Distinctiveness and Correlation in Conceptual Structure: Behavioral and Computational Studies

Billi Randall, Helen E. Moss, Jennifer M. Rodd, Mike Greer, and Lorraine K. Tyler
University of Cambridge

Patients with category-specific deficits have motivated a range of hypotheses about the structure of the conceptual system. One class of models claims that apparent category dissociations emerge from the internal structure of concepts rather than fractionation of the system into separate substores. This account claims that distinctive properties of concepts in the living domain are vulnerable because of their weak correlation with other features. Given the assumption that mutual activation among correlated properties produces faster activation in the normal system, the authors predicted a disadvantage for the distinctive features of living things for unimpaired adults. Results of a speeded feature verification study supported this prediction, as did a computational simulation in which networks mapped from orthography to semantics.

One of the most intriguing consequences of certain kinds of brain damage is that patients' conceptual knowledge may be selectively impaired for specific categories of objects while others remain relatively preserved. The most frequently reported dissociation is a deficit for living things such as fruits, vegetables, and animals relative to nonliving things (e.g., Sheridan & Humphreys, 1993; Warrington & Shallice, 1984). The reverse pattern, in which conceptual knowledge of nonliving things is more severely impaired, has also been observed (Moss & Tyler, 2000; Warrington & McCarthy, 1987; for reviews, see Gainotti, Silveri, Daniele & Giustolisi, 1995; Saffran & Schwartz, 1994). This double dissociation forces us to ask how conceptual knowledge is organized such that damage to the system can result in selective impairment of living or nonliving things.

A number of theories have been proposed to account for this double dissociation. One type of account is that conceptual knowledge is explicitly organized into domain-specific subsystems, each of which may be selectively impaired by brain damage (e.g., Caramazza & Shelton, 1998; Santos & Caramazza, 2002). However, the breakdown of conceptual knowledge does not always respect domain boundaries. For example, some patients with a living things deficit also have difficulties with musical instruments, gemstones, or food, whereas other patients with a nonliving things deficit show poor performance on parts of the body (e.g., Silveri & Gainotti, 1988; Warrington & Shallice, 1984). This has motivated a second class of account in which the major organizing principle of the conceptual system is type of semantic feature rather than category or domain per se. The simplest version of this account proposes two semantic subsystems, one for sensory and one for functional features. On the assumption that objects in the living domain are primarily distinguished by sensory features while nonliving things rely on functional features, these domains will be disproportionately impaired by damage to the sensory and functional subsystems, respectively. This would explain why certain categories may not pattern with others in their domain; for example, if musical instruments and gemstones are distinguished mainly on sensory rather than functional grounds, then they should pattern with living things rather than nonliving things. However, the sensory–functional account does not accommodate all of the neuropsychological data. Most importantly, the predicted association between a loss of sensory features and a living things deficit is absent in many patients (e.g., Laiacona, Barbarotto, & Capitani, 1993; Lambon Ralph, Howard, Nightingale, & Ellis, 1998; Moss, Tyler, Durrant-Peatfield, & Bunn, 1998).

Partly in response to the limitations of these accounts, there have been recent attempts to determine the extent to which patterns of category-specific deficits can be accounted for by damage to a single, unitary distributed semantic system (Durrant-Peatfield, Tyler, Moss, & Levy, 1997; Gonnerman, Andersen, Devlin, Kempler, & Seidenberg, 1997; Tyler & Moss, 2001; Tyler, Moss, Durrant-Peatfield, & Levy, 2000). These distributed accounts share the assumption that a concept is represented by a set of features that may be shared by a large number of concepts (e.g., many things are white or green) or only by a few (e.g., few objects have a serrated edge or are spotted). The crucial aspect of these models is that they assume that any apparent differences between domains or categories emerge because of differences in the internal structure of the different types of concepts rather than fractionation of the conceptual system into separate substores according to category of object or type of feature.

In this article we focus on one distributed account—the conceptual structure account—to test one of its core and distinctive assumptions. In keeping with other similar models, the conceptual structure account (Tyler & Moss, 2001; Tyler et al., 2000) claims that each concept is represented by a set of features, some of which are distinctive (occurring only in one or two concepts) and some of which are shared across many concepts. Earlier studies by Keil

(1986) and Malt and Smith (1984) claimed that features tend to be distributed differently across concepts, with the pattern of correlations between features (i.e., which features co-occur with each other) differing systematically across concepts. These claims have been largely supported and extended by data from property norm studies (e.g., McRae, de Sa, & Seidenberg, 1997), which show that concepts in the living domain have many shared properties that are correlated with one another. For example, if a concept has the feature, *has eyes,* it is also likely to have the features *has a mouth, can see, has ears, can move, has fur* (Durrant-Peatfield et al., 1997; McRae et al., 1997; Tyler & Moss, 2001; Tyler et al., 2000). The notion that within-category similarity is higher for living things than for nonliving things has also been captured within other approaches to object representation (Dixon, Bub, & Arguin, 1997; Humphreys & Forde, 2001), including classic exemplar models (Lamberts & Shapiro, 2002), although these accounts do not explicitly model the correlations among features. An important aspect of the conceptual structure model, and the major difference from other similarity and correlation-based models that have implemented some of the same fundamental assumptions (e.g., Devlin, Gonnerman, Andersen, & Seidenberg 1998; Humphreys & Forde, 2001; Lamberts & Shapiro, 2002; McRae et al., 1997), is that we incorporate a set of claims about the interaction between distinctiveness and correlation in the living and nonliving domains. Our central claim is that there are few correlations among the distinctive features of living things; that is, the occurrence of any specific distinctive property for a concept does not strongly predict the occurrence of other properties for the same concept. We have argued that this is at least partly due to the nature of form–function relations for living things (see Tversky & Hemenway, 1984); for living things, variations in form tend not to be functionally significant, at least not as far as the layperson is concerned (e.g., a lion's mane does not inform our knowledge of the lion's behavior; Tyler & Moss, 1997; Tyler et al., 2000). In contrast, nonliving things have distinctive forms, which are consistently associated with the distinctive functions for which they were created (De Renzi & Lucchelli, 1994; Keil, 1986, 1989; Madole, Oakes, & Cohen, 1993).[1] This is not to say that there are no form–function correlations for living things, but rather that they primarily occur among the shared rather than distinctive properties of concepts, often concerning biological functions (*eyes–see; legs–move; lungs–breathe*; Tyler & Moss, 1997). In sum, the conceptual structure account claims that the distinctive features of living things (e.g., *a lion's mane*) are only weakly correlated with other features, whereas the shared features are strongly correlated. For nonliving things, both shared and distinctive features tend to be strongly correlated, although they are relatively fewer in number (Durrant-Peatfield et al., 1997; Greer et al., 2001; McRae et al., 1997; Tyler & Moss, 2001).

Although the conceptual structure account is similar in many respects to several other models that stress the importance of the differing patterns of similarity and correlation across domains, the specific claim about the differential patterns of correlation within the distinctive and shared properties of living and nonliving things seems to us to be a unique property of our model. Other correlation-based models, such as Devlin et al. (1998) and McRae et al. (1997), emphasize that living things have more shared correlated properties than do nonliving things but do not specifically predict a difference in the occurrence of correlations among the distinctive properties of the two domains. Similarity-based accounts, including exemplar models (e.g., Lamberts & Shapiro, 2002), capture the greater overlap among living things but also have little to say about the predicted patterns of correlation among distinctive features.

On the conceptual structure account, these systematic differences in the correlational structure of shared and distinctive properties have important implications for the way in which concepts are impaired by damage to the conceptual system. On the basis of the assumption that strongly correlated properties will be relatively robust to damage due to mutual activation (Devlin et al., 1998), we predict that in most cases, the shared properties of both living and nonliving things will be relatively preserved (perhaps even more so for living things). Conversely, the weakly correlated distinctive properties of living things will be more vulnerable to damage than those of nonliving things, which are supported by form–function correlations. This account not only predicts the prevalence of living things deficits but also captures the finding that such patients typically have well-preserved knowledge of shared information about living things, along with widespread loss of the distinctive properties that would allow them to distinguish one animal or fruit from another (Moss, Tyler, & Devlin, 2002; Moss et al., 1998).

Elsewhere we have reported evidence consistent with the conceptual structure account from a wide variety of studies with patients and from attempts to simulate the effects of damage to the conceptual system in computational models that instantiate our main theoretical claims (Greer et al., 2001; Tyler et al., 2000; Tyler & Moss, 2001). To date, the main focus of the account has been the explanation of neuropsychological data. However, the central question in the present study concerns the implications of the effects of damage for the nature of the normal, intact system. Although many insights into the normal system can be gained from studying the effects of damage, it is also important to carry out direct investigations of conceptual representation and processing in the unimpaired adult population.

The present experiment tests the central prediction of the conceptual structure account—that there will be a consistent disadvantage in processing the distinctive properties of living things relative to other kinds of features, due to their lack of correlation with other information—even in the normal conceptual system. The basis for predicting a similar pattern in the intact as in the impaired system is as follows. As we have already mentioned, correlation supports features in a damaged system through mutual activation. There is also evidence that this mutual activation produces faster processing times for correlated features in the intact system. McRae and colleagues have shown that correlational structure can affect the speed with which the meaning of a concept is

---

[1] We use the term *form* here to refer to all visual perceptual properties of objects, including color, shape, and visible parts. Similarly, we use the term *function* to refer to human-related design and purpose as well as biological functions and behaviors. These uses are broadly consistent with the division into perceptual and functional features that is commonly made in the neuropsychological literature. Although there are undoubtedly other possible conceptual features—for example, nonvisual perceptual features such as sound and smell, and abstract encyclopedic knowledge (e.g., tigers are an endangered species)—these seem to be relatively few in number. Therefore, we have adhered to the neuropsychological convention and focused on visual and functional properties and the patterns of correlation among them.

activated by healthy participants (McRae, Cree, Westmacott, & de Sa, 1999; McRae et al., 1997). In a series of feature verification tasks, participants were asked to verify features as true or false of a concept. Features were either highly correlated with other features of the concept (e.g., *deer–is hunted*) or only weakly correlated with other features (e.g., *duck–is hunted*). Strength of correlation was a significant predictor of reaction time (RT), with more highly correlated features verified more quickly. These results are interpreted as showing that correlated features of concepts are activated and reach a stable state more quickly than weakly correlated features, again as a result of mutual activation from other regularly co-occurring information. Thus, although we would not expect unimpaired adults to make errors on distinctive properties of living things in the kinds of tests that are used with patients (because they will perform at ceiling), we predict a parallel effect in terms of RTs in a speeded task.

In summary, on the assumption that speed of activation increases as a function of correlational strength, the conceptual structure account makes the following prediction: Distinctive properties of living things will be activated more slowly than those of nonliving objects. For shared properties we predict little difference across domains, because both living and nonliving things have groups of correlated shared properties. If anything, the pattern should be in the opposite direction, as living things have a greater number of shared properties. To test this prediction, we use a speeded feature verification task, in which participants indicate as quickly as possible whether features are true of concepts, as our index of speed of activation. We also developed a computational model that instantiated our assumptions about the representational structure of living and nonliving things and carried out a simulation of the speeded verification experiment. The model allowed us to examine the effects of distinctiveness and correlation in isolation from other lexical and semantic variables, which may be influencing the results of the behavioral experiment.

Behavioral Experiments

We used a speeded feature verification task, in which participants decided whether a feature (e.g., *HAS LEGS*) was true of a specific concept (e.g., *ELEPHANT*). The concepts were selected from the domains of living and nonliving things, and the features were either shared (true of many category members, e.g., *HAS LEGS*) or highly distinctive (true of only one or two members of a category, e.g., *HAS A TRUNK*). To maximize the probability of tapping into the early activation of features in the semantic system before it settled into a steady state, we presented the concept names for a brief duration followed by a backward mask. A deadline was also introduced to encourage participants to respond as rapidly as possible, again to reflect early online processing of the concept's features.

The rationale for the speeded feature task is that when participants read the concept name (e.g., *ELEPHANT*), its various semantic features (*HAS LEGS, IS GREY, IS BIG, HAS A TRUNK, EATS GRASS,* etc.) begin to be activated, until a stable state, corresponding to the *ELEPHANT* concept, is reached. Following McRae et al. (1997), we assume that some features are activated more quickly than others and that this rate of activation is affected by correlational structure. Importantly, McRae et al. demonstrated that correlational structure determined participants' performance in a speeded feature verification task but not in an untimed sa-

liency rating task, suggesting that correlation affects initial activation rate but not necessarily the final level of activation of features once the stable state has been achieved. This pattern was confirmed in a computational simulation using distributed feature representations for concepts: The effect of feature correlation peaked early in the activation of a concept and then tailed off. On these grounds, the earliness with which we could tap into semantic activation was a major aim in our verification paradigm.

As discussed in the introduction, the conceptual structure account claims that weakly correlated features will be activated most slowly, and therefore predicts that the distinctive features of living things will be verified with significantly longer RTs and higher error rates in the speeded feature verification task compared with the other three conditions (living shared, nonliving shared, and distinctive). Additionally, we can contrast the pattern of results predicted by the conceptual structure account against the predictions of an alternative account of conceptual knowledge outlined earlier. Although the sensory–functional account does not make any specific claims about the speed of activation of different feature types, it would be plausible to expect that the sensory features of living things and the functional features of nonliving things would be verified most quickly, because these are held to be the most salient or distinguishing features for each domain. Some evidence consistent with this prediction was found in a priming study in which functional features of nonliving things were facilitated at an earlier point in the duration of a spoken concept name than were sensory features (Moss, McCormick, & Tyler, 1997). However, living things were not tested in this study, so we cannot tell whether they would show the reverse pattern, with earlier priming for sensory than functional features. We therefore tested for sensory–functional by domain interactions in the present verification experiment by coding each feature, not only as distinctive or shared but also as either a functional or sensory property.

The prediction of the conceptual structure account that distinctive properties of living things will be activated more slowly in the normal system is predicated on the central assumption that these features are weakly correlated relative to the shared properties of living things and both the distinctive and shared properties of nonliving objects. To confirm this basic claim and to select appropriate materials for the verification experiment, our first step was to analyze data from a property norm study, in which a large number of participants listed as many features as they could for concepts in the living and nonliving domains. Distributional statistics were then computed over the resulting feature lists to determine the patterns of distinctiveness and correlation in the two domains (aspects of these norms have been described in part elsewhere; Greer et al., 2001; Moss et al., 2002).

Property Norm Study

Method

Forty-five participants were asked to list all the features they could think of for a set of 93 concepts from the living and nonliving domains, including animals, fruits, vehicles, and tools. These categories were chosen because they are standardly used as test cases in neuropsychological studies of category-specific deficits. We included 31 animals, 16 fruits, 22 tools, and 24 vehicles reflecting the relative numbers of well-known items in each category. Concept familiarity (Coltheart, 1981) was matched across the living and nonliving domains (Ms = 5.1 and 5.16, respectively). An additional 47 filler items from a range of other categories (e.g., birds,

musical instruments, clothing) were included in the test list to add variety to the task.

Participants' responses were then compiled to produce a feature list for each concept, excluding idiosyncratic properties generated by fewer than 5 participants The data were then cleaned according to criteria based on those used by McRae et al. (1997). First, quantifiers (e.g., *generally, usually*) were removed. Second, certain key words were used to code the properties; for example, *is a tool* would indicate a category label; *is brown* would indicate an adjective. Third, adjective–noun properties, such as *has four wheels,* were divided into *has wheels* and *has 4 wheels* on the assumption that the participant has given two pieces of information. Fourth, disjunctive properties (e.g., *is green and red*) were also divided up into *is green* and *is red.* Fifth, verb–noun properties, such as *eats grass,* were divided into *eats* and *eats grass.* The resulting feature lists were used to calculate the following statistics.

1. The distinctiveness of each feature, in which distinctiveness is defined as the inverse of the number of concepts in which the property is listed. For example, a unique feature (such as *HAS A TRUNK*) that is present in only one concept has a high distinctiveness score of 1, whereas a shared feature (such as *HAS LEGS*) that is present in 29 of the 93 concepts has a low distinctiveness of 1/29. This enabled us to compute the mean distinctiveness of properties of concepts in each domain. As predicted, mean distinctiveness was higher for nonliving things (0.45) than for living things (0.33), $t(1055) = 5.4$, $p < .001$.[2]

2. The number of correlated property pairs in each domain. We defined a correlated property pair (CPP) as any pair with a Pearson product–moment correlation having a significance of $p < .05$. As predicted, there were more correlations overall for living things than nonliving things (60 vs. 25), $t(91) = 7.2$, $p < .001$, but crucially, the proportion of each concept's CPPs that were between two distinctive properties (i.e., properties occurring for two or fewer concepts) was significantly smaller for living things than for nonliving things (16% vs. 30%), $t(91) = 3.3$, $p < .001$.

3. The correlation strength for features within a concept. This was computed for each feature–concept pair as the mean correlation between that feature and all other features in the concept. As predicted, the correlation strength among distinctive properties (those occurring for two or fewer concepts) was significantly greater for nonliving things than living things (.58 vs. .50, respectively), $t(369) = 6.7$, $p < .001$. This was not the case for the shared properties, in which there was a greater correlation strength for living things (.35) than nonliving things (.32), $t(684) = 3.6$, $p < .001$.[3]

Summary

The property norm analyses show that although there are many correlations among properties in the living domain, only a relatively small proportion of these occur among distinctive features. Moreover, the correlational strength of distinctive features of living concepts is significantly lower than that of nonliving things. This provides support for our starting assumption that distinctive properties of living things are weakly correlated.

Speeded Feature Verification

Method

Participants

Twenty-six members of the Centre for Speech and Language (Cambridge, United Kingdom) participant pool took part in this experiment for which they received payment of £5 (approximately U.S. $8.80). All were native speakers of British English, were between the ages of 18 and 40 years, and had normal or corrected-to-normal eyesight.

Design and Materials

The experimental stimuli consisted of 80 written word pairs, which were presented in four blocks. In each pair, the first word denoted the concept and the second word described a distinctive or shared feature (e.g., *WOLF–FIERCE*). The concepts were selected from two categories in the domain of living things (animals and fruits) and two categories in the domain of nonliving things (tools and vehicles). Features were selected on the basis of the property generation study described earlier. Distinctive features were defined as those that were listed for no more than two concepts, whereas shared features were listed for three or more. We selected only those features that could be unambiguously expressed by a single content word. This was to encourage fast, consistent reading times across trials.

Concept words were matched as closely as possible across the four conditions (living/nonliving × shared/distinctive) for word length, lemma per million frequency (Baayen, Piepenbrock, & Van Rijn, 1995), familiarity and imageability (Medical Research Council Psycholinguistic database; Coltheart, 1981), and number of senses (Parks, Ray, & Bland, 1998; Rodd, Gaskell, & Marslen-Wilson, 2002), as were feature words. It was also important to control for a number of additional variables concerning the relation between the concept and feature. First, we matched for production frequency, that is, the number of participants who listed that feature for the concept in the property norming study. Second, we matched for the rated semantic relatedness of the concept–feature pairs in each condition. Ratings were collected in a pretest in which 15 participants indicated on a 9-point scale how related in meaning they considered each word pair, with $1 = $ *not related at all* and $9 = $ *very related.* We matched the free association strength of the concept and feature words and chose weakly associated pairs as far as possible to avoid a major contribution of priming on the basis of lexical co-occurrence (Moss & Older, 1996; Nelson, McEvoy, & Schreiber, 1998). Finally, we calculated the mean correlation strength for the distinctive and shared properties of living and nonliving things to confirm that we had the same pattern of correlation in this set, compared with the full property norm set.

Table 1 lists the mean values for each of these variables across the four conditions. Although we matched as closely as we could, it can be seen from Table 1 that some differences remained. It was not possible to match on all variables because of the limited set of potential items. Also, some variability is simply inherent to particular feature types. For example, distinctive features, because they apply to fewer objects in the real world, tend to be of lower frequency than shared features. To address this potential

---

[2] Although significant, the true difference in distinctiveness across domains may be even greater than this, because of the underrepresentation of shared information in the property generation paradigm. Many researchers have pointed out that participants are unlikely to repeatedly list all the common properties of category members (e.g., *breathes, can move*), focusing instead on the properties that distinguish among them (e.g., Murphy & Medin, 1985). Thus a high proportion of shared information—for living things especially—may not be revealed in this paradigm. We have explored various ways of compensating for any underrepresentation of shared information, such as adding properties generated directly in response to superordinate category names (e.g., we asked participants to list all the properties of animals or vehicles), and have found that the main patterns of results were little changed. In general, any differences are in the direction predicted by our model (e.g., more shared, correlated information for living things). Therefore, for the sake of simplicity, we report the data from the basic unsupplemented property lists here.

[3] Note that the correlation strength is higher overall for distinctive than shared properties in this set of results. This is because of the very high weighting that is given to correlations among the most distinctive properties that occur for one concept only. If two such properties occur for one concept and no others, then their correlation is 1. However, the important point here is the domain difference in correlation strength *within* the distinctive set rather than the shared–distinctive comparison.

Table 1
Descriptive Statistics for the Stimuli

Condition	N	Letters	Frequency[a]	Familiarity[b]	Imagability[b]
Concept variables					
Living					
Distinctive	13	5.0	23	522	613
Shared	29	5.5	18	502	624
Nonliving					
Distinctive	17	6.2	16	509	610
Shared	20	5.8	27	503	590
Feature variables					
Living					
Distinctive	13	5.5	23	533	521
Shared	29	5.0	107	548	483
Nonliving					
Distinctive	17	4.8	76	507	541
Shared	20	4.9	147	581	448

Condition	N	Correlation strength	Semantic relatedness[c]	Production frequency	No. of senses[d]	Association strength[e]
Concept–feature variables						
Living						
Distinctive	13	.55	5.98	9	4.1	6.6
Shared	29	.30	5.35	13	4.2	1.28
Nonliving						
Distinctive	17	.42	7.04	22	3.9	8.19
Shared	20	.35	5.60	16	5.4	3.72

Note. [a] Celex Lexical Database (Baayen et al., 1995). [b] Medical Research Council psycholinguistic database (Coltheart, 1981) and Centre for Speech and Language (CSL) laboratory pretests. [c] Pretest conducted in CSL laboratory. [d] Wordsmyth English Dictionary–Thesaurus (Parks et al., 1998). [e] Birkbeck association norms (Moss & Older, 1996) and University of South Florida free association norms (Nelson et al., 1998).

problem, we entered any unmatched variables that correlated with RTs or error rates as covariates in the data analysis.

To ensure an equal proportion of true and false trials throughout the experimental list, we constructed a further 80 concept–feature pairs as foils. These were pairs in which the feature was *not* generally true of the concept (e.g., *APPLE–SHY*). We included the same number of living and nonliving concepts in the foil items as in the experimental items and did not repeat concepts within a presentation block. Approximately 75% of the false features were taken from concepts within the same domain (e.g., *GIRAFFE–GROWL*—a giraffe does not growl, but the feature is true of other members of the animal category, such as bear and lion), and 25% of features were taken from concepts in the opposite domain (e.g., *ELEPHANT–WHEELS*). The high proportion of within-domain foils was intended to avoid the possibility that the verification task would be so easy that participants would lose interest in responding quickly. There was the same number of repeated foil features within a block as there were repeated test features. This was to ensure that the repetition of a feature was not predictive of a "yes" response.

We divided the test and foil pairs into four experimental blocks; each block contained 20 experimental items requiring a "yes" response and 20 foil items requiring a "no" response. In each block, the type of relationship between concept and feature was held constant. For example, in one block the feature always referred to a part of the concept's referent, as in *HORSE–(HAS A) TAIL* and *CAR–(HAS) WHEELS*. At the beginning of each block, participants were informed as to the type of relation to expect. In this way we were able to reduce the feature to a single word, thereby facilitating the very fast presentation that we considered necessary to tap into the early stages of activation of the semantic system. If different types of relation were presented in random order, a single word feature could be

ambiguous (e.g., *KNIFE–HAND* could mean that a knife is held in the hand, which is true, or that a knife has a hand, which is false).

For two blocks the feature was always a sensory characteristic of the concept (e.g., *LEMON* (IS) *JUICY*). In the third block the feature was again a sensory one, but this time referring to a visible part of the concept (e.g., *HORSE* (HAS A) *TAIL*). For the fourth block the feature word referred to the typical use or function of the concept's referent (e.g., *PENCIL* (IS USED TO) *WRITE*). The same number of living and nonliving things items required a "yes" and "no" response in each block. The following are examples of each type of block:

1. Living: CHERRY round Non-living: NEEDLE thin Relation: IS

2. Living: MONKEY legs Non-living: KNIFE blade Relation: HAS

3. Living: DONKEY ride Non-living: PEN write Relation: USED TO

We used a 2 × 2 independent measures design, in which the independent variables were domain (living and nonliving) and distinctiveness (distinctive and shared). The dependent variables were RT and error rate.

Procedure

The task was speeded feature verification, in which participants were presented with a written word denoting a concept followed by a word denoting a feature. The experiment began with written instructions describing the kind of concept–feature relation to be verified in the block of items

Table 2
Speeded Feature Verification Experiment: Mean Retransformed Reaction Times and Error Rates for Distinctive and Shared Features of Living and Nonliving Things

	Living		Nonliving	
Measure	Distinctive (*n* = 13)	Shared (*n* = 29)	Distinctive (*n* = 17)	Shared (*n* = 20)
Reaction time (ms)				
M	547 (15)	500 (10)	475 (12)	500 (13)
Estimated marginal *M*	553 (14)	484 (9)	506 (13)	494 (11)
Error rate (%)				
M	33 (4.6)	22 (3.0)	17 (4.0)	20 (3.8)
Estimated marginal *M*	36 (4.5)	18 (3.0)	19 (4.3)	20 (3.6)

Note. Standard errors are in parentheses.

about to be presented. Participants were also shown examples of the type of item to be presented in the block. They then pressed the space bar to begin.

On each trial, participants saw a fixation point in the center of a computer screen for 500 ms. This was followed by a concept word, presented in uppercase letters for 60 ms, followed by a backward mask (a series of #) for 150 ms to reduce the potential for continued semantic processing of the concept after the word was removed from view. After an interval of 100 ms, the feature word was presented in lowercase letters for 100 ms. After a further 450 ms, a tone sounded over headphones. Participants were asked to try to "beat the tone" by pressing the "yes" response button if the words were related and the "no" button if they were not. If they gave the correct answer they were given feedback that included their RT; if they gave an incorrect answer they were given the message "wrong." There was an intertrial interval of 1,000 ms.

The 4 blocks of items were randomly interleaved with 13 other blocks from another experiment. The experiment began with an additional practice block, and each block began with three lead-in items. The duration of the experiment was approximately 1 hr.

Results

Analyses were conducted on 79 items, as 1 item was lost due to experimenter error. Before analysis, the raw response times for all correct verification trials were inverse transformed to minimize the effects of outlying data points without introducing truncation biases (Ulrich & Miller, 1994). Error analyses were conducted on the proportions of errors (false negatives) calculated over items and participants in each condition.[4]

Reaction Times

Table 2 shows the retransformed (harmonic) mean latencies for correct "yes" responses in each condition. Table 2 also gives the estimated marginal means to allow for the influence of significant covariates (as discussed below).

Analyses were conducted on means in two ways: over all of the items in a condition for each participant to give the participants' means (F_1) and over all of the participants for each item to give the item means (F_2). A 2 (domain) $\times$ 2 (distinctiveness) repeated measures analysis of variance (ANOVA) on the participant (F_1) means and a 2 (domain) $\times$ 2 (distinctiveness) independent measures ANOVA on the item (F_2) means for correct responses only showed that RTs to living things ($M = 513$ ms, $SE = 8.90$) were slower overall than to nonliving things ($M = 488$ ms, $SE = 9.00$), $F_1(1, 25) = 24.01$, $p < .001$; $F_2(1, 75) = 5.18$, $p < .05$. There was

no main effect of distinctiveness (F_1 and $F_2 < 1$), with RTs to distinctive pairs averaging 504 ms ($SE = 9.9$) compared with 500 ms for shared pairs ($SE = 7.9$). However, distinctiveness interacted significantly with domain, $F_1(1, 25) = 50.79$, $p < .001$; $F_2(1, 75) = 8.25$, $p = .005$. To explore this interaction, we conducted further ANOVAs. First, we conducted an analysis on the living and nonliving items separately. For living things, the 47-ms advantage for shared features (500 ms, $SE = 9.94$) over distinctive features (547 ms, $SE = 14.90$) was significant, $F_1(1, 25) = 25.85$, $p < .001$; $F_2(1, 41) = 6.76$, $p < .025$. The 25-ms advantage for distinctive features (475 ms, $SE = 11.65$) over shared features (500 ms, $SE = 13.06$) for nonliving things was significant only in the participant analysis, $F_1(1, 25) = 13.94$, $p < .001$; $F_2(1, 34) = 2.22$, $p > .1$. We then carried out ANOVAs on the shared and distinctive properties separately. For the distinctive properties, RTs to nonliving things were significantly faster than to living things, $F_1(1, 25) = 61.01$, $p < .001$; $F_2(1, 28) = 13.93$, $p = .001$, whereas for the shared properties, RTs to living and nonliving things were identical.

Finally, using the Newman–Keuls statistic we compared means across all four conditions (living distinctive, living shared, nonliving distinctive, and nonliving shared). This analysis showed that there was no difference among the RTs for living shared, nonliving distinctive, and nonliving shared. Only the living distinctive items differed from the other three conditions.

Because it was not possible to fully match the materials across conditions on all the relevant variables, we also carried out a correlational analysis to determine whether any potentially confounding stimulus variables were significant predictors of RT. We entered the mean RT for each concept–feature pair and all the descriptive variables shown in Table 1 into a series of bivariate correlations. The results showed that three variables correlated significantly with RT: production frequency, $r(79) = .32$, $p = .005$; rated semantic relatedness between concept and feature, $r(79) = .40$, $p = .001$; and association strength between concept and feature words, $r(79) = .27$, $p < .025$. These variables were entered into a further analysis of covariance as covariates. In this

---

[4] In the analyses reported here, we include the data from all participants and all items. We also carried out additional sets of analyses in which the data from very slow participants and items that attracted high error rates were removed. The pattern of results was the same as in the analyses reported here.

analysis, there were two independent measures (domain: living/nonliving and distinctiveness: shared/distinctive) and three covariates (production frequency, semantic relatedness of concept and feature, and association strength of concept and feature). We found no main effect of domain ($F_2 < 1$), although distinctiveness was significant, $F_2(1, 72) = 9.89$, $p < .01$. The crucial interaction between domain and distinctiveness remained significant, $F_2(1, 72) = 5.69$, $p < .025$. Separate analyses showed that for living things, distinctive properties (553 ms) were responded to more slowly than shared properties (484 ms), $F_2(1, 38) = 20.17$, $p < .001$. For nonliving things there was no significant difference in RTs to distinctive (506 ms) and shared properties (494 ms), $F_2 < 1$. For distinctive properties only, RTs for living things were significantly slower than for nonliving things, $F_2(1, 25) = 4.29$, $p < .05$. For shared properties, there was no difference between living and nonliving concepts ($F_2 < 1$).

Error Rates

Errors are informative about the activation of semantic information, on the assumption that participants made errors in cases in which they attempted to respond before the deadline because of insufficient correct information, rather than reflecting deliberate incorrect decisions on difficult items. This assumption is supported by the finding that RTs on error trials (false negatives) were just as fast as those on correct trials ($M = 523$ ms, $SE = 10.2$ and $M = 542$ ms, $SE = 5.3$, respectively).

A 2 (domain) $\times$ 2 (distinctiveness) repeated measures ANOVA on the participant means and a 2 (domain) $\times$ 2 (distinctiveness) independent measures ANOVA on the item means showed a main effect of domain, $F_1(1, 25) = 18.33$, $p < .001$; $F_2(1, 75) = 3.38$, $p < .05$, with more errors for living things ($M = 25.0\%$, $SE = 2.6$) than for nonliving things ($M = 18.2\%$, $SE = 2.8$). Distinctiveness was significant only on the participant analysis (M distinctive $= 23.6\%$, $SE = 3.8$; M shared $= 20.9\%$, $SE = 2.0$), $F_1(1, 25) = 4.54$, $p < .05$; $F_2(1, 75) = 1.10$, $p > .2$. However, just as in the RT analysis, there was a Domain $\times$ Distinctiveness interaction that was highly significant in the participant analysis and marginally significant in the items analysis, $F_1(1, 25) = 19.30$, $p < .001$; $F_2(1, 75) = 3.38$, $p = .07$. A separate analysis showed that for living things, the distinctive properties incurred significantly more errors than the shared (33% vs. 22%), $F_1(1, 25) = 14.60$, $p = .001$; $F_2(1, 41) = 4.31$, $p < .05$. There was no difference between the distinctive (17%) and shared (20%) properties for the nonliving things, $F_1(1, 25) = 2.55$, $p > .1$; $F_2 < 1$. For the distinctive properties, there were significantly more errors for living things than for nonliving things, $F_1(1, 25) = 25.89$, $p < .001$; $F_2(1, 28) = 5.14$, $p < .05$. For shared properties, there was no significant difference between living things and nonliving things (both $Fs < 1$). A Newman–Keuls test showed that only the distinctive features of living things produced more errors than any other condition.

Once again, because the conditions were not completely matched on all the relevant variables, we looked to see if any variable correlated with error rate using the item means. We conducted a series of bivariate correlations to see if any of the descriptive variables shown in Table 1 were related to error rate. Two variables correlated significantly with error rate: production frequency, $r(79) = -.25$, $p = .05$, and semantic relatedness, $r(79) = -.35$, $p = .0025$. We repeated the items ANOVA but included production frequency and semantic relatedness as covari-

ates. The adjusted means, following this analysis, showed the same pattern as in the original analysis. We found no main effect of domain ($F_2 < 1$) but a significant effect of distinctiveness, $F_2(1, 73) = 6.59$, $p < .025$. With covariates added, the interaction between distinctiveness and domain was not significant, $F_2(1, 73) = 1.29$, $p > .2$, However, there were many more errors for the living things distinctive features ($M = 32.8\%$, $SE = 4.5$) than for the shared features ($M = 18.3\%$, $SE = 3.0$), and this difference was statistically reliable, $F_2(1, 39) = 5.97$, $p < .025$.

The results of both the RT and error analyses show that participants had most difficulty with the distinctive properties of living things. In all the analyses there was a significant difference between living and nonliving things for the distinctive feature conditions but little or no difference for the shared features. This pattern of results is precisely as predicted by the conceptual structure account (Tyler & Moss, 2001; Tyler et al., 2000).

Regression Analyses

In addition to the factorial ANOVA analyses, we also examined the data using multiple regression techniques, for two reasons. First, because there were different numbers of items in each of the different conditions in this experiment, the power of the factorial design may have been reduced. Second, we classified distinctive features as those that were generated for one or two concepts in the property norm study, and shared features as those that were generated for three or more concepts. However, in reality, distinctiveness is a continuous rather than binary variable, which may be better modeled in a regression analysis. We entered four factors into the regression analysis: semantic relatedness, production frequency, association strength (as we know these variables correlate with RT), and distinctiveness. In a simultaneous multiple regression, we found that distinctiveness was a significant predictor of RT for the total data set (unstandardized beta coefficient $= -0.20$, $SE = 0.07$), $t(78) = -3.00$, $p < .005$. When we examined the two domains separately, the effect for living and nonliving things differed. For living things only, the regression analysis showed that distinctiveness remained a significant predictor (unstandardized beta coefficient $= -0.5$, $SE = 0.08$), $t(42) = -4.13$, $p < .001$, whereas it was not significant for the nonliving things (unstandardized beta coefficient $= -0.04$, $SE = 0.11$), $t(35) = -0.37$, $p > .7$. We then carried out a parallel regression analysis, with error rates rather than RTs as the dependent variable, and found the same pattern. Overall, distinctiveness was a significant predictor (unstandardized beta coefficient $= 0.18$, $SE = 0.06$), $t(78) = 3.31$, $p = .001$. When living and nonliving things were examined separately, distinctiveness was only a significant regressor for the living things: living things, unstandardized beta coefficient $= 0.22$, $SE = 0.07$, $t(42) = 3.11$, $p = .005$; nonliving things, unstandardized beta coefficient $= 0.15$, $SE = 0.09$, $t(35) = 1.63$, $p > .1$.

The results of these regression analyses confirm the results of the ANOVAs and support the predictions of the conceptual structure account. For living things, the more distinctive a feature is, the slower the RT in this speeded verification task. There was no such effect of distinctiveness for nonliving things.

Sensory–Functional Account

When designing the materials for this experiment, we coded items for the sensory or functional nature of the feature. This

enabled us to analyze the data to determine whether the functional and sensory features of living or nonliving things produced different patterns of performance, as would be predicted by the sensory–functional account.

RTs. The mean RTs are shown in Table 3 and were explored in an ANOVA with two independent measures (domain: living and nonliving; type of feature: sensory or functional). RTs to nonliving things ($M = 488$ ms, $SE = 8.97$) were faster than to living things ($M = 513$ ms, $SE = 8.81$), $F_1(1, 25) = 16.16$, $p < .001$; $F_2(1, 75) = 4.36$, $p < .05$, and faster to functional features ($M = 494$ ms, $SE = 9.54$) than to perceptual features ($M = 510$ ms, $SE = 8.61$). This difference was significant only in the participant analysis, $F_1(1, 25) = 7.63$, $p < .025$; $F_2(1, 75) = 1.72$, $p > .1$. There was no interaction between type of feature and domain, $F_1(1, 25) = 1.49$, $p > .2$; $F_2 < 1$.

When we repeated the analysis on the item means, including variables that covary with RT (semantic relatedness, production frequency, and association strength), the effects of domain and type of feature were lost (domain: $F_2 < 1$; type of feature: $F_2 < 1$). Most importantly, there was still no interaction between domain and type of feature ($F_2 < 1$).

Error rates. Participants made more errors for living things ($M = 25.0\%$, $SE = 2.6$) than nonliving things ($M = 18.2\%$, $SE = 2.8$). This difference was significant by participants and marginally significant by items, $F_1(1, 25) = 12.89$, $p = .001$; $F_2(1, 75) = 3.23$, $p < .08$. However, there was no difference in number of errors to perceptual ($M = 21.8\%$, $SE = 2.5$) or functional features ($M = 22.0\%$, $SE = 2.9$; both $Fs < 1$) and no interaction between type of feature and domain (both $Fs < 1$). When semantic relatedness and production frequency were added as covariates in the analysis, there were no significant effects at all: domain, $F_2 < 1$; type of feature, $F_2(1, 73) = 1.38$, $p > .2$; Domain $\times$ Type of Feature interaction, $F_2(1, 73) = 1.04$, $p > .3$. Thus, both RT and error analyses fail to provide any support for the sensory–functional hypothesis. In contrast, the results of this speeded feature verification study generate a pattern of RTs and errors that are predicted by the conceptual structure account.

Unspeeded Feature Verification

The results of the speeded feature verification task appear to support the predictions of the conceptual structure account, but it is also possible that the pattern of RTs and error rates reflects differences in how well participants perceive the feature to match the concept. It is possible that we selected a set of distinctive features for the living things that are not as well fitted to their concepts as features from the other conditions, and hence are simply more difficult to relate to each other. If this were the case, we should find that these properties are less reliably judged to be true of their concepts than those in the other conditions, even in a task in which speed of activation is not a factor. To test whether this was the case, we carried out a further study in which participants were asked to judge whether each of the feature–concept pairings in the experiment was true or false.

We constructed a booklet in which each of the experimental items and the foils from the speeded feature verification task were listed in the same blocks, and with the same relationship between concept and feature as in the original experiment. Fifteen participants were asked to judge each concept–feature pair as true or false. The results are shown in Table 4. An ANOVA showed that there are no statistical differences in these judgments: domain, $F < 1$; distinctiveness, $F(1, 75) = 3.11$, $p = .08$; Domain $\times$ Distinctiveness: $F < 1$. A Newman–Keuls post hoc test revealed that there are no differences between any of the separate conditions in this rating task. The results of this task show that the results we obtained in the speeded feature verification task cannot be attributed to differences in the difficulty of judging how well the feature matches the concept. Although there was a marginal effect of distinctiveness, such that error rates were higher for distinctive features than shared features, there was no evidence of a difference between the distinctive properties of living and nonliving things.

Computational Model

According to our conceptual structure account, the differences between living and nonliving concepts—seen both in patients and in the performance of healthy participants in the present study—reflect systematic differences in correlational structure of concepts across domains. Although we have focused on distinctiveness and correlation throughout the development of the model, we would not, of course, suggest that these are the only factors that determine the processing and breakdown of conceptual representations; concept familiarity, autobiographical relevance, property salience and cue validity, knowledge of causal relations, and many other factors may also play a part. Our strategy has been to start with just the two factors of distinctiveness and correlation to see how far these can account for the data before adding further complexity.

Table 3

Speeded Feature Verification Experiment: Mean Retransformed Reaction Times and Error Rates for the Functional and Perceptual Features of Living and Nonliving Things

Measure	Living		Nonliving	
	Functional ($n = 21$)	Perceptual ($n = 22$)	Functional ($n = 18$)	Perceptual ($n = 18$)
Reaction time (ms)				
M	510 (14)	517 (10)	476 (13)	501 (13)
Estimated marginal M	504 (14)	517 (10)	496 (14)	499 (12)
Error rate (%)				
M	25 (4.3)	25 (3.2)	19 (4.1)	18 (3.9)
Estimated marginal M	23 (4.4)	22 (3.3)	26 (4.3)	17 (3.8)

Note. Standard errors are in parentheses.

With this in mind, a potential counterargument to our account is that the results of the present study are due to one or more of these additional factors rather than to correlational structure. Although we partialed out several important lexical and semantic factors in our covariate and regression analyses, we cannot claim to have ruled out all possible relevant variables. The aim of the computational model is to address this theoretical possibility by testing the effect of correlational structure in a more tightly controlled, albeit artificial context, by using a computational simulation. In the simulation, we can ensure that the network does not "know" anything else about the concepts other than their internal structure—that is, the trained pattern of activation over feature vectors. Information such as familiarity, salience, lexical association, and so on is not available to the network.

The computational simulation reported here uses the set of concepts and properties generated in the property generation study described earlier in this article and allows us to test whether an interaction between domain and distinctiveness can emerge from differences in the correlational structure of living and nonliving things, in the absence of any other information about the concepts or properties. Although this would not entail that correlational structure also plays a part in the human system, it would demonstrate the plausibility of an account in which correlation is a key explanatory variable.

Method

Network Architecture

A feed-forward network was trained on the mapping between orthography and semantics. The model's architecture is presented in Figure 1. The network mapped from 200 orthographic to 396 semantic feature units, via 50 hidden units.

Training Patterns

Word form units. The orthography of each of the 93 concepts was represented by turning on 50 of the 200 orthographic units. Each word had a unique orthographic pattern that was generated pseudo-randomly, with the constraint that all concepts differed by at least 6 units. Therefore, these input patterns capture two important characteristics of the mapping from form to meaning. First, it is largely arbitrary. Second, because each word form unit has a one in four chance of being activated for each concept, the patterns are highly overlapping.

Semantic feature units. The semantic representations for the 93 concepts were taken directly from the property norms described earlier. There were exemplars from four categories (31 animals, 16 fruits, 22 tools, and 24 vehicles). Each concept retained all of the features that had been listed by at least 5 of 45 participants during property norming, resulting in a total of 396 semantic features. Each of the 396 semantic units in the semantic feature layer corresponds to one of these features. On average, concepts
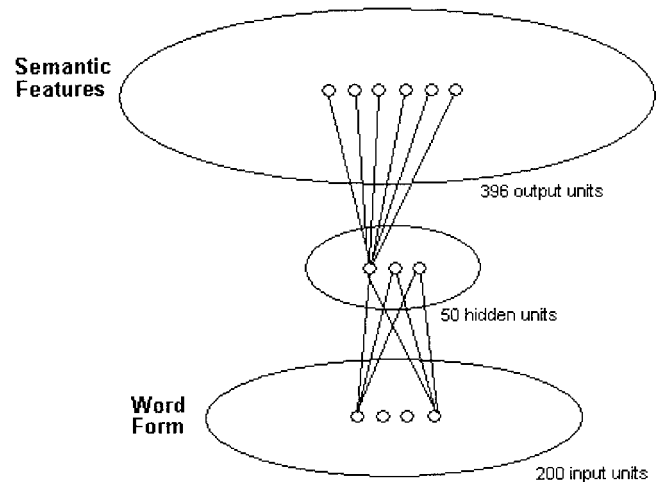

Figure 1. Architecture of the feed-forward network. Not all connections are shown. Where connections are shown, full connectivity was used.

activated just 11.3 of the 396 features, making these representations relatively sparse. As for the speeded feature verification analysis, each semantic feature was classified as distinctive if it was present in only one or two concepts, and as shared if it was present in at least three concepts. The living concepts had 196 distinctive features and 451 shared features; the nonliving things had 175 distinctive features and 235 shared features.

Training

During each learning trial, the network was presented with the orthographic and semantic representations that correspond to one of the 93 concepts, and the strength of the connections between the units was modified according to the back-propagation learning algorithm (Rumelhart, Hinton, & Williams, 1986; learning rate = 0.25, momentum = 0.5). The values of these parameters were chosen to optimize the network's performance across the whole set of features; that is, they were all set such that the network was able to learn the whole training set as quickly as possible. The network was trained in a series of epochs in which all 93 concepts were presented once in random order. The network continued to train until the activation of all features for every concept was within a threshold of 0.1 (i.e., at least 0.9 if it should be activated and less than 0.1 if it should be off).[5] One hundred individual networks with different small initial random weights (± 0.05) were trained using this procedure.[6] On average, these networks took 221 epochs to reach threshold.

Testing: Error Scores

For each of the 93 concepts, the orthographic input pattern was presented, and the resulting output at the semantic units was measured. For every concept, the error score for each feature was calculated as the difference between the target activation (0 or 1) and its actual activation. In the following analyses we focus on the network's performance on the features that should be turned on for a particular concept (e.g., *LION–CAN ROAR*), because these correspond to the "yes" responses in the feature verification experiment. In total there are 1,057 concept–feature pairs of this type. In all cases (because of the threshold used during training), the output for these features was above 0.9, and therefore the errors are very

Table 4
Unspeeded Feature Verification Experiment: Mean True Ratings

	Living		Nonliving	
	Distinctive ($n = 13$)	Shared ($n = 29$)	Distinctive ($n = 17$)	Shared ($n = 20$)
% true ratings				
M	87 (2.5)	94 (1.9)	88 (6.7)	94 (2.5)

Note. Standard errors are in parentheses.

_____

[5] This simulation has been repeated with different values for the activation threshold; the pattern of results is stable across these changes.

[6] Three additional networks that did not reach the threshold performance after 5,000 epochs were excluded from the analysis.

small. Following Seidenberg and McClelland (1989) we interpret these error scores as an analogue of both RT and accuracy in the behavioral experiment; items that have lower error scores are those for which we would expect faster and more accurate performance under time pressure.

The error scores for the distinctive and shared features of the living and nonliving concepts are shown in Figure 2. The mean error rates for the different conditions were averaged across features and networks, and these mean values were subjected to ANOVAs that mirror those reported for the feature verification experiment. (For these analyses, F_1 represents conditions averaged across the different networks and F_2 represents conditions averaged across the different items in the training set.) These analyses show that the error scores were higher for living things ($M = 1.25\%$, $SE = 0.00$) than for nonliving things ($M = 1.12\%$, $SE = 0.00$), $F_1(1, 99) = 14.2$, $p < .001$; $F_2(1, 1053) = 21.4$, $p < .001$. There was also a main effect of distinctiveness, with higher error scores for distinctive features ($M = 1.42\%$, $SE = 0.00$) than for shared features ($M = 0.96\%$, $SE = 0.00$), $F_1(1, 99) = 444.0$, $p < .001$; $F_2(1, 1053) = 252.0$, $p < .001$. This main effect was not present in the speeded feature verification experiment, as we discuss below. Most importantly, there was an interaction between distinctiveness and domain, $F_1(1, 99) = 156.4$, $p < .001$; $F_2(1, 1053) = 21.9$, $p < .001$. The form of the interaction was that there was a significant effect of domain for the distinctive features ($M = 1.55\%$, $SE = 0.001$ for living concepts; $M = 1.28\%$, $SE = 0.001$ for nonliving concepts), $t_1(99) = 6.3$, $p < .001$; $t_2(369) = 9.0$, $p < .001$, but not for the shared features ($M = 0.96\%$, $SE = 0.00$ for both living and nonliving concepts), $t_1(99) = 0.06$, $p > .9$; $t_2(684) = 0.05$, $p = .97$. This is the same pattern as was seen in the speeded feature verification experiment.

Testing: Cascaded Activation

One possible difficulty in relating these results to the results of the feature verification experiment is that we are using an error score as the measure of the network's performance. Although it is a well-established method to use these error scores to simulate participants' performance on a speeded task, we decided to evaluate the network's performance in a second way that arguably provides a more transparent analysis of the time to activate features in the human system, using a technique proposed by Cohen, Dunbar, and McClelland (1990). This method allows us to look at the time course of semantic activation by cascading activation through the network. Information from lower levels gradually propagates upward, allowing the activation of the semantic units to develop over time. This gives us a more direct measure of how long it would take the network to produce the appropriate output for different types of features.

The results of this cascaded approach replicate the interaction between domain and distinctiveness reported earlier. These effects were present
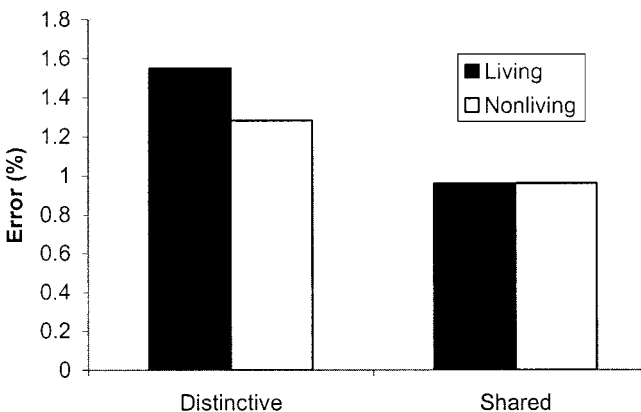

Figure 2. Network error scores (%) for the distinctive and shared features of living and nonliving things.

throughout the settling of the network. We analyzed in detail the number of update cycles taken by the network to reach an activation level of 0.7.[7] These analyses showed that although the living concepts settled more slowly ($M = 66.0$ cycles, $SE = 1.09$) than the nonliving concepts ($M = 65.4$ cycles, $SE = 1.13$), this difference was not significant (both $Fs < 1$). There was a main effect of distinctiveness, with longer settling times for distinctive features ($M = 70.7$ cycles, $SE = 1.12$) compared with shared features ($M = 60.7$ cycles, $SE = 0.74$), $F_1(1, 99) = 377$, $p < .001$; $F_2(1, 1053) = 160$, $p < .001$. Most importantly, there was an interaction between distinctiveness and domain, $F_1(1, 99) = 10.0$, $p < .01$; $F_2(1, 1053) = 21.9$, $p < .001$, such that for the distinctive features there was a significant disadvantage for the living things ($M = 72.3$ cycles, $SE = 1.20$ for living concepts; $M = 69.2$ cycles, $SE = 1.22$ for nonliving concepts), $t_1(99) = 3.4$, $p < .001$, but not for the shared features ($M = 59.8$ cycles, $SE = 1.08$ for living concepts; $M = 61.7$ cycles, $SE = 1.15$ for nonliving concepts), $t_2(99) = 1.1$, $p > .2$. Although there is a main effect of distinctiveness in the model, this is a very similar pattern to the one in the speeded feature verification experiment.

Effect of Parameters on Network Performance

The results described above show that a simple back-propagation network can show the predicted interaction between domain and distinctiveness. This is an important demonstration that the interaction can result from the distributional properties of the semantic features of these concepts, without the need for additional assumptions about differences between semantic domains. However, one question that arises is how stable is this result to changes in the network. In this section we independently vary each of the parameters used during the training and testing of the network (while keeping all other variables constant) and investigate the effect of these changes on the network's performance. For each value of each parameter, we repeated the training and testing of 100 independent models and analyzed the error scores as described above.

Hidden Units

The number of hidden units in a back-propagation model can have important consequences on its behavior (e.g., see Seidenberg & McClelland, 1989). In the simulations reported above, the model had 50 hidden units. This was the minimum number of units with which the network could successfully train. We investigated the effect of increasing the number of hidden units by varying this parameter from 50 to 200 (this was the number of input units) in steps of 10.

Learning Rate

This variable is used during training and controls the magnitude of the change that is made to the weights on each learning trial (Rumelhart et al., 1986). It was set to 0.25 in the original simulation. This was the maximum value at which the network could successfully train. We investigated the effect of decreasing this value by varying its value from 0.05 to 0.25 (in steps of 0.05).

Momentum

This variable is used during training. Each time the weights are updated, the change made to each connection strength incorporates a proportion of the change that was made on the previous update (Rumelhart et al., 1986). This variable specifies this proportion and was set to 0.5 in the original simulation. This was the maximum value at which the network could

---

[7] The pattern of the network's behavior was stable over a range of threshold activation values between 0.5 and 0.9.
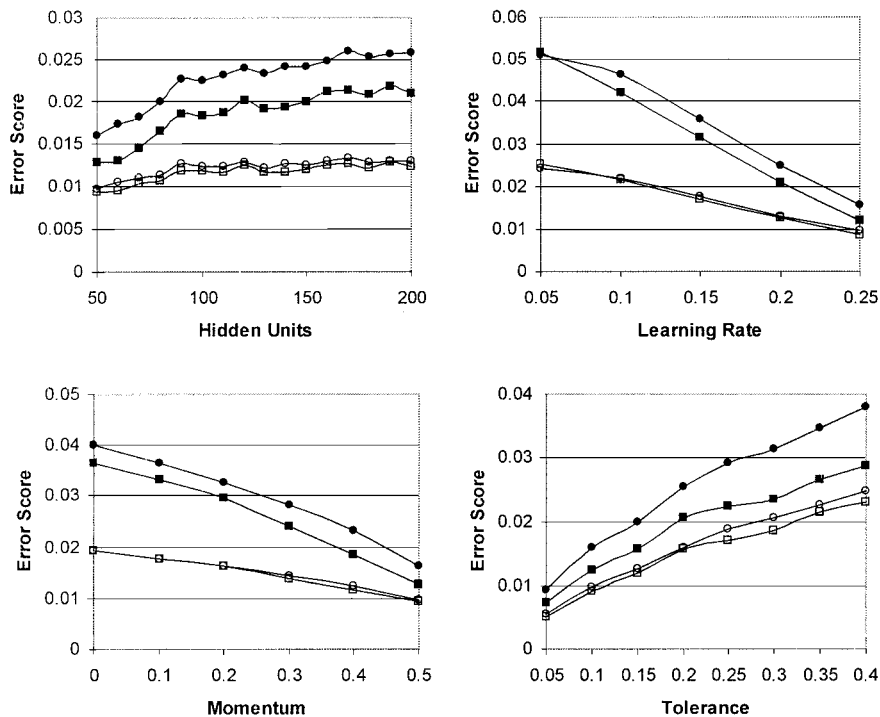
Figure 3. The effects of different parameters on network performance. Solid circles = living distinctive; solid squares = nonliving distinctive; open circles = living shared; open squares = nonliving shared.

successfully train. We investigated the effect of decreasing this value by varying its value from 0 to 0.5 (in steps of 0.1).

Tolerance

This variable specifies when the network stops training. In the original simulation it was set to 0.1, which means that for all training patterns, the activation of each output unit had to be within 0.1 of its target activation, that is, greater than 0.9 for semantic features that are present for a particular concept and less than 0.1 for all other features. We investigated the effect of this parameter by varying its value from 0.05 to 0.4 (in steps of 0.05).

Results

The results showed that for three of the four parameters (hidden units, momentum, and tolerance) the pattern of differences between the four conditions remains constant for all values of the parameter (see Figure 3). In particular, a significant interaction between domain and distinctiveness was always present (all $ps < .01$), with higher error scores for the distinctive properties of living things compared with the distinctive properties of nonliving things (all $ps < .001$). This difference was always larger than any domain difference for the shared properties.

It is interesting to note that the only parameter that did have a significant effect on the pattern of the network's performance was learning rate. Although the pattern of results was stable across most of the range, when the learning rate was set to be very small (0.05) the interaction between domain and distinctiveness disappeared. In particular, the disadvantage for the distinctive features of living things disappeared. This suggests that the effect of correlational structure is reduced when the network is making very small changes to its weights on each trial.[8]

In summary, both the original model and the cascaded activation model replicated the Distinctiveness × Domain interaction that was observed in the speeded feature verification experiment. In both cases, the distinctive properties of living things were more error-prone, or took longer to settle than the distinctive properties of nonliving things, while there was no such difference for the shared properties. This shows that a system with *no* other sources of information can demonstrate clear processing differences for living and nonliving things, which must be driven by the differences in correlational structure inherent in concepts in the two domains (or at least as estimated by our property generation results). Further, the pattern of behavior shown by the model is stable across a wide variety of changes to the model's parameters.

However, unlike the behavioral experiment, the model also showed a significant main effect of distinctiveness, with distinctive properties being more error-prone overall than shared properties. This is perhaps not surprising, because by definition, the model encounters shared properties more frequently than distinctive properties during training (i.e., shared properties are present in many concepts, distinctive in only a few), and so they will be better learned. This will lead to an overall advantage for shared

---

[8] It is also possible that this effect of learning rate relates to the sparsity of the patterns. In general, more sparse patterns need a higher learning rate because they have fewer units contributing to the error score. On average, the nonliving concepts have fewer features than the living concepts, therefore learning the more sparse nonliving concepts will be particularly difficult at small learning rates. This may, in part, be responsible for the relative improvement in performance for the living concepts when the learning rate is small.

properties. The more puzzling aspect of the results perhaps is that human participants do not also show the same overall advantage for the more frequently occurring shared properties. We suggest that the explanation for this lack of main effect of distinctiveness in humans is likely to be due to the complex interactions of the additional semantic variables, which we highlighted in the introduction to the modeling section. Although more infrequently encountered in absolute terms—for humans just as for the model—distinctive properties are perhaps more salient, more important in causal explanations and in their relevance to people's everyday experience; thus, their relatively infrequent occurrence across the entire domain is at least partially compensated for, and error rates reduced. Because the simulations had no such additional information (by definition, as this was their *raison d'etre*), the relatively low incidence of distinctive properties was directly revealed in their overall error rate. While this explanation may be somewhat speculative, the unpredicted distinctiveness effect does not seriously undermine the contribution of the simulation, because (a) the distinctiveness main effect is readily explained within the model by the relative frequency of exposure to distinctive and shared properties during training, and (b) our specific prediction was about the Distinctiveness × Domain interaction, which is reliably shown across all the simulations.

General Discussion

The purpose of the experiments reported here was to investigate one of the key claims of the conceptual structure account—namely, that the distinctive properties of living things will be slower to activate in the semantic representation compared with the distinctive features of nonliving things. This prediction was confirmed in a speeded feature verification task, which showed a significant interaction between domain and property type, with the distinctive properties of living things being verified more slowly than the distinctive properties of nonliving things.

These results provide support for one of the major, and perhaps counterintuitive, claims of the conceptual structure account. Tyler et al. (2000) and Tyler and Moss (2001) argued that because the distinctive features of living things are only weakly correlated with other features, this makes them particularly slow to be activated in comparison with the distinctive features of nonliving things, which tend to be correlated with each other. This weaker activation is predicted to reveal itself in slower RTs and in higher error rates, as was the case in the present experiment. The slower RTs for the distinctive properties of living things suggest that their access is slower than to shared features. Within the conceptual structure account, we claim that what will delay access and make performance more error-prone is the fact that the distinctive features of living things are not highly correlated with other features. Their activation is not facilitated by the strong coactivation of other features.

We also found that error rates and RTs to the shared features of living things and nonliving things were very similar. Given that our property norms show that living things have more shared features (and therefore more correlated shared features), we might have predicted differences here, with faster RTs to the shared properties of living things compared with the shared properties of nonliving things. However, the finding that the distinctive properties of living things generated significantly slower RTs and more errors than any of the other three experimental conditions, and that

the mean RT of the three other conditions was very fast (481 ms), suggests that participants were operating at their limits in this speeded paradigm. At these very fast latencies, small differences might not be detectable. Thus, although it remains a possibility that differences in the degree to which shared properties are correlated for living and nonliving things may be reflected in differences in the speed with which they are activated, it may not have been possible to observe these differences using the present paradigm.

The same interaction between domain and distinctiveness that was present in the speeded feature verification task was also observed in our computational simulation, in which we trained a network on the mapping from orthography to semantics. With two different measures of the network's performance, there was a significant effect of domain for the distinctive but not the shared features. Because the internal structure of the concepts is the only source of variation available to the model, this provides strong evidence that this interaction arises because of differences in the structure of the concepts.

The specific pattern of results that we have obtained in the two studies reported here—the Domain × Distinctiveness interaction—cannot be accounted for on any other existing model of category-specific deficits. In the domain-specific account (Caramazza & Shelton, 1998), conceptual knowledge is organized by category, with evolutionary demands generating specialization for a small number of categories such as living things and plant life. On this account we might expect performance on living things to be facilitated (i.e., overall faster RTs and fewer errors) relative to performance on nonliving things. Our results do not support this account. Neither are our results compatible with the claims of the sensory–functional account. Warrington and Shallice (1984) argued that perceptual features are more salient in the representation of living things compared with nonliving things and that functional features are more salient in the representation of nonliving things. This distinction would predict more errors and slower RTs for the functional properties of living things and for the perceptual features of nonliving things. However, we found no interaction between type of feature and domain. Finally, the distributed model of Gonnerman et al. (1997) claims that living things have more shared correlated features and fewer distinctive features than nonliving things overall, but the model does not specifically identify the distinctive features of living things as being weakly correlated and therefore slow to activate in the normal system.

The results of the speeded feature verification study provide strong evidence in support of feature-based accounts of conceptual knowledge (e.g., Caramazza, Hillis, Rapp, & Romani, 1990; Forde, Francis, Riddoch, Rumiati, & Humphreys, 1997; Humphreys, Riddoch, & Quinlan, 1988; McNamara & Miller, 1989). The results of the computational simulation further suggest that these features are represented in a nonfractionated distributed semantic system and that conceptual structure emerges out of this system as a result of correlational structure. This type of account clearly offers certain advantages in attempting to develop theoretical explanations for the detailed patterns of impaired and preserved knowledge exhibited by patients who have semantic deficits. For example, it provides a natural explanation for the ubiquitous finding that brain damage does not selectively impair a particular type of property or category in an all-or-none manner. Instead, deficits tend to be graded, as predicted by distributed connectionist accounts that show graceful degradation when damage is simulated (Hinton & Sejnowski, 1986). Second, it provides

a framework for interpreting the fine-grained patterns of preserved and impaired properties in patients following brain damage to the semantic system. Explaining category-specific deficits in terms of damage to distinct types of knowledge (whether categories or properties) does not in itself elucidate the structure of concepts. We suggest that this is a necessary step toward developing a theoretical account of so-called category-specific impairments.

Finally, the conceptual structure account is one of a number of unitary distributed models of semantics, such as the organized unitary content hypothesis (OUCH, Caramazza et al., 1990). However, it represents an advance over these earlier models in that it addresses one of the criticisms that has been leveled against them—they are so flexible that they can explain any pattern of deficit and are therefore theoretically unhelpful (e.g., Caramazza & Shelton, 1998). By developing specific claims about conceptual structure, based on well-supported claims in the psychological literature (e.g., Keil, 1986, 1989; Malt & Smith, 1984), we have been able to constrain the power of the account and make falsifiable predictions, thus overcoming this kind of objection.

References

Baayen, R. H., Piepenbrock, R., & van Rign, H. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology, 7,* 161–189.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience, 10,* 1–34.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Coltheart, M. (1981). The MRC psycholinguistics database. *Quarterly Journal of Experimental Psychology, 33A,* 497–505.

De Renzi, E., & Lucchelli, F. (1994). Are semantic systems separately represented in the brain? The case of living category impairment. *Cortex, 30,* 3–25.

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience, 10,* 77–94.

Dixon, M., Bub, D. N., & Arguin, M. (1997). The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive Neuropsychology, 14,* 1085–1130.

Durrant-Peatfield, M., Tyler, L. K., Moss, H. E., & Levy, J. (1997). The distinctiveness of form and function in category structure: A connectionist account. In *Proceedings of the 19th annual conference of the Cognitive Science Society* (pp. 193–198). Mahwah, NJ: Erlbaum.

Forde, E. M. E., Francis, D., Riddoch, M. J., Rumiati, R. I., & Humphreys, G. W. (1997). On the links between visual knowledge and naming: A single case study of a patient with a category-specific impairment for living things. *Cognitive Neuropsychology, 14,* 403–458.

Gainotti, G., Silveri, M. C., Daniele, A., & Giustolisi, L. (1995). Neuroanatomical correlates of category-specific semantic disorders: A critical survey. *Memory, 3,* 247–264.

Gonnerman, L. M., Andersen, E. S., Devlin, J. T., Kempler, D., & Seidenberg, M. S. (1997). Double dissociation of semantic categories in Alzheimer's disease. *Brain and Language, 57,* 254–279.

Greer, M. J., van Casteren, M., McLellan, S. A., Moss, H. E., Rodd, J., Rogers, T., & Tyler, L. K. (2001). The emergence of semantic categories from distributed featural representations. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the Cognitive Science Society* (pp. 358–363). Mahwah, NJ: Erlbaum.

Hinton, G., & Sejnowski, T. J. (1986). Learning and relearning in Boltzman machines. In D. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing* (pp. 282–317). Cambridge, MA: MIT Press.

Humphreys, G. W., & Forde, E. M. E. (2001). Category-specific deficits: A major review and presentation of the hierarchical interactive theory (HIT). *Behavioural and Brain Sciences, 24,* 453–465.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67–103.

Keil, F. (1986). Conceptual domains and the acquisition of metaphor. *Cognitive Development, 1,* 73–96.

Keil, F. (1989). *Concepts, kinds and cognitive development.* Cambridge, MA: MIT Press.

Laiacona, M., Barbarotto, R., & Capitani, E. (1993). Perceptual and associative knowledge in category specific impairment of semantic memory: A study of two cases. *Cortex, 29,* 727–740.

Lamberts, K., & Shapiro, L. (2002). Exemplar models and category-specific deficits. In E. M. E. Forde & G. W. Humphreys (Eds.), *Category specificity in brain and mind* (pp. 291–311). Hove, UK: Psychology Press.

Lambon Ralph, M. A., Howard, D., Nightingale, G., & Ellis, A. W. (1998). Are living and non-living category specific deficits causally linked to impaired perceptual or associative knowledge? Evidence from a category-specific double dissociation. *Neurocase, 4,* 311–338.

Madole, K., Oakes, L., & Cohen, L. (1993). Developmental changes in infants' attention to function and form–function correlations. *Cognitive Development, 8,* 189–209.

Malt, B. C., & Smith, E. E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior, 23,* 250–269.

McNamara, T. P., & Miller, D. L. (1989). Attributes of theories of meaning. *Psychological Bulletin, 106,* 335–376.

McRae, K., Cree, G. S., Westmacott, R., & de Sa, V. R. (1999). Further evidence for feature–feature correlations in semantic memory. *Canadian Journal of Experimental Psychology, 53,* 360–373.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99–130.

Moss, H. E., McCormick, S. F., & Tyler, L. K. (1997). The time course of activation of semantic information during spoken word recognition. *Language and Cognitive Processes, 12,* 695–731.

Moss, H., & Older, L. (1996). *Birkbeck word association norms.* Hove, UK: Taylor & Francis.

Moss, H. E., & Tyler, L. K. (2000). A progressive category-specific semantic deficit for non-living things. *Neuropsychologia, 38,* 60–82.

Moss, H. E., Tyler, L. K., & Devlin, J. (2002). The emergence of category specific deficits in a distributed semantic system. In E. M. E. Forde & G. W. Humphreys (Eds.), *Category specificity in brain and mind* (pp. 115–148). Hove, UK: Psychology Press.

Moss, H. E., Tyler, L. K., Durrant-Peatfield, M., & Bunn, E. M. (1998). "Two eyes of a see-through": Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase, 4,* 291–310.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* Retrieved from http://www.usf.edu/FreeAssociation/

Parks, R., Ray, J., & Bland, S. (1998). *Wordsmyth English dictionary–thesaurus.* Chicago: University of Chicago Press. Retrieved February 1, 1999, from http://www.wordsmyth.net/Chicago

Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language, 46,* 245–266.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning

internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing* (Vol. I, pp. 318–364). Cambridge, MA: MIT Press.

Saffran, E. M., & Schwartz, M. F. (1994). Of cabbages and things: Semantic memory from a neuropsychological perspective: A tutorial review. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconscious information processing* (pp. 507–536). Cambridge, MA: MIT Press.

Santos, L. R., & Caramazza, A. (2002). The domain-specific hypothesis: A developmental and comparative perspective on category-specific deficits. In E. M. E. Forde & G. W. Humphreys (Eds), *Category-specificity in brain and mind* (pp. 1–24). Hove, UK: Psychology Press.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Sheridan, J., & Humphreys, G. W. (1993). A verbal-semantic category-specific recognition impairment. *Cognitive Neuropsychology, 10,* 143–184.

Silveri, M. C., & Gainotti, G. (1988). Interaction between vision and language in category specific impairment. *Cognitive Neuropsychology, 5,* 677–709.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113,* 169–197.

Tyler, L. K., & Moss, H. E. (1997). Functional properties of word meaning: Studies of normal and brain-damaged patients. *Cognitive Neuropsychology, 14,* 511–545.

Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences, 5,* 244–252.

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

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction-time analysis. *Journal of Experimental Psychology: General, 123,* 34–80.

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain, 110,* 1273–1296.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107,* 829–854.