looked like and could be used as a bird feeder, they were happy to allow that it was no longer a coffee pot. The majority of our participants would appear to agree. Nonetheless, even when categorizing by current function, the appearance of the object continued to have an effect.

For biological kinds, a common view is that categorization is driven by some notion of “essence” (Gelman, 2003; Murphy, 2002). Hampton et al. (2007) argued that there are two discrete notions of essence. One relates to the causal processes at work deep within an organism that lead to its appearance and behavior, and the other relates to the notion of a germ-line that is passed from parent to offspring. In our scenarios, each of these two types of information was available to indicate whether the essence was that of the category in question. The data were very clear in indicating that not only both of these more essentialist criteria, but also the appearance and behavior of the organism itself were all treated as relevant sources of information for categorization. There was no clear “winner” in terms of “innards” essentialism versus “germ-line” essentialism or for that matter outward appearance features.

Conclusion

The results presented here constitute the first demonstration of a key difference between biological and artifact kinds. For biological kinds the evidence for kind-ship is integrated in a non-additive fashion, with the result that a chimeric creature with aspects of more than one species is more likely to be classified in neither than in both. For artifact kinds, features are apparently combined in an additive fashion, and items with hybrid features may be more likely to be in both categories than in neither. Returning to the problem of the lighthouse-bell tower in Collioure, a Google search on 28/8/8 gave 17,800 hits for “Collioure phare” (lighthouse) and 21,900 hits for “Collioure clocher” (bell-tower). Current function has it by a short whisker.

References

Ahn, W. K., & Dennis, M. J. (2001). Dissociation between categorization and similarity judgment: Differential effect of causal status on feature weights. In Unhand & M. Ramscar (Eds.), *Similarity and Categorization* (pp. 87-107). New York, NY: Oxford University Press.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60,* 1-29.

Bloom, P. (1998). Theories of artifact categorization. *Cognition, 66,* 87-93.

Bonini, N., Osherson, D. N., Viale, R., & Williamson, T. (1999). On the psychology of vague predicates. *Mind and Language, 14,* 377-393.

Boyd, R. (1999). Homeostasis, Species, and Higher Taxa. In R.A.Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141-186). Cambridge: MIT Press.

Chaigneau, S. E., Barsalou, L. W., & Sloman, S. A. (2004). Assessing the causal structure of function. *Journal of Experimental Psychology: General, 133,* 601-625.

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.

Dhami, M. K., & Harries, C. (2001). Fast and frugal versus regression models of human judgement. *Thinking and Reasoning, 7,* 5-27.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429-433.

Estes, Z. (2004). Confidence and gradedness in semantic categorization: Definitely somewhat artifactual, maybe absolutely natural. *Psychonomic Bulletin and Review, 11,* 1041-1047.

Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought.* Oxford: Oxford University Press.

Gelman, S. A., & Bloom, P. (2000). Young children are sensitive to how an object was created when deciding what to name it. *Cognition, 76*, 91-103.

Gelman, S. A., & Wellman, H. (1991). Insides and Essences: Early Understandings of the Non-obvious. *Cognition, 38*, 213-244.

Hampton, J. A. (1979). Polymorphous Concepts in Semantic Memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 441-461.

Hampton, J. A. (1995). Testing Prototype Theory of Concepts. *Journal of Memory and Language, 34,* 686-708.

Hampton, J. A. (2001). The Role of Similarity in Natural Categorization. In M. Ramscar, U. Hahn, E. Cambouropolos, & H. Pain (Eds.), *Similarity and categorization* (pp. 13-28). Oxford: Oxford University Press.

Hampton, J. A. (2006). Concepts as Prototypes. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 46* (pp. 79-113). Amsterdam: Elsevier.

Hampton, J. A. (2008). Concepts in human adults. In D. Mareschal, P. Quinn, and S.E.G. Lea (Eds.) *The Making of Human Concepts.* Oxford: Oxford University Press (in press).

Hampton, J. A., Estes, Z., & Simmons, S. G. (2007). Metamorphosis: Essence, Appearance and Behavior in the Categorization of Natural Kinds. *Memory & Cognition, 35,* 1785-1800.

Hampton, J.A., & Simmons, C.L. (2003). Feature Independence in Natural Categories.

Poster presented to the Annual Convention of the Psychonomic Society, Vancouver BC, November.

Jaswal, V. K. (2006). Preschoolers favor the creator's label when reasoning about an artifact's function. *Cognition, 99,* B83–B92.

Kalish, C. W. (1995). Essentialism and graded membership in animal and artifact categories. *Memory & Cognition, 23*, 335-353.

Keil, F. C. (1986). The acquisition of natural kind and artifact terms. In W. Demopoulos & A. Marras (Eds.), *Language learning and concept acquisition: Foundational issues* (pp. 133-153). Norwood, NJ: Ablex.

Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory and Language, 31,* 195-217.

Margolis, E. & Laurence, S. (2007). *Creations of the mind: Theories of artifacts and their representation*. Oxford: Oxford University Press.

Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition*, *78*, 1-26.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 179-195). Cambridge: Cambridge University Press.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254-278.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence.

Psychological Review, 92, 289-316.

Murphy, G.L. (2002). *The Big Book of Concepts*. Cambridge MA: MIT Press.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 700-708.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3,* 382-407.

Rehder, B. (2003). Categorization as Causal Reasoning. *Cognitive Science, 27,* 709-748.

Rips, L. J. (1989). Similarity, typicality and categorization. In S.Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 21-59). Cambridge: Cambridge University Press.

Rosch, E. R. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192-232.

Rosch, E. R., & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Ruts, W., Storms, G., & Hampton, J. A. (2004). Linear Separability in Superordinate Natural Language Concepts. *Memory & Cognition, 32,* 83-95.

Sloman, S. A., & Malt, B. C. (2003). Artifacts are not ascribed essences, nor are they treated as belonging to kinds. In H. E. Moss & J. A. Hampton (Eds.), *Conceptual Representation* (pp. 563-582). Hove: Psychology Press.

Sloman, S. A., Love, B. C., & Ahn, W. K. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22*, 189-228.

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

U. P.

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1411-1436.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition, 74*, 149-175.

Tyler, L. K., Moss, H. E., Durrant-Peatfield, M. R., & Levy, J. P. (2000). Conceptual structure and the structure of concepts: A distributed account of category-specific deficits. *Brain and Language*, *75*, 195-231.

Appendix A: Sample of materials used in Experiment 1 (Biological kinds) and 2 (Artifacts)

Biological kinds

Concept A	Concept B	Feature-type	Concept A Feature	Concept B Feature
Crab	Lobster	Appearance	A creature with legs and claws that looks and acts just like a crab	A creature with a long tail and claws that looks and acts just like a lobster
		Innards	The scientists found that the structure of its eyes was identical to that typically found only in crabs	The scientists found that the structure of its eyes was identical to that typically found only in lobsters
		Offspring	They found that the creature had offspring that looked and acted just like crabs	They found that the creature had offspring that looked and acted just like lobsters
Mosquito	Wasp	Appearance	A small flying insect with transparent wings that bites people, and looks and acts just like a mosquito	A striped flying insect that stings people, and looks and acts just like a wasp
		Innards	The scientists found that the chemistry of its blood was just like that normally only found in mosquitoes	The scientists found that the chemistry of its blood was just like that normally only found in lobsters
		Offspring	They found that the eggs laid by the creature developed into offspring that looked and acted just like mosquitoes	They found that the eggs laid by the creature developed into offspring that looked and acted just like wasps
Oak	Pine	Appearance	A tall tree that loses its leaves in winter and that looks just like an oak	A tall tree that keeps its needles all year round and that looks just like a pine
		Innards	The scientists found that the micro-structure of the wood fibers was just like that only typically found in oaks	The scientists found that the micro-structure of the wood fibers was just like that only typically found in pines
		Offspring	They found that when the tree reproduced, new trees grew that looked just like oaks	They found that when the tree reproduced, new trees grew that looked just like pines

Concept A	Concept B	Feature-type	Concept A Feature	Concept B Feature
Grape	Cherry	Appearance	A small round green fruit which looks and tastes just like a grape	A small, dark red fruit, which looks and tastes just like a cherry.
		Innards	The scientists found that its cellular potassium metabolism is just like that which is normally only observed in grapes.	The scientists found that its cellular potassium metabolism is just like that which is normally only observed in cherries.
		Offspring	They found that when the seeds of this fruit are planted, a vine grows, yielding fruit which look and taste just like grapes.	They found that when the seeds of this fruit are planted, a tree grows, yielding fruit which look and taste just like cherries.

Artifacts

Concept A	Concept B	Feature-type	Concept A Feature	Concept B Feature
Church	Art Gallery	Appearance	A large building with stained glass windows, and a steeple with a cross on the top, which looks just like a church.	A large gothic building with white interior walls on which paintings are hung, and which looks just like an art gallery.
		Original Function	It was originally built just to be a place of Christian worship, and had that function in the past.	It was originally built just to be an exhibition hall for displaying large works of art, and had that function in the past.
		Current Function	It is presently occasionally used for Christian services, and has no other function.	It is presently occasionally used for the public exhibition of painting and sculpture, and has no other function.

Concept A	Concept B	Feature	Concept A	Concept B
Bank-notes	Postage stamps	Appearance	Rectangular pieces of paper with a colored design and an embedded metallic strip which look just like bank notes.	Small rectangular pieces of paper with serrated edges and sticky backs which look just like postage stamps.
		Original Function	Originally, these were produced as a kind of money, and they served that function in the past.	Originally, these were produced for sticking on letters as postage, and they served that function in the past.
		Current Function	They are now **just** sometimes used for buying or selling things, and have no other use.	They are now **just** sometimes stuck to envelopes to pay for postage, and have no other use.
Taxi	Ambulance	Appearance	A motor vehicle which is black with a yellow light on the top, has a diesel engine and looks just like a London taxi.	A motor vehicle which is white and green with a flashing blue light on the top, and which looks just like an ambulance.
		Original Function	It was originally intended and used to provide transport for small groups of people to their desired destination.	It was originally intended and used to carry sick or injured people to hospital for urgent medical attention.
		Current Function	Now, when it is used, its only use is to take people wherever they want to go in exchange for money.	Now, when it is used, people use it **only** in the case of medical emergencies when the driver takes people to the hospital
Tie	Scarf	Appearance	An item sewn from a long piece of patterned silk fabric which looks just like a man’s tie.	An item which is made of a long thin piece of knitted wool, and looks just like a scarf.
		Original Function	Originally, it was intended to be tied around the collar of a shirt as a form of decoration, and in the past it had this function.	Originally, it was intended to be wrapped around the neck for protection against the cold when outside, and in the past it had this function.
		Current Function	Now, when used at all, it is **only ever** worn with shirts and suits by male members of the community as a part of formal dress.	Now, when used at all, it is **only ever** worn round the neck by members of the community for keeping warm when outdoors in winter.

Appendix B: Full set of Concept pairs used in Experiments

Biological Kinds (Expt 1)		Artifacts (Expt 2)	
Concept A	Concept B	Concept A	Concept B
Crab	Lobster	Church	Gallery
Mosquito	Wasp	Banknote	Stamp
Pigeon	Crow	TV	Monitor
Lizard	Snake	Cinema	Theatre
Tiger	Wolf	Ferryboat	Warship
Shark	Dolphin	Taxi	Ambulance
Rabbit	Squirrel	Tie	Scarf
Horse	Cow	Vase	Carafe
Rose	Dandelion	Saucepan	Helmet
Mint	Onion	Beer glass	Jar
Oak	Pine	Sketchbook	Diary
Grass	Moss	Roof tile	Drain cover
Grape	Cherry	Drum	Waste basket
Apple	Orange	Nightshirt	Dress
Pumpkin	Watermelon	Rug	Blanket
Carrot	Potato	Chimneypot	Flowerpot

Author Note

Enquiries concerning this paper should be addressed to James A. Hampton, Department of Psychology, City University, Northampton Square, London EC1V OHB or by email to hampton@city.ac.uk. We acknowledge the helpful comments of contributors to the 3rd Leuven Workshop on Concepts and Categories, held in Leuven in 2006.

Footnotes

1. We refer to the three aspects as “features” for simplicity of exposition, although in fact each aspect may be composed of multiple features. Since the aim was to contrast appearance with deep properties, appearance features were taken together as a single set.
2. Although through an oversight the instructions did not mention plants, in fact half the biological kinds included were plants. No participant mentioned that they noticed this omission, and as there was no important difference between responses to plants and to animals, it was assumed that they took the instructions to apply equally to all the items in this section of the booklet.

Table 1

Mean (and SD across Items) for Probability of Categorization and z-scores for each of the 8 Types of Exemplar in each Experiment.

Experiment 1: Biological kinds (*N =30)*

	Offspring appearance +				Offspring appearance –			
	Innards +		Innards –		Innards +		Innards –	
	Appearance +	Appearance –	Appearance +	Appearance –	Appearance +	Appearance –	Appearance +	Appearance –
Probability	.96 (.04)	.63 (.14)	.84 (.09)	.30 (.11)	.51 (.13)	.16 (.07)	.22 (.09)	.12 (.08)
z-score	1.77 (.30)	0.35 (.38)	1.06 (.39)	-0.56 (.38)	0.03 (.33)	-1.04 (.31)	-0.83 (.36)	-1.24 (.44)

Experiment 2: Artifacts (*N =31)*

	Current Function +				Current Function –			
	Original Function +		Original Function –		Original Function +		Original Function –	
	Appearance +	Appearance –	Appearance +	Appearance –	Appearance +	Appearance –	Appearance +	Appearance –
Probability	.95 (.07)	.84 (.13)	.79 (.12)	.59 (.18)	.41 (.14)	.30 (.17)	.11 (.09)	.07 (.08)
z-score	1.65 (.40)	1.11 (.53)	0.92 (.51)	0.24 (.55)	-0.21 (.41)	-0.59 (.59)	-1.27 (.54)	-1.46 (.55)

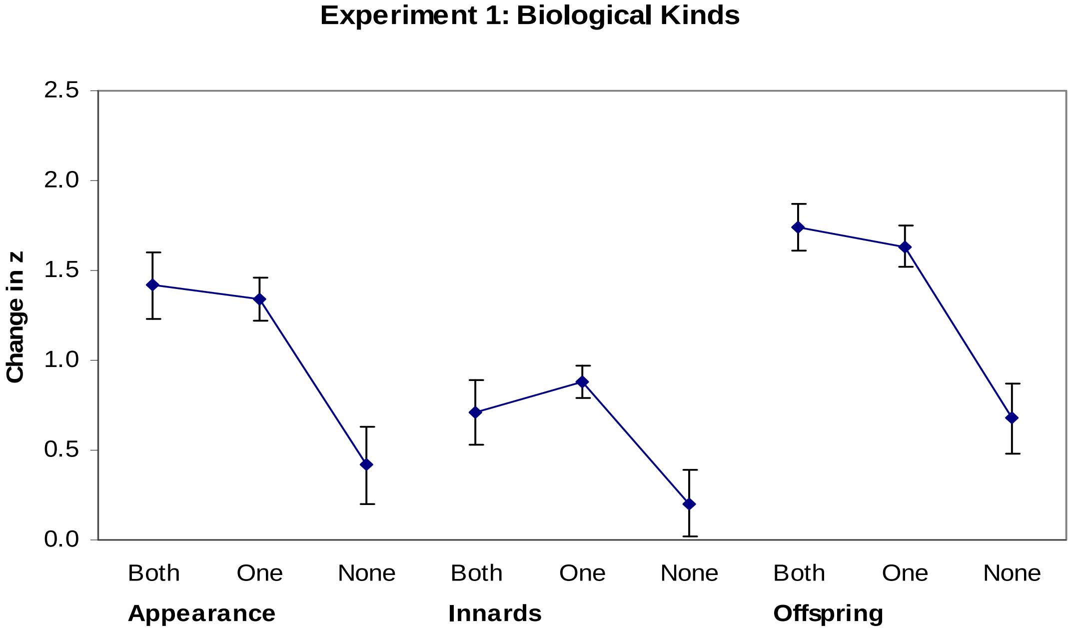
Figure Captions

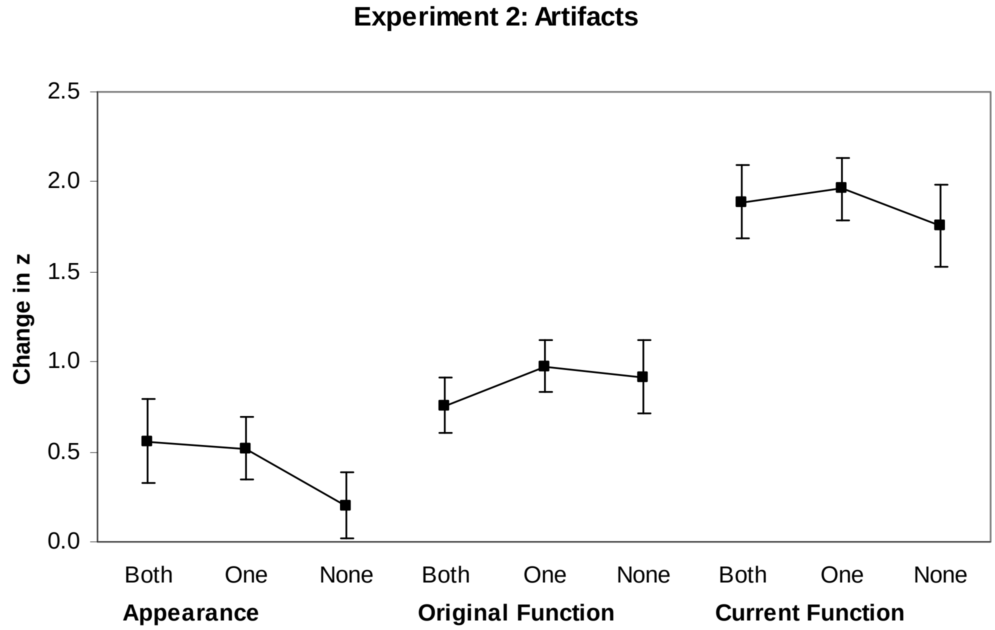
Figure 1. The light-house bell-tower in Collioure, South-West France.

Figure 2. Effect of changing each feature in Experiment 1 (Biological kinds) and Experiment 2 (Artifacts) on the z transformed probability of categorization when both, just one or neither of the other two features are present.

Figure 1

Figure 2





Note: Figure shows the change in categorization probability (z-transformed) resulting from a change in the named Feature-type when either Both, just One or None of the other features are present.