Category-Specific Semantic Deficits: The Role of Familiarity and Property Type Reexamined

Elaine M. Bunn and Lorraine K. Tyler
Birkbeck College, University of London

Helen E. Moss
University of Glasgow

Category-specific deficits for living things have been explained variously as an artifact due to differences in the familiarity of concepts in different categories (E. Funnell & J. Sheridan, 1992) or as the result of an underlying impairment to sensory knowledge (E. K. Warrington & T. Shallice, 1984). Efforts to test these hypotheses empirically have been hindered by the shortcomings of currently available stimulus materials. A new set of stimuli are described that the authors developed to overcome the limitations of existing sets. The set consists of color photographs, matched across categories for familiarity and visual complexity. This set was used to test the semantic knowledge of a classic patient, J.B.R. (E. K. Warrington & T. Shallice, 1984). The results suggest that J.B.R.’s deficit for living things cannot be explained in terms of familiarity effects and that the most severely affected categories are those whose identification is most dependent on sensory information.

Category-specific impairments of semantic memory force us to consider how our stored knowledge about the world may be structured such that brain damage can produce a disproportionate deficit for items from certain conceptual categories. The most commonly reported pattern is for the broad category of living things to be more impaired than nonliving things (Farah, McMullen, & Meyer, 1991; Hart & Gordon, 1992; Laitacona, Barbarotto, & Capitani, 1993; Sartori & Job, 1988; Sheridan & Humphreys, 1993; Silveri & Gainotti, 1988; Warrington & Shallice, 1984), although the opposite pattern has also been observed (Hillis & Caramazza, 1991; Sacchetti & Humphreys, 1992; Warrington & McCarthy, 1983, 1987; see Saffran and Schwartz, 1994, for a review). These dissociations between knowledge of living and nonliving things could be taken to suggest that the organization of semantic knowledge in the brain honors these categorical distinctions.

Category Effects and Confounding Variables

An alternative view is that apparently selective deficits for natural kinds arise because living things tend to be at a disadvantage relative to nonliving things on several dimensions that are known to affect processing difficulty (Funnell & Sheridan, 1992; Riddoch & Humphreys, 1987; Stewart, Parkin, & Hunkin, 1992). In support of this hypothesis, Stewart et al. (1992) and Funnell and Sheridan (1992) have described patients whose category-specific naming deficits were eliminated when factors such as word frequency, concept familiarity, and visual complexity were taken into account. Furthermore, Funnell and Sheridan claimed that the selective deficit for living things demonstrated by a classic patient, J.B.R. (Warrington & Shallice, 1984), can also be explained in terms of the effects of concept familiarity.

However, it may be the case that category effects cannot be explained entirely in terms of confounding factors. For example, although Funnell and Sheridan (1992) demonstrated that J.B.R.’s performance was influenced by familiarity, this does not rule out the possibility that J.B.R.’s performance was also affected by category. Indeed, in the midrange of familiarity (i.e., where items would not be passed or failed simply because of their very high or very low familiarity), J.B.R. was able to define 8 of 10 items from a nonliving category (tools) that had a mean familiarity rating very similar to those of two living categories on which he failed to score (fruits and vegetables).¹

The controversy as to whether some patients demonstrate semantic deficits that are genuinely category specific has been perpetuated by the fact that (as Funnell & Sheridan, 1992, have noted) living things are generally of lower familiarity than nonliving things in one of the most widely used sets of published pictures (Snodgrass & Vanderwart, 1980). The difficulty of obtaining well-controlled sets of stimuli makes any investigation of the effect of semantic category per se on patients’ performance problematic.

A potential solution to this problem is to select small, ¹Tools had a mean familiarity rating of 3.2; fruits and vegetables had mean familiarity ratings of 3.4 and 3.2, respectively.
familiarity-matched subsets of living and nonliving things from the Snodgrass and Vanderwart set. Using two such subsets, Funnell and Sheridan (1992) found their patient S.L.’s naming to be influenced by familiarity but not by category. However, both of these subsets are problematic. Funnell and Sheridan noted that in the first subset, the familiarity of living things is confounded with semantic subcategory (i.e., fruits/vegetables vs. animals), making it difficult to separate the effects of these two variables. The second subset was designed to overcome this problem, but its small number of stimuli means that it lacks statistical power. Furthermore, in both subsets, the nonliving things consist of items from a mixture of different subcategories, which could disguise category differences. Perhaps most important, a third of the nonliving things are musical instruments. This is potentially of some significance because it has been suggested that, for patients with selective deficits for living things, performance on musical instruments may be atypical of performance on nonliving things in general (Warrington & Shallice, 1984; see later discussion).

The use of small, matched subsets to overcome the problem of potentially confounding variables has a number of disadvantages. An alternative approach is to partial out the effects of these factors using statistical techniques (Farah et al., 1991; Laiacona et al., 1993; see Gaffan and Heywood, 1993, for a critique). However, this type of approach also has its drawbacks. One problem is that the relevant variables (such as word frequency and concept familiarity) tend to be significantly correlated with one another (Carroll & White, 1973; Gilhooley & Logie, 1980; Stratton, Jacobus, & Brinley, 1975). This makes their use as predictors in, for example, a multiple regression analysis difficult, because this multicollinearity suggests that these variables load onto the same factor to a large degree. Furthermore, the use of this type of analysis is dependent on other factors, such as being able to produce a linear relationship between predictor and dependent variables and having sufficient data points to justify the use of such powerful techniques.

A more satisfactory solution to the problem of confounding variables would be to develop a large set of pictures containing familiarity-matched subcategories of living and nonliving things. Such a set would allow more adequate investigations of the effects of semantic category and would also show whether general category effects vary across different subcategories of living and nonliving things.

### Category Effects and Semantic Attributes

A second motivation for developing a new set of picture stimuli is to test hypotheses that explain category-specific deficits in terms of an underlying impairment to knowledge about particular types of semantic properties. For example, Warrington and Shallice (1984) suggested that category-specific deficits for living and nonliving things could arise as a result of damage to semantic systems based on sensory and functional features, respectively (see also, Warrington & McCarthy, 1983, 1987). They were led to this conclusion by certain anomalies in J.B.R.’s pattern of performance. Although J.B.R. generally appeared to have impaired knowl-

\[2\] A standard assumption across many models of semantic memory is that access to some form of presensational representation of input must precede the retrieval of semantic information. In the case of pictorial input, these representations would take the form of structural descriptions, that is, abstract representations of object surfaces (Marr, 1982). These structural descriptions would allow the semantic system to be accessed from vision, even if sensory information within the semantic system itself was impaired.
to which knowledge about different types of attributes are dependent on one another may differ across semantic categories (De Renzi & Lucchelli, 1994; Gonnerman, Andersen, Devlin, Kempler, & Seidenberg, 1997; Tyler, Durrant-Pagefield, Levy, Voice, & Moss, 1996). Hence, results from this type of study are open to a number of alternative interpretations. Poor knowledge of living relative to nonliving things, regardless of the type of information tested, may be indicative of an underlying impairment that is category specific rather than attribute specific, or it may arise because it is difficult to measure a patient's sensory and functional knowledge independently of one another.

An alternative approach is to examine a patient's ability to identify stimuli from subcategories of living and nonliving things that appear to differ in terms of the extent to which their identification is reliant on sensory or functional characteristics. This approach may be useful, because it does not rely on an examination of the sensory and functional information to which a patient has explicit access. Warrington and Shallice's (1984) examination of J.B.R.'s performance on body parts and musical instruments is an example of this type of approach. However, the results obtained are undermined by the differences in familiarity between the subcategories used. Body parts tend to be more familiar than most other living things, whereas musical instruments tend to be less familiar than many other artifacts. By contrast, the subsets selected from the Snodgrass and Vanderwart (1980) pictures by Funnell and Sheridan (1992) do control for the effects of familiarity, but they are too small to allow comparisons between different subcategories of living and nonliving things. For comparisons of interest to be made, a picture set is required that includes a reasonable number of familiarity-matched items from theoretically important subcategories.

The first aim of the present research was to develop a large set of colored photographs of familiarity-matched living and nonliving things for use with neurological patients. One of the main requirements of this set was that it should include sets of items from the subcategories of living and nonliving things that have been identified as being of potential theoretical importance.

The picture set was then used to investigate the semantic knowledge of a patient, J.B.R., who was originally described by Warrington and Shallice (1984). The set was used to investigate whether J.B.R.'s apparently category-specific impairment can be explained in terms of the effects of familiarity alone. In addition, the set was used to examine whether J.B.R.'s performance across familiarity-matched subcategories of living and nonliving things supports the claim that deficits for living things are the result of damage to sensory knowledge. These results were then compared to results obtained with the Snodgrass and Vanderwart (1980) stimuli that have been widely used in earlier research into semantic impairments. Finally, to establish whether any effects of semantic category on J.B.R.'s ability to name pictures were the result of a central semantic impairment rather than a more peripheral deficit, we tested J.B.R. on a number of tasks that did not involve pictorial input and on tasks that did not require a naming response.

Patient Background

J.B.R. is a 40-year-old man who sustained brain damage as a result of herpes simplex encephalitis at the age of 23. He was 1 of 4 patients with category-specific deficits described by Warrington and Shallice (1984) and has subsequently been a participant in a number of further studies (Funnell, 1996; Funnell & Sheridan, 1992; Wilson, 1994, 1997; Wilson, Clare, Young, & Hodges, in press). J.B.R. currently appears to be essentially unchanged since his assessment by Warrington and Shallice (1984). His spontaneous speech is fluent, with no syntactic problems. However, he does demonstrate some word-finding difficulties. He is also densely amnesic. A CT scan presented by Warrington and Shallice (1984) showed widespread low attenuation, maximal in both temporal lobes.

A magnetic resonance imaging (MRI) scan was performed on J.B.R. in February 1996 at the Wellcome Institute for Cognitive Neurology (see Figure 1). On a T1-weighted sequence, the anterior halves of the temporal lobes were hypointense, and this involved the medial and lateral temporal cortex to an equal degree. The hypointensity extended about 1 cm more posteriorly on the right than on the left. The lateral and fourth ventricles and the lateral sulci were bilaterally and symmetrically enlarged. The entire right hippocampus was destroyed, and the posterior part of the left hippocampus was preserved. The amygdala was totally destroyed.

General Cognitive Functioning

On the Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981), J.B.R. has a verbal IQ of 80, a performance IQ of 80, and a full-scale IQ of 80 (Wilson, 1997). These scores are within the low normal range. However, as Wilson has noted, it is likely that J.B.R. was of at least good average ability premorbidly, because he had been an undergraduate student before his illness.

J.B.R. has poor memory, with a screening score of 0 of 12 on the Rivermead Behavioural Memory Test (Wilson, 1997; Wilson, Cockburn, & Baddeley, 1985). In addition, he scored only 3 of 36 when attempting to recall the Rey-Osterrieth Complex Figure Design (Osterrieth, 1944), despite achieving a score of 34 of 36 when copying the figure (which is at the 70th percentile; Lezak, 1995). He has a forward digit span of 6 (Wilson, 1997).

Wilson (1994) has described J.B.R.'s reading in detail. J.B.R. has an acquired surface dyslexia and has more difficulty with irregular than regular words. His spelling is also very poor. Finally, J.B.R. has normal visuospatial abilities, scoring a maximum score of 20 on the Visual Object and Space Perception Battery (Warrington & James, 1991) screening test (Wilson, 1997).

Semantic Memory

Warrington and Shallice (1984) reported that J.B.R. failed to score on the Graded Naming Test (McKenna & Warrington, 1983). He also scored very poorly on the Peabody Picture Vocabulary Test and the Shallice and McGill Con-
crete/Abstract Picture Word Matching Test. Clinical observation suggested that he might have more difficulty with items from certain semantic categories than from others. Warrington and Shallice (1984) therefore carried out a number of studies that demonstrated that J.B.R. had an impairment of semantic memory that was much more severe for living than for nonliving things. This general pattern of performance was observed using both visual and verbal stimuli.

Experimental Investigations

*J.B.R. Picture Naming I: Snodgrass and Vanderwart (1980) Set*

To investigate whether J.B.R. still shows the category effect reported by Warrington and Shallice (1984) when a standard set of materials is used, we initially tested his ability to name pictures using the whole of the Snodgrass and Vanderwart (1980) stimulus set.

*Results and Discussion*

On this set, J.B.R. named 93 of 160 (58%) artifacts correctly, as compared to only 25 of 97 (26%) natural kinds. This difference is highly significant, \( \chi^2(1, N = 257) = 24.17, p < .001 \). He also named three of three (100%) foods (which seem to cross the boundary between natural kinds and artifacts). These results are broadly similar to those reported by Warrington and Shallice (1984) who found (using colored pictures) that J.B.R. was able to name 67% of inanimate objects correctly, as compared to only 6% of living things.

As reported by Warrington and Shallice (1984), J.B.R.’s score on the category of body parts was not typical of his scores for living things in general; he named 100% of items from this category correctly. However, in contrast to the results obtained by Warrington and Shallice, his score for musical instruments (50%) seemed to be more similar to his performance on other nonliving things than on living things.3

Overall, therefore, when the Snodgrass and Vanderwart (1980) stimuli are used, J.B.R. demonstrates a semantic deficit that appears to be disproportionately severe for living things.

---

3 Warrington and Shallice’s results for these subcategories were obtained by asking J.B.R. to define the spoken names of these items.
things. However, this category effect may reflect the fact that living things are generally of lower familiarity than nonliving things in this set. As previously noted, Funnell and Sheridan (1992) have selected familiarity-matched subsets of stimuli from this set to try and deal with the problem of familiarity effects. J.B.R.'s performance on these subsets was therefore examined to determine whether these subsets do in fact provide a satisfactory solution to the problem.


The first subset used by Funnell and Sheridan (1992) consists of two sets of 12 living and 12 nonliving things taken from the Snodgrass and Vanderwart (1980) corpus. One of the sets contains low-familiarity items, while the other contains high-familiarity items. Items within each set are matched across categories for familiarity, visual complexity, and word frequency.

**Results and Discussion**

On the low-familiarity set, J.B.R. scored 1 of 12 for living things and 3 of 12 for nonliving things. This difference is not statistically significant, \( \chi^2(1, N = 24) = 0.30 \). On the high-familiarity set, he scored 2 of 12 for living things and 9 of 12 for nonliving things, which is a significant difference, \( \chi^2 = 6.04, p < .02 \). Collapsing across the high- and low-familiarity sets, there was a marginally significant effect of category, \( \chi^2(1, N = 48) = 2.88, p < .10 \); collapsing over living and nonliving things, there was a marginally significant effect of familiarity, \( \chi^2(1, N = 48) = 3.49, p < .10 \). J.B.R.'s scores on this subset therefore appear to suggest that there is only a minor effect of semantic category on his performance on high-familiarity items. These results contrast markedly with those of Funnell and De Mornay Davies (1996; see also, Funnell, 1996), who reported that it was only on low-familiarity items that J.B.R.'s naming was significantly affected by semantic category.

However, as previously noted, Funnell and Sheridan (1992) recognized that one problem with this subset was that a category effect confined to either the low- or high-familiarity set could reflect the fact that the low-familiarity living things consist entirely of animals, whereas the high-familiarity living things consist mainly of fruits and vegetables. Hence, these authors assembled a further set of 12 animal and 12 object pictures from the Snodgrass and Vanderwart (1980) set, in which half of the animals had a familiarity rating greater than 3, and half had a familiarity rating of less than 3. Nonliving things were matched with the living things for familiarity, complexity, and frequency.

J.B.R.'s scores on this second subset of items were also examined. He was able to name five of six animals and six of six nonliving things from the higher familiarity set and zero of six animals and zero of six nonliving things from the lower familiarity set. Collapsing over high- and low-familiarity items, there was no effect of semantic category, \( \chi^2(1, N = 24) = 0 \); but collapsing over categories, there was a significant effect of familiarity, \( \chi^2(1, N = 24) = 16.78, p < .001 \). J.B.R.'s scores on this second subset, therefore, suggest that his ability to name pictures is influenced by familiarity rather than semantic category.

Once again, these results are difficult to interpret. This is because J.B.R. appears to be virtually at ceiling on the higher familiarity items and at floor on the lower familiarity items. Hence, although there does appear to be a strong effect of familiarity on his performance, this subset does not provide any information as to what might happen in the midrange of familiarity (i.e., where items would not be passed or failed simply because of their very high or very low familiarity). It may be only in this range that semantic category would be seen to exert an effect.

**Conclusions on Current Stimulus Sets**

These results illustrate some of the problems inherent in using the Snodgrass and Vanderwart (1980) pictures to investigate apparently category-specific deficits of semantic memory. Across the set as a whole, deficits for living things may be explained in terms of differences in concept familiarity between categories. The smaller, familiarity-matched subsets selected by Funnell and Sheridan (1992) are also problematic. For example, in the first of these subsets, concept familiarity is confounded with semantic subcategory, whereas the second subset is too small to investigate performance across the full range of familiarity. Finally, a major drawback of these subsets is that they are too small to allow any investigation of whether similar patterns of results are obtained across different subcategories of living and nonliving things.

Because of the difficulty in obtaining a satisfactory set of stimuli from the Snodgrass and Vanderwart (1980) corpus, a new picture set was developed.

**Development of the Centre for Speech and Language (CSL) Set**

Our aim in developing a new picture set was to produce a large set of picture stimuli that would allow comparisons to be made between familiarity-matched sets of living and nonliving things. An additional requirement of the set was that it should include reasonable numbers of items from theoretically interesting subcategories of these two domains.

**Method**

**Materials**

A total of 731 colored photographs of everyday objects and living things was collected from books and magazines for pretesting with control participants. Each picture was glued onto a plain white card measuring 15 × 10 cm. The cards were put into photograph albums, which were given to participants to work through at their own pace.

**Participants**

Each pretest was completed by 40 participants. The participants were undergraduate students attending Birkbeck College who were
between 19 and 42 years old. Participants were each paid £5 for their participation.

**Design and Procedure**

Three pretests were carried out to obtain (a) normative naming data, (b) ratings of concept familiarity, and (c) ratings of visual complexity.

**Naming.** Participants were asked to write down the name of the item in each picture. If they did not recognize the target object, or if they recognized it but did not know its name, they were asked to indicate that this was the case.

**Familiarity ratings.** Participants were asked to rate the familiarity of each item. The instructions given to participants were based on those used by Snodgrass and Vanderwart (1980). Hence, participants were asked to rate how unusual or unfamiliar the object was from their experience, that is, how often they came into contact with the object in real life and how often they thought about the object. Ratings were made on a 5-point scale, where 1 indicated *very unfamiliar* and 5 indicated *very familiar*. Participants were instructed that they should rate the familiarity of the object itself, rather than the way in which it was pictured.

**Complexity ratings.** Participants were asked to rate the visual complexity of each item. Once again, the instructions given to participants were based on the instructions used by Snodgrass and Vanderwart (1980). Complexity was defined as “the amount of detail, or the intricacy of line in the picture.” Ratings were made on a 5-point scale, where 1 indicated *very simple* and 5 indicated *very complex*. Participants were instructed that they should rate the complexity of the picture itself, rather than the complexity of the real-life object that it represented.

For the ratings of both familiarity and complexity, participants were asked to try and make use of the whole range of the scale.

**Selection of Pictures for the CSL Set**

For each item, the percentage of participants who had produced an acceptable name (as agreed by the experimenters) and the mean familiarity and complexity ratings were calculated. Any items for which fewer than 70% of participants had produced an acceptable response were excluded. From the remaining items, sets of living and nonliving things were selected.

The first requirement of the new picture set was that it should include sets of living and nonliving things that were closely matched for familiarity and visual complexity and that were matched as far as possible for percentage correct naming and word length (where word length was taken as the length in letters of the most frequently given name for the object). A second aim was that the set should include reasonable numbers of items from subcategories of living and nonliving things that have been identified as being of particular theoretical importance. For example, as previously noted, it has been suggested that patients’ performance on body parts and musical instruments may be atypical of their performance on living and nonliving things in general (Warrington & Shallice, 1984); items from these subcategories were therefore included in the new picture set. It has also been claimed that performance on small, manipulable objects may dissociate from performance on larger, less manipulable objects (Warrington & McCarthy, 1987). The subcategory of furniture was therefore included to represent the latter form of object. Finally, the subcategory of foods is of particular interest, because it crosses the boundary between natural kinds and artifacts. Furthermore, many herpes simplex encephalitis patients have particular difficulty with food items (Greenwood, Bhalla, Gordin, & Roberts, 1983). This subcategory is scarcely represented in the Snodgrass and Vanderwart (1980) set (which includes only three foods), and as a result has often been overlooked in studies of patients with semantic deficits. A large number of items from this subcategory were therefore selected for inclusion in the new picture set.

**Characteristics of the CSL Picture Set**

The final picture set consists of a total of 227 colored photographs taken from 13 different subcategories of living and nonliving things. The subcategories are: animals (n = 20), insects (n = 12), birds (n = 9), fruits (n = 16), vegetables (n = 21), foods (n = 34), body parts (n = 22), vehicles (n = 20), toys (n = 12), household tools (n = 16), clothing (n = 21), musical instruments (n = 10), and furniture (n = 14).

It was not possible to select the sets such that all of the subcategories were familiarity-matched with one another, because there was considerable variation between subcategories in terms of the familiarity range into which their members fell. However, four of the subcategories of living things (animals, insects, fruits, and vegetables) are closely matched (pairwise) for familiarity and visual complexity (and matched as far as possible for percentage correct naming and word length) with four subcategories of nonliving things (vehicles, toys, household tools, and clothing, respectively). This means that performance on each of these subcategories of living things can be directly compared with performance on a matched subcategory of nonliving things. Furthermore, taken together, these subcategories provide a relatively large corpus of familiarity-matched living and nonliving things that may be used to compare patients’ performance across these two domains.

The set also incorporates a further five subcategories (birds, foods, body parts, furniture, and musical instruments) that are not directly matched with a subcategory from the opposite domain. (A summary of the properties of the picture set is given in Appendix A.) However, a subset of 21 foods is matched with the subcategories of vegetables and clothes, and the subcategory of furniture is matched with subsets of 14 fruits and household tools. This allows further comparisons to be made across subcategories.

This set of pictures has a number of advantages over sets that have been used previously to investigate category effects demonstrated by neurologically impaired patients. Unlike the Snodgrass and Vanderwart (1980) set, the four main subcategories of living and nonliving things are matched for concept familiarity and visual complexity. This means that any effect of the living–nonliving variable obtained with this set cannot be explained in terms of these potentially confounding factors. In addition, the total number of items is much larger than in the subsets selected by Funnell and Sheridan (1992), and hence the set has a number of advantages over those subsets. First, the larger number of items means that the present set gives greater statistical power. The large number of stimuli also allows a patient’s performance to be examined over a wider range of familiarity and across a number of different subcategories of living and nonliving things. This allows a more adequate examination of a patient’s performance on theoretically interesting subcategories (such as body parts, musical instruments, and foods) than other stimulus sets. For example, the fact that J.B.R. named 100% of the foods in the Snodgrass and Vanderwart (1980) set correctly may be misleading, because this score is based on his responses to only three items. The large number of stimuli in the CSL set also means that there is less likelihood of a particular range of familiarity being occupied solely by members of a particular subcategory. Furthermore, there is a reduced likelihood that any results obtained will be an artifact of the particular items
selected to represent a subcategory. Finally, the CSL set has the 
advantage that it consists of colored photographs rather than 
black-and-white line drawings (cf. Snodgrass & Vanderwart, 
1980). This is an important factor, because it is possible that using 
black-and-white line drawings as stimuli puts items from certain 
subcategories where color is an important distinguishing attribute 
(such as fruits and vegetables) at a particular disadvantage.

Results

Control Participants

Normative naming data for the CSL picture set was 
obtained from both young and old control participants to 
provide a baseline against which to interpret the performance of neurological patients.

Young controls. The data obtained from the pretests 
using 40 young control participants were used to examine 
whether the mean percentage of participants giving a correct 
name for each item varied across the matched subcategories. 
Overall, living things were named significantly more poorly 
than nonliving things (93% vs. 96%), t(68) = -2.40, p < .05. No significant difference was found between participants’ naming of animals and vehicles (97% correct vs. 
94%), fruits and household tools (95% vs. 98%), or vegetables and clothes (91% vs. 95%). However, naming of insects was significantly poorer than naming of toys (86% vs. 94%), t(11) = -2.54, p < .05.

Naming of the subset of foods that were matched with 
vegetables and clothes (95%) did not differ significantly 
from naming of either of these subcategories. Furthermore, 
naming of furniture (97%) did not differ significantly from 
either of the subsets of fruits (96%) and household tools 
(98%) with which this subcategory was matched.

Older controls. Eight older control participants (mean 
age = 64.6, SD = 3.7) were also asked to name the CSL 
picture set aloud. Overall, their mean score on living things 
(M = 95%, SD = 5%) was not significantly different from 
their mean score on nonliving things (M = 95%, SD = 2%). However, there were significant differences between their mean scores for animals and vehicles (M = 99%, SD = 2% 
vs. M = 94%, SD = 4%), t(7) = 3.81, p < .01, and insects 
and toys (M = 87%, SD = 11% vs. M = 96%, SD = 5%), 
t(7) = -2.55, p < .05. It is noteworthy that, in the case of 
animals versus vehicles, items from the living category were 
named significantly better. The differences between participants’ mean scores on fruits and household tools (M = 97%, 
SD = 5% vs. M = 93%, SD = 4%) and between vegetables 
and clothes (M = 93%, SD = 7% vs. M = 97%, SD = 4%) were not statistically significant.

Naming of the subset of foods matched with vegetables 
and clothes (M = 94%, SD = 2%) did not differ significantly from naming of vegetables, but it was significantly worse than naming of clothes, t(7) = -2.38, p < .05. Naming of furniture (M = 98%, SD = 3%) did not differ significantly from naming of the subset of fruits (M = 98%, 
SD = 3%), but it was significantly better than the naming of the subset of household tools (M = 92%, SD = 5%) with which this subcategory was matched, t(7) = 2.96, p < .05.

J.B.R. Picture Naming III: The CSL Set

J.B.R. was asked to name each item from the CSL picture 
set aloud. This was carried out over two testing sessions, 
with half of the stimuli being presented in each session. The 
stimuli were presented in random order. The results obtained 
are shown in Table 1. Table 1 also shows his performance on 
the equivalent subcategories from Snodgrass and Vander-
wart (1980).

It can be seen that, for the matched sets of living and 
nonliving things from the CSL picture set, there was a large 
effect of semantic category on J.B.R.’s performance. He 
named living things significantly more poorly than nonliv-
ing things, $\chi^2(1, N = 138) = 27.99, p < .001$. Although the 
young control participants were also significantly poorer at 
naming living than nonliving things, the magnitude of this 
difference (3%) was much smaller than that demonstrated by 
J.B.R. (45%). As previously noted, there was no significant 
difference between the older controls’ naming of living and 
nonliving things. Furthermore, the difference between 
J.B.R.’s total scores for living and nonliving things was 
more than 2$SD$ above the mean of the older controls 
(M = 0.4%, SD = 4.6).

J.B.R. named living things more poorly than nonliving 
things for all four of the main pairs of matched subcatego-
ries. This effect was found to be statistically significant for 
the sets of animals versus vehicles, $\chi^2(1, N = 40) = 3.84, 
p < .05$, and vegetables versus clothing, $\chi^2(1, N = 42) = 
18.84, p < .001$. The differences between insects and toys 
and between fruits and household tools did not reach significance, $\chi^2(1, N = 24) = 2.70$ and $\chi^2(1, N = 32) = 
2.07$, respectively.

J.B.R. correctly named 5 of 21 (24%) of the subset of 
foods that were matched with vegetables and clothes. This 
score was not significantly different from his score for 
vegetables, $\chi^2(1, N = 42) = 0.69$, but it was significantly 
poorer than his score for clothes, $\chi^2(1, N = 42) = 11.55, 
p < .001$. This suggests that the subcategory of foods may 
cluster with living more than nonliving things. J.B.R.’s score 
for the category of furniture was not significantly different 
from his scores on the matched subsets of fruits (4 of 
14 = 29%), $\chi^2(1, N = 28) = 2.30$, and household tools (8 
of 14 = 57%), $\chi^2(1, N = 28) = 0$. There is therefore no 
evidence that his performance on small, manipulable objects 
differs significantly from his performance on larger, less 
manipulable objects.

Neither set of control participants demonstrated any of the 
significant differences between the main paired subcatego-
ries that were shown by J.B.R. Furthermore, although both 
the young and older control participants had shown signifi-
cantly poorer naming of insects than toys, the differences 
between J.B.R.’s scores for all four of the main paired 
subcategories of living and nonliving things were more than 
2$SD$ above the mean of the older controls. For the remaining 
subcategories, the young control participants did not dem-
strate any of the significant effects shown by J.B.R. The only 
effect that was also shown by the older control participants
was the significantly poorer naming of the subset of foods than clothes. However, the magnitude of this difference for J.B.R. (57%) was again more than 2 SD above the mean of the older controls ($M = 3.0\%, SD = 4.8\%$).

Finally, J.B.R.’s scores for both body parts (82%) and musical instruments (50%) appeared to cluster with his performance on nonliving things more than with living things. However, these results do not rule out the possibility that his good naming of body parts reflects the fact that members of this subcategory are highly familiar.

Overall, therefore, J.B.R. shows a marked deficit for living things. Furthermore, when the differences between his scores for the four main paired subcategories of living and nonliving things are compared with those of the older control participants, it can be seen that this disproportionate deficit for living things applies across all four subcategories.

**Comparison of Results Obtained With the CSL Set and the Snodgrass and Vanderwart (1980) Set**

It can be seen from Table 1 that J.B.R.’s total scores for living and nonliving things for the CSL picture set and for the Snodgrass and Vanderwart (1980) set are very similar when equivalent subcategories are used. However, the CSL picture set has the advantage that the difference between his scores on living and nonliving things cannot be explained in terms of the effects of concept familiarity or visual complexity.

By contrast, his performance on subcategories of living and nonliving things appears, at first, to differ widely across the two sets. On further examination of the sets, however, a number of possible explanations for these differences become apparent. The members of some subcategories (e.g., clothes) appear to be less familiar in the Snodgrass and Vanderwart (1980) set than in the CSL picture set; whereas for other categories (e.g., vehicles and body parts), the reverse is true. The variation in J.B.R.’s performance on equivalent subcategories from the two picture sets may be a result of these differences in familiarity. This highlights the importance of making comparisons only between familiarity-matched subcategories. Another interesting possibility is that J.B.R.’s superior naming of fruits and vegetables when the CSL picture set is used reflects the fact that these subcategories are particularly dependent on information about color, texture, and other attributes available from color photographs that is not available from black-and-white line drawings.

**The CSL Set and the Effects of Concept Familiarity**

As noted earlier, Funnell and De Mornay Davies (1996; Funnell, 1996) have recently reported that the category-specific impairment demonstrated by J.B.R. is only found for low-familiarity items. Furthermore, Funnell (1996) has argued that this undermines the view that category effects reflect underlying organizing principles of semantic memory. Our examination of J.B.R.’s performance on the subsets selected from the Snodgrass and Vanderwart (1980) set by Funnell and Sheridan (1992) did not support this claim,
because we found the effect to be limited to items of higher familiarity. However, any interactions between the effects of familiarity and category found using the Funnel and Sheridan subset could be the result of the confounding of levels of familiarity with different semantic subcategories. To investigate Funnel and De Mornay Davies' claim further, we divided items from the eight main subcategories of the CSL picture set into four familiarity bands. J.B.R.'s ability to name the living and nonliving things that fell into each of these bands was then examined.

For the most familiar items (i.e., items with a familiarity rating of 4.1–5), J.B.R. named only 6 of 17 living things, as compared to 14 of 17 nonliving things. This effect of category was found to be statistically significant, \( \chi^2(1, N = 28) = 5.95, p < .02 \). There was also a significant effect of category on his scores for the next two familiarity bands. For items with familiarity ratings between 3.1 and 4, he named 4 of 27 living and 19 of 27 nonliving things correctly, \( \chi^2(1, N = 54) = 14.84, p < .001 \); whereas for items with familiarity ratings between 2.1 and 3, he named 0 of 17 living and 6 of 16 nonliving things correctly, \( \chi^2(1, N = 33) = 5.47, p < .02 \). It was only for the least familiar items (items with a familiarity rating of 1.1–2) that the difference between his scores for living things (zero of eight) and nonliving things (two of nine) was not statistically significant, \( \chi^2(1, N = 17) = 0.44 \). The most parsimonious explanation for the latter finding is that it reflects the fact that J.B.R.'s performance was virtually at floor for this set of stimuli.

Overall, there was a significant effect of concept familiarity on J.B.R.'s picture naming. Collapsing over living and nonliving things, J.B.R.'s performance on items from the highest familiarity band (59% correct) was significantly better than his performance on items from the lowest familiarity band (7% correct), \( \chi^2(1, N = 51) = 8.40, p < .01 \).

Once again, therefore, our findings with regard to the relationship between the effects of semantic category and concept familiarity appear to contradict those reported by Funnel and De Mornay Davies (1996; Funnel, 1996). Using the CSL picture set, we did not find any evidence to support their claim that J.B.R.'s category-specific deficit is only evident for low-familiarity items. The interaction between category and familiarity observed by Funnel and De Mornay Davies may reflect the fact that while 13 of 18 (72%) of their high-familiarity living things belonged to the subcategory of body parts, none of their low-familiarity living things belonged to this subcategory. This confounding of subcategory with different levels may be significant because, using the CSL set, we found some evidence to suggest that J.B.R.'s naming of body parts is better than his naming of other subcategories of living things, even when stimuli are matched for familiarity (see later discussion).

The CSL Set and the Effects of Semantic Property Type

As previously noted, patients' performance on certain subcategories may be particularly useful in determining whether deficits for living things can be explained in terms of a primary impairment to sensory knowledge. If, as Warrington and Shallice (1984) have argued, musical instruments are atypical of most artifacts in that their identification is largely dependent on sensory information, whereas body parts are atypical of most living things in that their identification is more dependent on functional information, performance on these subcategories may help to tease apart the effects of semantic category and property type.

It can be seen from Table 1 that when the CSL stimuli are used, J.B.R.'s naming of musical instruments and body parts appears to cluster with his performance on nonliving things more than living things. However, this finding is difficult to interpret because it may be an artifact of familiarity effects. For this reason, further subsets of items were selected post hoc to be familiarity matched with the subcategories of musical instruments and body parts (see Appendix B). J.B.R.'s naming of these additional subcategories is shown in Table 2. The subcategory of musical instruments was familiarity matched with subsets of animals, vegetables, vehicles, clothes, and foods. J.B.R.'s naming of musical instruments was not significantly different from his naming of any of these subsets apart from foods, \( \chi^2(1, N = 20) = 4.27, p < .05 \). Hence, J.B.R.'s naming of musical instruments appears to be intermediate between his naming of living and other nonliving things. A subset of 12 body parts were familiarity matched with subsets of fruits, vegetables, household tools, clothes, furniture, and foods. J.B.R.'s naming of the subset of body parts was found to be significantly different only from his naming of the subset of vegetables, \( \chi^2(1, N = 24) = 4.29, p < .05 \). This finding lends some support to the view that the subcategory of body parts tends to cluster with subcategories of nonliving rather than living things.

<table>
<thead>
<tr>
<th>Musical instruments and matched subsets</th>
<th>% correct</th>
<th>Body parts and matched subsets</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musical instruments</td>
<td>50</td>
<td>Body parts</td>
<td>67</td>
</tr>
<tr>
<td>Animals</td>
<td>20</td>
<td>Fruit</td>
<td>25</td>
</tr>
<tr>
<td>Vegetables</td>
<td>10</td>
<td>Vegetables</td>
<td>17</td>
</tr>
<tr>
<td>Food</td>
<td>0</td>
<td>Food</td>
<td>50</td>
</tr>
<tr>
<td>Vehicles</td>
<td>60</td>
<td>Tools</td>
<td>67</td>
</tr>
<tr>
<td>Clothes</td>
<td>70</td>
<td>Clothes</td>
<td>92</td>
</tr>
</tbody>
</table>

*Subcategory n = 10.  Subcategory n = 12.
Determining the Locus of J.B.R.'s Category-Specific Naming Deficit

A final point to consider is that picture-naming deficits do not necessarily reflect central semantic impairments and may instead be the result of impairments at a more peripheral level (such as in accessing semantic information from vision or in accessing output phonology from semantics). Performance on picture-naming tasks will not, therefore, necessarily reflect the organization of semantic memory. For this reason, a number of further tests were developed from the CSL picture set to establish whether J.B.R.'s difficulties with living things extend to tasks that do not require naming responses and to tasks that do not involve pictorial input and, hence, provide further information about the locus of J.B.R.'s semantic deficit.

Word–Picture Matching

The stimuli for this task were 96 picture arrays, each of which consisted of 4 pictures arranged in a square. The four pictures consisted of a target item, an item from the same semantic subcategory as the target, and two distractors from the opposite domain to the target (which came from the same subcategory as one another). For example, the target kangaroo appeared with the within-category distractor bear and the two across-domain distractors tricycle and skateboard. The targets consisted of 12 items from each of four subcategories of living things (animals, insects, fruits, and vegetables) and four subcategories of nonliving things (vehicles, toys, household tools, and clothing). The distractors in each array were familiarity matched with the target. In addition, the target items for each living category were familiarity matched with the target items from one of the nonliving categories to allow performance on the two domains to be compared. The positions of the targets and the distractors in the arrays were counterbalanced across trials and across subcategories. J.B.R. was asked to match a spoken word to one of the four pictures.

J.B.R. scored a total of 32 of 48 (67%) correct for living things and 43 of 48 (90%) for nonliving things. This difference was statistically significant, $\chi^2(1, N = 96) = 6.10, p < .02$. Of the 16 incorrect choices that J.B.R. made for living things, 15 (94%) were within-category distractors and 1 (6%) was an across-domain distractor. For nonliving things, two of five (40%) of his errors were within-category distractors and three of five (60%) were across-domain distractors. Nine older control participants did not make any errors on this task.

Picture Sorting

The stimuli for this task consisted of the target items from six of the subcategories of living and nonliving things that were used in the word–picture matching test (animals, fruits, and vegetables for the living things and vehicles, household tools, and clothing for the nonliving things). J.B.R. was asked to sort pictures of these items into two piles—one containing living things and one containing nonliving things. He scored 21 of 36 (58%) for living things and 34 of 36 (94%) for nonliving things. This difference was highly significant, $\chi^2(1, N = 72) = 11.09, p < .001$.

Word Sorting

The stimuli for this task consisted of the written names of the items used in the picture-sorting task. Again, J.B.R. was asked to sort these words into two piles—one of living and one of nonliving things. He scored 19 of 36 (53%) for living things and 35 of 36 (97%) for nonliving things. Once again, this difference was highly significant, $\chi^2(1, N = 72) = 16.67, p < .001$.

It can be seen from these results that J.B.R.'s deficit for living things does not only affect picture naming but also affects a word–picture matching task, a purely verbal task (word sorting), and a nonverbal task (picture sorting). These results suggest that his category-specific deficit for living things is indeed the result of an impairment at the semantic level.

General Discussion

An examination of J.B.R.'s picture naming has demonstrated that theoretically interesting results may be obtained using the CSL picture set in a case where results obtained with other stimulus sets (e.g., Funnell & Sheridan, 1992; Snodgrass & Vanderwart, 1980) are difficult to interpret.

First, the CSL picture set has been used to demonstrate that category effects can be observed even when stimuli are matched for familiarity and visual complexity. The present study therefore provides an extension to Warrington and Shallice's (1984) original description of J.B.R., which has been criticized on the basis that their stimuli were not familiarity matched. Furthermore, the effect of category was observed across all levels of familiarity, other than where J.B.R.'s performance was at floor. These results allow us to reject the hypothesis that category-specific deficits are an artifact of familiarity effects and hence do not reveal anything novel about the way in which semantic knowledge is organized (Funnell & Sheridan, 1992; Funnell, 1996; Funnell & De Mornay Davies, 1996).

However, the fact that the effects of concept familiarity cannot explain the selective impairment demonstrated by J.B.R. cannot necessarily be generalized to other cases of apparently category-specific deficits. There was a strong overall effect of familiarity on J.B.R.'s performance, and hence it is possible that some cases of apparently category-specific impairment may indeed be explicable in terms of this variable. The importance of using familiarity-matched sets when making comparisons between categories is therefore been underlined.

It was noted previously that Warrington and Shallice (1984) have suggested that deficits for living things may be the result of impaired sensory knowledge and that this hypothesis was based on J.B.R.'s atypical performance on
certain subcategories of living and nonliving things. However, it has been argued that these apparent anomalies may simply reflect the atypical familiarity of these subcategories (Funnell & Sheridan, 1992). Again, the CSL picture set gave us the opportunity to extend Warrington and Shallice’s findings by allowing us to make comparisons between subcategories that were matched for familiarity.

J.B.R.’s naming of the CSL set provided some evidence to support the hypothesis that his category-specific deficit cannot be explained in terms of a simple distinction between either living and nonliving things or functions versus perceptual knowledge. First, his naming of most living things (whose identification would appear to be particularly dependent on sensory information) was more impaired than his naming of most nonliving things. In addition, we replicated the finding that his naming of body parts (which Warrington and Shallice, 1984, have argued are reliant on functional information) clusters with his performance on nonliving rather than on living things. Furthermore, his naming of foods tended to cluster with his performance on living things, despite the fact that this subcategory crosses the boundary between natural kinds and artifacts. Results for the category of musical instruments were equivocal. His naming of this subcategory was not significantly different from his naming of other subcategories of either living and nonliving things, suggesting that it was at an intermediate level. This finding is not easily accounted for in terms of either the living–nonliving or sensory–functional distinctions. J.B.R.’s intermediate performance on musical instruments suggests that this subcategory’s reliance on sensory rather than functional information may be less extreme than initially proposed by Warrington and Shallice. Although the functions of various musical instruments may be less unique than those of other artifacts (in that they share the basic function of making music), these functions may nevertheless be more defined than those of most living things. This account is, of course, post hoc and is only speculative at present. We are currently in the process of collecting norms to determine the most important features for different subcategories of living and nonliving things to produce independent evidence that will reduce the need for speculation of this kind.

Overall, therefore, the CSL picture set has proved useful in examining two theoretical issues of considerable current interest that could not be addressed satisfactorily with standard stimuli. First, J.B.R. has been shown to demonstrate a category-specific semantic deficit that cannot be explained in terms of the effects of concept familiarity alone and that is not limited to low-familiarity items. Furthermore, there is some evidence that the subcategories that are most severely affected are indeed those whose identification would appear to be more reliant on sensory than functional information. We anticipate that well-controlled picture sets of this kind will prove crucial to the study of other neurological patients and to the investigation of further theoretical issues in the future.

References


## Appendix B

### Properties of Subsets Matched Post Hoc With Musical Instruments and Body Parts

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Musical instruments and matched subsets</th>
<th>Body parts and matched subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 10)</td>
<td>(n = 12)</td>
</tr>
<tr>
<td>Musical instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fam.</td>
<td>Com.</td>
</tr>
<tr>
<td>M</td>
<td>3.01</td>
<td>2.74</td>
</tr>
<tr>
<td>SD</td>
<td>0.36</td>
<td>0.64</td>
</tr>
<tr>
<td>Animals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.06</td>
<td>3.19</td>
</tr>
<tr>
<td>SD</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>Vegetables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.19</td>
<td>2.35</td>
</tr>
<tr>
<td>SD</td>
<td>0.46</td>
<td>0.36</td>
</tr>
<tr>
<td>Foods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.23</td>
<td>2.69</td>
</tr>
<tr>
<td>SD</td>
<td>0.26</td>
<td>0.59</td>
</tr>
<tr>
<td>Vehicles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.10</td>
<td>3.18</td>
</tr>
<tr>
<td>SD</td>
<td>0.40</td>
<td>0.45</td>
</tr>
<tr>
<td>Clothes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.15</td>
<td>2.37</td>
</tr>
<tr>
<td>SD</td>
<td>0.50</td>
<td>0.62</td>
</tr>
<tr>
<td>Furniture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4.14</td>
<td>2.13</td>
</tr>
</tbody>
</table>

\(^a\)Length in letters.  \(^b\)Percentage correct naming by young controls.

**Note.** Fam. = familiarity; Com. = visual complexity.