Journal of Experimental Psychology: General
2011, Vol. 140, No. 1, 14–34

0096-3445/10/$12.00 DOI: 10.1037/a0021446

The Representation of Abstract Words: Why Emotion Matters

Stavroula-Thaleia Kousta, Gabriella Vigliocco, David P. Vinson, and Mark Andrews
University College London

Elena Del Campo
Università degli Studi di Padova

Although much is known about the representation and processing of concrete concepts, knowledge of what abstract semantics might be is severely limited. In this article we first address the adequacy of the 2 dominant accounts (dual coding theory and the context availability model) put forward in order to explain representation and processing differences between concrete and abstract words. We find that neither proposal can account for experimental findings and that this is, at least partly, because abstract words are considered to be unrelated to experiential information in both of these accounts. We then address a particular type of experiential information, emotional content, and demonstrate that it plays a crucial role in the processing and representation of abstract concepts: Statistically, abstract words are more emotionally valenced than are concrete words, and this accounts for a residual latency advantage for abstract words, when variables such as imageability (a construct derived from dual coding theory) and rated context availability are held constant. We conclude with a discussion of our novel hypothesis for embodied abstract semantics.

Keywords: abstract and concrete knowledge, emotion, semantic representations, lexical processing

Concrete entities exist in space-time and are independent of human minds/language; abstract entities, on the other hand, do not exist in space-time but their existence depends on human minds/language (Hale, 1988). "Concreteness," therefore, indexes a basic ontological distinction, dividing entities into these two kinds. This ontological distinction is reflected in our epistemologies, and concreteness is arguably an organizing principle of semantic knowledge. Up to the present, research into semantic and conceptual representation has focused almost exclusively on how concrete word meanings and concepts are represented and processed, to the exclusion of abstract word meanings and concepts. However, the ability to communicate through language about abstract concepts, such as "courage," "dignity," and "revenge," lies at the heart of what it means to be human, and no theory of semantic or conceptual representation is complete without an explicit account of how abstract knowledge is acquired, represented, and processed.

In this article we first demonstrate, by combining experiments with large-scale regression analyses of data from the English Lexicon Project (ELP; Balota et al., 2007), that the dual coding theory and the context availability hypothesis—two of the most popular accounts of differences in representation and processing between concrete and abstract words—do not exhaustively account for processing (and hence representational) differences between the two types of word meanings. In fact, once imageability and context availability (along with a large number of other lexical and sublexical variables) are controlled, there is a residual advantage for abstract word processing. We show that this advantage can be explained by differences in emotional valence between concrete and abstract words, and we discuss a new hypothesis of how the semantic system is organized with respect to the distinction between concrete and abstract concepts. Specifically, we propose that both concrete and abstract concepts bind different types of information: experiential information (sensory, motor, and affective) and also linguistic information. However, concrete and abstract semantic representations differ in terms of whether sensory, motor, or affective information have the greatest weight, with sensory-motor information being more preponderant for concrete concepts and affective information playing a greater role for abstract concepts. Thus, a central and novel element of this proposal is the idea that experiential information contributes to the representation of both concrete and abstract words. However, whereas sensory-motor information is statistically more important for the representation of concrete words, emotional content, a largely neglected type of experiential information in the literature on semantic representation/processing, contributes to word representation and processing, particularly for abstract concepts.

This article was published Online First December 20, 2010.

Stavroula-Thaleia Kousta and Mark Andrews, Research Department of Cognitive, Perceptual, and Brain Sciences, University College London, London, England; Gabriella Vigliocco and David P. Vinson, Research Department of Cognitive, Perceptual, and Brain Sciences and Deafness Cognition and Language Centre, University College London; Elena Del Campo, Dipartimento di Psicologia Generale, Università degli Studi di Padova, Padova, Italy.

This research was supported by an EU (FP6-2004-NEST-PATH) grant (028714) and by an ESRC grant (RES-062-23-2012) to Gabriella Vigliocco. We are grateful to Pasquale Della Rosa for his help with selecting stimuli for Experiment 1. Many thanks to Allan Paivio for a stimulating discussion on our results and on dual coding theory.

Correspondence concerning this article should be addressed to Gabriella Vigliocco, Research Department of Cognitive, Perceptual, and Brain Sciences, University College London, London WC1H 0AP, England. E-mail: g.vigliocco@ucl.ac.uk

The Concreteness Effect: Dual Coding Theory and the Context Availability Model

It has been demonstrated repeatedly, and with a variety of methodologies, that concrete words have a cognitive advantage over abstract words—an advantage labeled *the concreteness effect*. With respect to lexical processing, early demonstrations of a processing advantage for concrete over abstract words were provided by James (1975), Whaley (1978), and Rubin (1980). James showed that at least when low-frequency words are considered, concrete words are identified as words faster than abstract words are. Whaley and Rubin adopted a correlational approach, showing that there is a significant negative correlation between concreteness ratings and lexical decision reaction times (RTs) for the same items. This processing advantage has since then been replicated in both lexical decision (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Bleasdale, 1987; de Groot, 1989; Howell & Bryden, 1987; Kroll & Merves, 1986; Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Stowe, 1989) and word-naming (de Groot, 1989; Schwanenflugel & Stowe, 1989) tasks.

With respect to memory for concrete and abstract words, it has been again repeatedly demonstrated that concrete words have an advantage over abstract words in both long-term and short-term memory tasks—for example, paired-associate learning (Paivio, Yuille, & Smythe, 1966), serial recall (Allen & Hulme, 2006; Romani, McAlpine, & Martin, 2007; Walker & Hulme, 1999), free recall (Romani et al, 2007; Ter Doest & Semin, 2005), reconstruction of order (Neath, 1997), and recognition memory (Fliessbach, Weis, Klaver, Elger, & Weber, 2006).

Among the handful of proposals that have been put forward to explain the "concreteness effect," two have been particularly influential: dual coding theory (Paivio, 1971, 1986, 1991, 2007) and the context availability model (Schwanenflugel, 1991; Schwanenflugel & Shoben, 1983). In both of these accounts, concrete word representations are assumed to be richer than abstract word representations (see also Plaut & Shallice, 1993). According to dual coding theory, concrete words are represented in two representationally distinct but functionally related systems: a verbal, linguistic system and a nonverbal, imagistic system. Abstract concepts, on the other hand, are primarily or exclusively represented in the verbal system. The cognitive advantage for words referring to concrete concepts is attributed to the fact that they have access to information from multiple systems. According to the context availability model, both concrete and abstract concepts are represented in a single verbal code, and neither the representations nor the processes that operate on these representations differ for the two types of concepts. The argument here is that comprehension relies on verbal context (supplied by either the discourse or the comprehender's own semantic memory) in order to be effective. Accessing the meaning of a word involves accessing a network of associated semantic information, and the advantage for concrete words arises because they have stronger and denser associations to contextual knowledge than do abstract words. These two proposals have guided research on concrete/abstract semantics; results, however, have been inconclusive. The majority of recent work is neuroscientific in nature, employing either electrophysiological or neuroimaging techniques in order to determine the neural bases of the distinction between concrete and abstract words.

A series of studies using event-related potentials (ERPs) suggested combining dual coding theory and the context availability model in explaining the concreteness effect (i.e., context-extended dual coding theory; Holcomb, Kounios, Anderson, & West, 1999; West & Holcomb, 2000). ERP studies have identified two components associated with concreteness: the N400 and a late negative component peaking around 700–800 ms. With respect to the first component, all relevant studies have found that concrete words elicit a larger N400 than do abstract words (Holcomb et al., 1999; Kanske & Kotz, 2007; Kounios & Holcomb, 1994; Nittono, Suehiro, & Hori, 2002; van Schie, Wijers, Mars, Benjamin, & Stowe, 2005; West & Holcomb, 2000). The observation that the effect has an anterior maximum but is widely distributed across the scalp (West & Holcomb, 2000) and the failure to find any structural overlap between concreteness and visual object working memory on that component (van Schie et al., 2005) have led to the suggestion that the effect arises within a verbal semantic system that is common to both concrete and abstract words. This N400 component has been argued to reflect postlexical processing in a semantic memory system, possibly involving the integration of semantic information into higher level representations (Osterhout & Holcomb, 1995). According to the context availability model, concrete words are assumed to have stronger and denser interconnections with other concepts in semantic memory than do abstract words (Schwanenflugel & Shoben, 1983). In the electroencephalograph (EEG) literature, concrete words are said to activate the semantic network more extensively than do abstract words, and this extensive activation is reflected in an amplified N400 for concrete words. The second, later, component is assumed to reflect the contribution of mental imagery for concrete words: It is more sustained over time, peaking at around 700–800 ms poststimulus. It is said to be associated with the retrieval of mental imagery associated with concrete words and thus to be consistent with dual coding theory. Although the imagery-related component is consistent with dual coding claims that imagery has a late effect in processing, the greater N400 amplitude for concrete words is harder to reconcile with context availability claims. The stronger interconnections in semantic memory for concrete words according to the model lead to facilitated integration of information. The increased N400 amplitude for concrete words, however, has been interpreted as indexing difficulty in integrating appropriate information (see Kutas, Van Petten & Kluender, 2006, for a review). So the extent to which EEG data actually support the context-extended dual coding theory is questionable.

A case for a qualitative difference between concrete and abstract word meanings, thus compatible with dual-coding views, comes from neuropsychological studies in which a double dissociation between concrete and abstract words has been observed. Although cases in which concrete words are better preserved in the damaged/aging brain are the most frequently reported (see e.g., Coltheart, Patterson, & Marshall, 1980; Franklin, Howard, & Patterson, 1995; Katz & Goodglass, 1990; Martin & Saffran, 1992; Roeltgen, Sevush, & Heilman, 1983; Warrington, 1975), there are cases reporting better performance on abstract over concrete words (see e.g., Breedin, Saffran, & Coslett, 1994; Cipolotti & Warrington, 1995; Marshall, Pring, & Robson, 1996; Papagno, Capasso, Zerboni, & Miceli, 2007; Reilly, Peelle, & Grossman, 2007; Sirigu, Duhamel, & Poncet, 1991; Warrington, 1975; Warrington & Shallice, 1984).

In the imaging literature, although abstract word processing seems to involve activations in a more distributed network of brain regions than concrete word processing (Pexman, Hargreaves, Edwards, Henry, & Goodyear, 2007), there is converging evidence that abstract word processing is associated with higher activation in left hemispheric areas that are known to be involved in semantic processing, for example, the left inferior frontal gyrus (LIFG; Binder et al., 2005; Fiebach & Friederici, 2004; Jessen et al., 2000; Noppeney & Price, 2004; Perani et al., 1999) and the superior temporolateral cortex (Binder, Desai, Graves, & Conant, 2009; Binder et al., 2005; Kiehl et al., 1999; Mellet, Tzourio, Denis, & Mazoyer, 1998; Wise et al., 2000). With respect to greater activation for abstract over concrete words in the LIFG, this finding has been interpreted as indicating more effortful retrieval of semantic information for abstract words, a finding that has been interpreted in some studies as consistent with context availability predictions. Again, however, the majority of the studies use items matched on frequency but not on familiarity or other relevant variables. For instance, in one of the otherwise best controlled studies in the imaging literature (Binder et al., 2005), although items were matched on frequency, we found that concrete words were significantly more familiar than abstract words, with average familiarity ratings of 534 and 471, respectively, $t(98) = 3.956$, $p < .001$. It may well be that such differences in familiarity underlie some of the effects reported in the neuroimaging literature.

When concrete words are compared with abstract words, results have been extremely variable. Although some studies have found activations of left hemispheric regions associated with higher levels of visual processing such as the left fusiform gyrus (D'Esposito et al., 1997; Fiebach & Friederici, 2004; Mellet et al., 1998; Sabsevitz, Medler, Seidenberg, & Binder, 2005), consistent with the dual coding prediction that concrete word meanings activate relevant imagistic information, a number of studies have failed to find any regions at all that are activated more during concrete word processing (Friederici, Opitz, & von Cramon, 2000; Grossman et al., 2000; Kiehl et al., 1999; Krause et al., 1999; Noppeney & Price, 2004; Perani et al., 1999; Pexman et al., 2007; Tyler, Russell, Fadili, & Moss, 2001). Some studies have found more bilateral activations during concrete word processing (Binder et al., 2005; Sabsevitz et al., 2005), whereas other studies have shown that there is no right-hemisphere involvement in the processing of concrete words and that, if anything, there are more right-hemispheric activations for abstract rather than concrete words (see Fiebach & Friederici, 2004, for a review). One of the reasons for the lack of consistency in the results may be that the concrete words used within and across studies differ in terms of their featural composition, which quite reasonably leads to activation of different brain networks in different studies or to lack of consistent areas of activation within the same study. Thus, just as with the behavioral and EEG evidence reviewed earlier, imaging studies do not provide clear support for either dual coding or context availability calling for new theoretical directions and further empirical investigation.

The Concreteness Effect: Testing Dual Coding Theory and the Context Availability Hypothesis

In the literature, it is invariably assumed that the psycholinguistic constructs of concreteness and imageability tap into the same underlying theoretical construct (i.e., the ontological distinction between concrete, spatiotemporally bound concepts and abstract, nonspatiotemporally bound concepts). After all, when nothing else is taken into account, imageability ratings explain more than 72% of the variance in concreteness ratings, and up to now the variance that is not explained by imageability has been considered to be pure noise, due perhaps to the imprecise nature of subjective norms. This general assumption is illustrated in the following quote:

> Although imageability and concreteness are technically different psycholinguistic constructs, the correlation between these variables is so strong that many authors use the terms interchangeably. Here we make the same assumption of synonymy between imageability and concreteness in terms of theory (i.e., concreteness effects = imageability effects). (Reilly & Kean, 2007, p. 158)

In fact, concreteness and imageability ratings have been used interchangeably in most of the recent literature in the field (see e.g., Binder et al., 2005; Fliessbach et al., 2006; Giesbrecht, Camblin, & Swaab, 2004; Richardson, 2003). However, concreteness and imageability tap into, at least partially, different aspects of semantic representations if native speaker intuitions about them are taken seriously: Our analyses of ratings for more than 4,000 words in the MRC Psycholinguistic Database show that the frequency distribution of concreteness ratings is bimodal, with two distinct modes for abstract and concrete words (see also Cartwright & Nickerson, 1979; Nelson & Schreiber, 1992), whereas the distribution of imageability ratings is unimodal (see Figure 1). In other words, concreteness ratings capture the categorical ontological distinction between concrete and abstract words (and their underlying conceptual representations), whereas imageability ratings index a graded property that is meant to capture the differential association of words with sensory (primarily visual) properties.

Moreover, from a theoretical point of view, imageability ratings are a proxy for concreteness only in the dual coding theory and not, for example, in the context availability hypothesis, because only the former explains differences between concrete and abstract words in terms of whether (and to what extent) the nonverbal imagistic system is engaged (Fliessbach et al., 2006; Reilly & Kean, 2007). According to the context availability hypothesis, however, imageability would not exhaust the differences between concrete and abstract words—differences that, instead, arise as a consequence of different degrees of richness of semantic representation within a verbal system.

One approach to testing both hypotheses is to manipulate concreteness while controlling for both imageability and context availability. Both dual coding and context availability theories predict that concreteness effects will not be observed under these conditions.

Experiment 1

In this experiment, we contrast morphologically simple abstract and concrete words that have been matched for imageability and context availability (as well as a host of other noise variables).

Method.

Participants. Fifty-eight native English speakers (32 female; mean age = 28.69 ± 9.96 years) participated and were paid at a rate of £6 ($12) per hour. Three participants were replaced because of a high number of timed-out responses in their data.

Materials and design. Forty concrete and 40 abstract monomorphemic words were selected (the full item list appears in

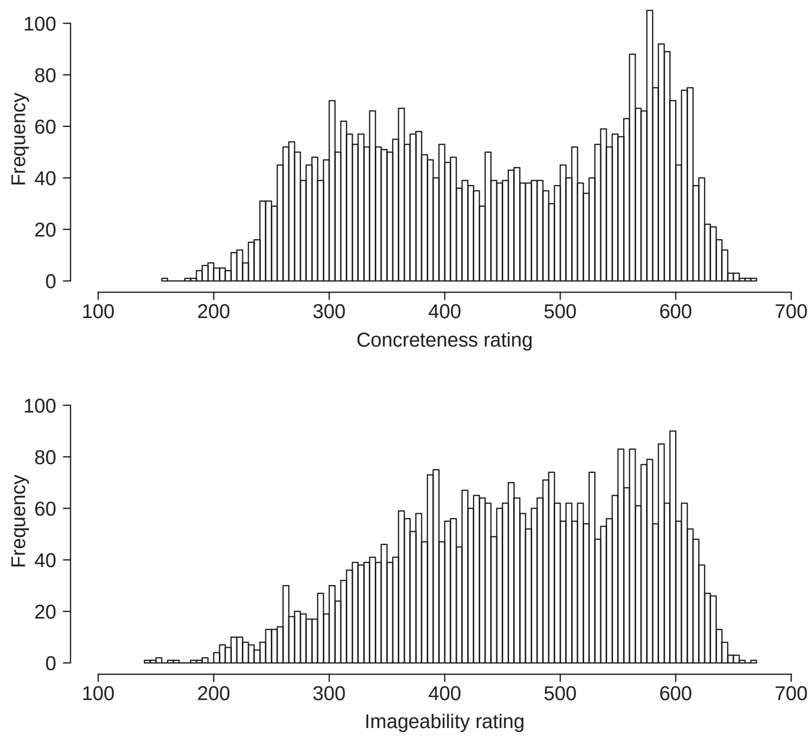

Figure 1. Frequency plots for concreteness and imageability ratings for 4,274 words from the MRC Psycholinguistic Database. Using the dip test (Hartigan & Hartigan, 1985), we rejected the hypothesis of unimodality for the concreteness distribution (dip = .0244, $p < .001$) but not for the imageability distribution (dip = .0058, *ns*).

Appendix A). The items differed on concreteness but were matched pairwise on 12 lexical and sublexical variables, including rated context availability (see Table 1). Imageability, familiarity, and age of acquisition ratings were obtained from the MRC Psycholinguistic Database (Coltheart, 1981). Items were also matched in length (in number of letters, phonemes, and syllables) and number of meanings (in terms of number of synsets in which a word appears in WordNet; Fellbaum, 1998). Frequency, orthographic neighborhood density, mean frequency of orthographic neighbors, and mean positional bigram frequency were taken from the ELP (Balota et al., 2007). Finally, we obtained context availability ratings by asking 47 native English speakers to rate words on a 7-point Likert scale according to how easy it is to come up with a particular context or circumstance in which they might appear. The instructions to participants were identical to those used by Schwanenflugel and Shoben (1983), with the exception of some of the examples given in order to anchor the ratings, which differed between the studies. We obtained these norms for 650 words (each word was rated by 22 or 25 speakers; see Table 1 for details and Appendix A for a full list of items used in the experiment).

Table 1
Item Averages (and Standard Deviations) for Experiment 1

Variable	Abstract	Concrete
Concreteness	345 (40)	552 (44)
Context availability	568 (46)	566 (52)
Imageability	500 (42)	505 (35)
Familiarity	504 (70)	505 (67)
Age of acquisition	385 (40)	390 (103)
Log frequency	9.02 (1.44)	9.03 (1.62)
Number of letters	5.55 (1.20)	5.63 (1.28)
Number of phonemes	4.71 (1.33)	4.55 (1.27)
Number of syllables	1.68 (0.57)	1.68 (0.70)
Mean positional bigram frequency	1,491 (959)	1,595 (943)
Number of orthographic neighbors	2.63 (3.90)	2.84 (4.02)
Mean neighbor frequency	4.86 (3.93)	4.26 (4.13)
Number of synsets	5.16 (3.25)	6.50 (6.86)

Note. The numbers reported here are based on 38 items per condition (two were excluded on the basis of low accuracy; see Results section).

We also selected 40 concrete and 40 abstract words matched with the experimental items in terms of concreteness to serve as the basis for creating pseudowords for the experiment. The pseudowords were created by altering a single letter in each of these words. We made an effort to select pseudowords with only one orthographic neighbor (the intended real word). In cases in which that was not possible (for all three-letter and some of the four-letter words), the intended word was the most frequent among the set of orthographic neighbors of the nonword. The resulting pseudowords were matched pairwise with the experimental items in terms of length and mean positional bigram frequency. Eighty additional words and 80 nonwords were also included as fillers.

Procedure. Participants were tested individually. Each trial began with a fixation cross presented in the middle of the screen for 400 ms, followed by presentation of the string for 2,000 ms or until a response was given (whichever was earlier). Participants were instructed to respond as quickly and accurately as possible using a serial response box. After response or time-out, the screen went blank and participants were instructed to press the space bar to continue with the next trial. Ten practice items were first presented, followed by the 320 words and nonwords presented in a different random order for each participant.

Results. In the analysis of RTs, we excluded all responses faster than 200 ms and slower than 2,000 ms (0.84% of the data). For two concrete words, accuracy rates did not differ from chance. We excluded these items as well as their paired abstract items from further analysis. We also removed outliers by excluding from analysis RTs 2.5 standard deviations above the mean per condition for each participant (2.04% of the data). Mean latencies can be found in Table 2. The analyses reported were carried out on correct responses only.

Abstract words were recognized as words faster than concrete words ($M_{\text{abstract}} = 568$ ms; $M_{\text{concrete}} = 590$ ms). This difference was significant both by participants, $F_1(1, 57) = 23.327$, $p < .001$, and by items, $F_2(1, 37) = 5.447$, $p < .05$. In the analysis of accuracy, there was a numerical advantage for abstract over concrete words ($M_{\text{abstract}} = 96.59\%$; $M_{\text{concrete}} = 95.48\%$), but the effect was not statistically reliable, $F_1(1, 57) = 3.166$, $p = .08$; $F_2 < 1$.

Discussion. In this experiment we found that abstract words were processed faster than concrete words. This finding forces us to reject the dual coding and context availability hypotheses as stated earlier, because we found differences between the concrete and abstract conditions that were matched for imageability and context availability. In order to further assess the generalizability of the effect, given that in contrast to previous work we found an advantage for abstract words, in the next section we report the results of large-scale regression analyses on lexical decision data from the ELP.

Regression Analyses 1 (903 words). In this set of analyses, we used context availability norms from Clark and Paivio (2004), who collected ratings for 925 words. We also included concreteness, imageability, and a number of variables that have been identified as relevant for visual word recognition: number of letters, mean positional bigram frequency, orthographic neighborhood density (the latter two from the ELP), number of morphemes, log frequency (based on the Hyperspace Analogue to Language (HAL) frequency counts as reported in the ELP), and age of acquisition (from the merged Bristol and MRC norms; Stadthagen-Gonzalez & Davis, 2006). We also coded each word according to whether its grammatical class is ambiguous. We did not include familiarity in this analysis due to the high correlation with context availability (.93 and .80 for the two sets of familiarity ratings reported in Clark & Paivio, 2004).

Table 2
Reaction Times and Accuracy Rates (and Standard Deviations) for Experiment 1

Variable	High concreteness	Low concreteness	Nonwords
Reaction time (in ms)	590 (99)	568 (88)	682 (140)
Accuracy rate (%)	95.48 (4.92)	96.59 (3.66)	94.19 (4.78)

The analyses reported next were carried out on 903 items for which lexical decision RTs (averaged across multiple participants) and accuracy data were available in the ELP. We tested whether concreteness explains any of the variance in the data after the effects of imageability, context availability, and other lexical and sublexical variables are removed. For the RT analyses, we logarithmically transformed the by-item mean RTs and then fitted an ordinary least squares linear regression model on the transformed data. For the accuracy analysis, we used maximum likelihood estimation models.

Latencies. In this and all subsequent regression analyses on latencies, the procedure was as follows: We first fitted a linear regression model including all the predictors. In all models, the relationship between latencies and the predictors was modeled as a linear combination of the relevant correlation coefficients. When fitting the linear model, we relaxed the assumption of linearity when considering the relationship between each individual predictor and the dependent variable. In order to model nonlinear relationships between individual predictors and the dependent variable, we used restricted cubic splines (Harrell, 2001). Cubic splines are piecewise polynomials used in curve fitting such that the relationship between a predictor and a dependent variable (X) is modeled by placing polynomials within intervals of X and connecting the polynomials across different intervals of X (Harrell, 2001). These intervals are called knots, and in our analyses we used the minimum number of knots necessary in order to model nonlinearities. Restricted cubic splines are spline functions that are constrained to be linear at the tails (i.e., before the first and after the last knot); the motivation for constraining the function to be linear at the tails is that cubic splines provide poor fit at the tails.

After fitting the initial model, we removed outliers (following the procedure in Baayen, Feldman, & Schreuder, 2006). We then refitted the model and used a bootstrap validation procedure (Harrell, 2001) to determine to what extent our model overfitted the data. We included a fast backward elimination algorithm in the validation procedure to eliminate nonexplanatory variables. We then refitted the model, excluding nonexplanatory variables. The results we report are from this final refitted model.

Of the total data, 3.10% were removed as outliers. Model optimism (an estimate of the degree of overfitting) was low (0.29%). Although context availability had a significant facilitatory effect on latencies, $F(1, 863) = 10.30$, $p < .01$, concreteness continued to have a significant inhibitory effect, $F(1, 863) = 5.51$, $p < .05$; final model $R^2 = .717$. The effect of imageability was also significant, $F(1, 863) = 8.72$, $p < .01$. The partial effects of these predictors are plotted in Figure 1.

Accuracy rates. All three variables of interest predicted probability of a correct response, consistent with the response latency data: context availability: $\chi^2(1) = 105.13$, $p < .001$; concreteness: $\chi^2(1) = 21.08$, $p < .001$; imageability: $\chi^2(1) = 12.82$, $p < .001$ (see Figure 2 for the partial effects).

The results of these analyses show, with a much larger set of items than the one used in Experiment 1, that concreteness has a small but significant effect on latencies and accuracy rates to the advantage of abstract words, when imageability and context availability are partialed out.

In the following regression analyses we provide a final test of dual coding, assessing the generalizability of these results for an

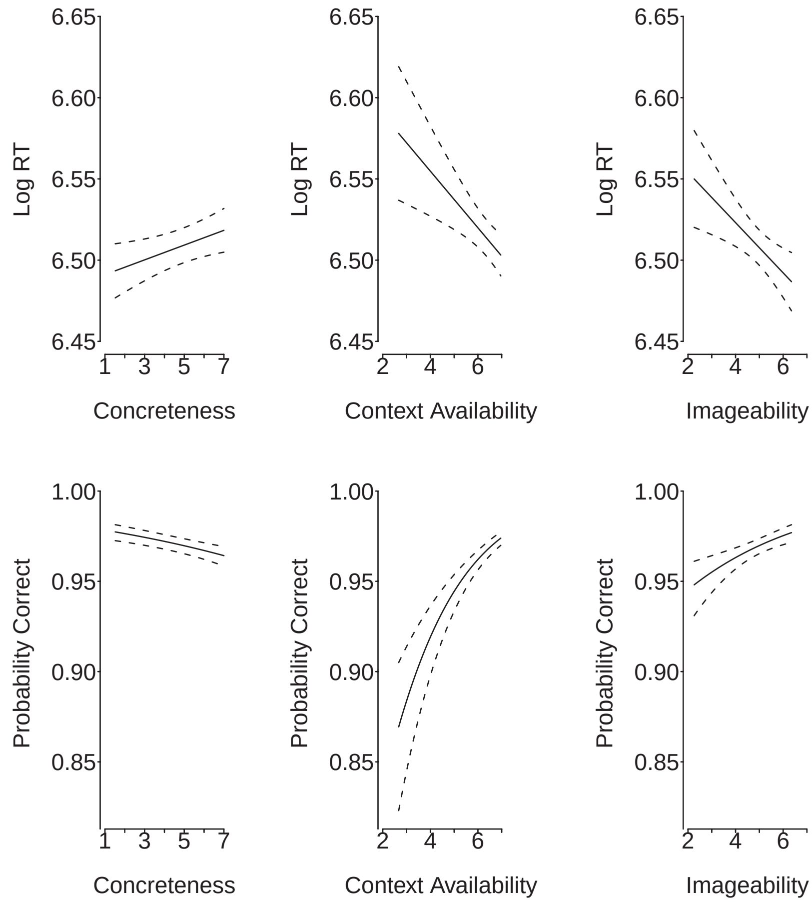

Figure 2. Plots of the partial effects of concreteness, context availability, and imageability in Regression Analyses 1 (upper panels indicate log reaction times [RTs]; lower panels indicate accuracy). Dashed lines represent 95% confidence intervals. The effects are adjusted to the median of all other continuous predictors and to class-ambiguous words.

even larger set of words from the ELP (achieved by leaving out context availability).

Regression Analyses 2 (2,330 words). We used the same predictors in these analyses as in Regression Analyses 1, excluding context availability. We also included familiarity (from the MRC Psycholinguistic Database) and part of speech. Again we fitted separate models for lexical decision latencies and accuracy.

Latencies. Of the total data, 1.63% were outliers and were removed. The validation procedure showed that model optimism was minimal (0.38%), confirming that our model was reliable. Both variables of interest were significant predictors of latencies: concreteness: $F(1, 2276) = 4.32$, $p < .05$; imageability: $F(2, 2276) = 3.75$, $p < .05$; nonlinear: $F(1, 2274) = 4.47$, $p < .05$; final model $R^2 = .6965$. Plots of the partial effects of the two predictors, which enable direct comparison of effect sizes, appear in Figure 3. For concreteness, slower response times are observed for the most concrete words, whereas for imageability, faster response times are observed for the most imageable words.

Accuracy. Both variables significantly predict correct responses: concreteness: $\chi^2(2) = 44.57$, $p < .001$; nonlinear: $\chi^2(1) = 11.64$, $p < .001$; imageability: $\chi^2(2) = 28.24$, $p < .001$; nonlinear: $\chi^2(1) = 18.54$, $p < .001$. Plots of the partial effects of these predictors can be found in Figure 3. Concreteness predicted greatest accuracy for abstract words, with a nonlinear decrease in accuracy with higher concreteness ratings. Imageability, on the other hand, predicted greatest accuracy for highly imageable words, whereas for low-imageability words (in the 200–400 range) the effect of imageability leveled off to similar accuracy rates.

Note that the zero-order correlation between concreteness and latency in this word set is negative ($r = -.22$), whereas in the earlier analyses we found a positive slope for concreteness. Cases in which the coefficient of a predictor variable reverses in sign when entered in a regression model have been associated with the phenomenon of enhancement (Bollen, 1989; Friedman & Wall, 2005; McFatter, 1979; Shrout & Bolger, 2002).[1] According to our hypothesis that differences in concreteness are not exhausted by differences in imageability, we assume that the coefficient for concreteness in this model represents the direct effect of concreteness, when the indirect effect of concreteness (through imageability) is held constant. Because the interpretation of such effects in

[1] Many thanks to Harald Baayen for bringing this phenomenon to our attention.

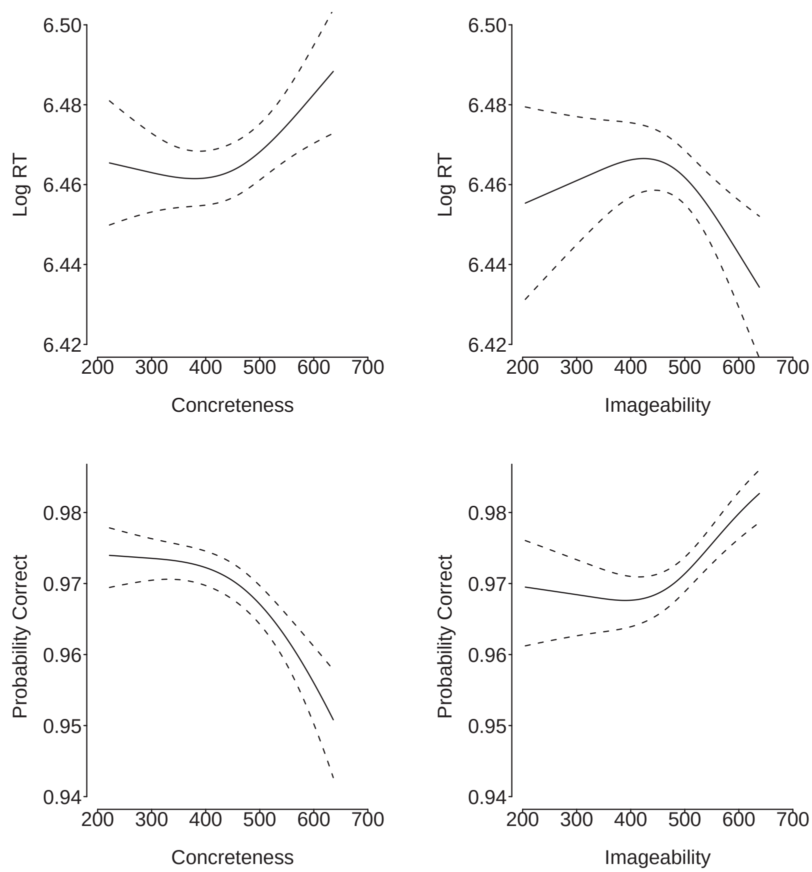

Figure 3. Plots of the partial effects of imageability and concreteness in Regression Analyses 2 (upper panels indicate log reaction times [RTs]; lower panels indicate accuracy). Dashed lines represent 95% confidence intervals. The effects are adjusted to the median of all other continuous predictors and to nouns and class-ambiguous words.

linear regression has been a matter of debate, in Appendix B we provide a detailed theoretical overview as well as a formal specification of the theoretical model that we assume underlies our data.

Thus, to summarize, we have presented here decisive evidence that neither dual coding nor context availability provides a full account of the representational and processing differences between concrete and abstract words.

The Abstractness Effect: The Role of Affect

Altarriba, Bauer, and Benvenuto (1999) were the first to note that affective association (and in particular the valence of words, namely whether words have positive, negative, or neutral connotation) may be confounded with, or rather interact with, concreteness. Altarriba et al. proposed that instead of treating concreteness as a dichotomous variable (concrete vs. abstract), it should be treated as a trichotomy (concrete vs. abstract vs. emotion words). This proposal was motivated by the finding that concrete words, abstract words, and words denoting emotional states consistently received different concreteness, imageability, and context availability ratings: Although emotion words were rated as more abstract than other abstract words, they were higher in imageability and context availability than other abstract words (and lower than concrete words).[2] On the basis of these findings, Altarriba et al. cautioned against including emotion words within the group of abstract words when concreteness effects are investigated, because their inclusion would be a confound. If we consider the items we used in Experiment 1 (reported in Appendix A), it appears indeed that our abstract words may have more affective associations than do the concrete words (although importantly, only a few refer directly to emotions, and the pattern of results does not change if they are removed from the analysis), leading to the possibility that the abstractness effect we observed there may be mediated by the confounding between concreteness and affective association (see also Altarriba & Bauer, 2004).

In previous related work (also combining carefully controlled experiments with regression analyses of ELP data), we have found that words with affective associations (regardless of whether they referred to positive or negative emotions) are processed faster than neutral words (Kousta, Vinson, & Vigliocco, 2009). Ratings of

[2] See also Altarriba and Bauer (2004, Experiment 1); in their free recall test, higher recall rates were observed for emotion words than for either concrete or abstract words.

affective association were obtained by merging the ANEW database (Bradley & Lang, 1999) with normative data we collected for an additional 1,200 words using the same instructions and procedure as in the original database (see Kousta et al., 2009, for details).[3] These findings by Kousta et al. (2009) are important because the processing advantage for words with affective associations provides a straightforward account of the abstractness effect: Abstract words have a processing advantage over concrete words because abstract words tend to be more emotionally loaded. Importantly, Kousta et al. showed that the processing advantage is not limited to words referring to emotions but also extends to other words with affective associations. Thus, these results suggest that affective association should be considered as a continuous variable spanning across words of all types (rather than a variable identifying the special category of emotion words, as originally hypothesized by Altarriba et al., 1999; but see also Altarriba, 2008, for a discussion concerning bilingualism). Note here that the processing advantage for both positive and negative words reported by Kousta et al. goes against evidence indicating that differences are observed between positive and negative words (e.g., advantage for positive words over negative words in immediate serial recall, greater for concrete words than for abstract words: Tse & Altarriba, 2009; or inhibition of negative words: e.g., Estes & Adelman, 2008). This discrepancy is addressed in Kousta et al., who show that in many instances, differences in results can be attributed to less stringent criteria for item selection or, as in the case of Estes and Adelman (2008), due to sampling differences for valence (see Kousta et al., 2009, pp. 474 and 478 for details).[4]

In order to make the link between abstract words and affective associations explicit, we first need evidence that abstract words tend to have more affective associations than do concrete words. Initial evidence in this direction is provided by Vigliocco et al. (2010), who showed that for a set of 1,446 words, valence ratings significantly predict concreteness ratings, even after imageability is taken into account. In other words, the more valenced a word is, the more abstract it tends to be, whereas the more neutral a word is, the more concrete it tends to be. In an fMRI study using items similar to those we used in Experiment 1, Vigliocco et al. further showed that for abstract words, ratings of affective association predicted modulation of BOLD signal in the rostral anterior cingulate cortex, an area associated with emotional processing on the basis of anatomical, physiological, and imaging results (see Bush, Luu, & Posner, 2000). Taken together, these findings provide the motivation for exploring abstractness effects in terms of words' emotional content.

The Abstractness Effect: Testing the Role of Affective Associations

In Experiments 2 and 3 detailed later, we directly test whether differences in affective associations can account for the abstractness effect. In Experiment 2 we used neutral words (arbitrarily defined as those words whose mean valence ratings ranged between 4.25 and 5.75 on a 9-point scale, where 1 = *negative*; 5 = *neutral*; and 9 = *positive*), but they spanned the entire range of the concreteness and imageability scales. In Experiment 3, we selected familiar words spanning the whole range of valence (and arousal) ratings, both concrete and abstract.

If the abstractness effect we observed in Experiment 1 and Regression Analyses 1 and 2 can be accounted for in terms of differences in affective associations between abstract and concrete words, then it should not be found in Experiment 2; and it should be present in Experiment 3 when affective associations (both valence and arousal) are not entered in the regression model but should be eliminated once affective associations are entered in the model.

Experiment 2

Method.

Participants. Forty-six undergraduate psychology students (30 female; mean age = 23.9) participated in the experiment and received monetary compensation of £10 ($20).

Materials and design. For this experiment, 774 words with valence ratings ranging between 4.25 and 5.75 ($M = 5.11 \pm 0.39$) were chosen.[5] Their concreteness ratings ranged from 217 to 646 ($M = 481 \pm 112$), imageability from 143 to 659 ($M = 475 \pm 106$), familiarity from 126 to 643 ($M = 476 \pm 88$), age of acquisition from 164 to 700 ($M = 413 \pm 124$), length from 3 to 14 letters ($M = 6.26 \pm 2.32$), and log frequency (HAL) from 2.71 to 12.99 ($M = 8.41 \pm 1.85$). We created 774 pronounceable nonwords by changing one letter in random position within real words. Pseudowords were matched to the experimental items in terms of length and bigram frequency (using WordGen; Duyck, Desmet, Verbeke, & Brysbaert, 2004). A single presentation list was generated for the experiment. The data were analyzed using linear regression models (the same procedures as in Regression Analyses 1 and 2, except that in Experiments 2 and 3 we did not average RTs across participants but instead analyzed trial-level data), including the following predictors: familiarity, length, log frequency, age of acquisition, orthographic neighbors, bigram frequency, part of speech, and number of morphemes in addition to concreteness, imageability, and valence.

Procedure. The procedure was the same as in Experiment 1.

Results.

Latencies. Analyses were conducted on only those 706 words for which age of acquisition data were available. We additionally excluded 35 words for which accuracy was less than 60% correct

[3] We prefer to talk about affective associations, namely, considering valence and arousal together, rather than distinguishing between the two here. This is because, first, in all of our studies conducted to date, valence has the larger effect, but arousal also has a modest role. Second, although there are neuroanatomical and theoretical reasons to distinguish between the two constructs (see e.g., Lewis, Critchley, Rotshtein, & Dolan, 2007), there is no clear rationale to expect dissociations between them for abstract words.

[4] Although the study by Tse and Altarriba (2009) is not discussed by Kousta et al. (2009), a similar point can be made because the authors did not control for age of acquisition, which varies by valence and concreteness (for concrete words, those that are positively valenced tend to be acquired earlier than are neutral or negative words; for abstract words, those that are positively or negatively valenced tend to be acquired earlier than are neutral words).

[5] Words on the more negative end of this interval include *golf, chop,* and *transfer*; words on the more positive end of this interval include *post, menu,* and *theme.*

across participants. Finally, we excluded from analyses all responses faster than 200 ms and slower than 2,000 ms, as well as RTs more than three standard deviations from the mean for each participant (2.4% of the data). An additional 2.35% of data points were excluded as outliers (as in Regression Analyses 1 and 2). In the analysis of RTs (using the predictors listed earlier), concreteness and imageability were not significant predictors: concreteness: $F(1, 29413) = 0.09$, $p = .76$; imageability: $F(1, 29413) = 1.37$, $p = .24$); final model $R^2 = .0954$.[6] Despite valence ratings' being limited to a restricted range (4.25–5.75), this predictor was significant, $F(1, 29413) = 3.85$, $p < .05$ (see Figure 4, upper panels).

Accuracy. Concreteness was not a significant predictor, $\chi^2(1) = 0.01$, $p = .91$, nor was imageability, $\chi^2(1) = 0.01$, $p = .94$, whereas valence was significant, $\chi^2(1) = 4.88$, $p < 05$ (see Figure 4, lower panels).

Discussion. No abstractness effect was observed in this experiment using neutral words, thus supporting the suggestion that the apparent abstractness advantage in Experiment 1 and Regression Analyses 1 and 2 is due to a confounding effect of valence. By reducing the range of valence in the item set (while retaining the full range of concreteness and imageability), we were able to eliminate any abstractness effect. A further, unexpected result from this study is the finding of a significant effect of valence even for the subtle extent of variation among neutral words, a result further underscoring the graded (rather than categorical) nature of valence effects.

Experiment 3

In order to seek converging evidence with the results of Experiment 2, in this final experiment we chose a set of items to cover the full range of emotional valence and arousal, allowing other variables to vary freely, following the same logic as earlier. Because so many items had to be excluded from Experiment 2 due to low accuracy, and because words rated low in imageability often also tend to be particularly low in familiarity, we were also more selective in choosing items—picking only words with average or high familiarity and that yielded accurate lexical decisions in the ELP.

Method.

Participants. Forty-seven undergraduate psychology students who were native English speakers (33 female; mean age = 20.34 ± 4.59) participated as part of a class requirement.

Materials and design. In this experiment, 480 words were chosen from the set of items for which valence, arousal, concreteness, age of acquisition, and other such variables were available, but any items with low familiarity (i.e., ratings below 350 on the 100–700 scale) or lexical decision accuracy (i.e., less than 70% correct in the ELP) were not included. We started by including 111 words from Kousta et al. (2009)—37 neutral, 37 positive, and 37 negative words closely matched for other lexical variables—plus 40 additional words that were randomly selected from the valence intervals not used in that study (i.e., 20 words from the gap between negative and neutral categories defined by Kousta et al. and 20 words from the gap between positive and neutral). Finally, 329 words were chosen randomly from the remaining set. Concreteness ratings of the 480 words ranged from 219 to 634 ($M = 459 \pm 115$); valence from 1.56 to 8.44 ($M = 5.21 \pm 1.46$); arousal from 2.67 to 7.67 ($M = 4.86 \pm 0.93$); imageability from 213 to 637 ($M = 488 \pm 95$); familiarity from 351 to 645 ($M = 506 \pm 66$); age of acquisition from 152 to 692 ($M = 389 \pm 112$); length from 3 to 14 letters ($M = 6.29 \pm 2.31$); and log frequency (HAL) from 2.77 to 12.47 ($M = 8.84 \pm 1.61$). The 480 pseudowords were created by selecting an unused word from the set, matched in valence and length to the actual words, and changing one letter (or two letters for source words longer than eight letters). The data were analyzed as in Experiment 2.

Procedure. The procedure was the same as in Experiments 1 and 2.

Results. We first excluded the data from one participant who did not complete the task and one who was less than 65% correct. All other participants were above 75% correct. We then excluded nine words for which average accuracy was less than 60% correct in this study (*prairie, herdsman, theologian, giver, nozzle, havoc, impediment, adherence, furnace*), leaving us with data from 45 participants and 471 words.

Latencies. We excluded from analyses responses faster than 200 ms and slower than 2,000 ms and RTs 2.5 standard deviations from the mean for each participant (1.44% of the data). A further 3.00% of data points were excluded as outliers following the same procedure used in previous analyses. In the analysis of RTs, concreteness and imageability were not significant predictors: concreteness: $F(1, 20790) = 2.40$, $p = .12$; imageability: $F(1, 20790) = 0.02$, $p = .89$); final model $R^2 = .0469$. Valence was significant, however: $F(2, 20790) = 3.60$, $p < .05$; nonlinear: $F(1, 20790) = 4.79$, $p < .05$ (see Figure 5, upper panels). Arousal also approached significance, $F(2, 20790) = 3.35$, $p = .067$.

Accuracy. Concreteness was not a significant predictor, $\chi^2(1) = 2.27$, $p = .13$, nor was imageability, $\chi^2(1) = 0.05$, $p = .82$, whereas valence was significant, $\chi^2(2) = 9.11$, $p < 05$; nonlinear: $\chi^2(1) = 6.97$, $p < 01$ (see Figure 5, lower panels). Again, arousal approached significance, $\chi^2(1) = 3.17$, $p = 075$.

Although the results of these analyses clearly show that emotional variables are significant predictors of RTs and accuracy in Experiment 3, whereas concreteness and imageability are not, this is not yet sufficient evidence to claim that the abstractness effects we observed in Experiment 1 and Regression Analyses 1 and 2 are actually the product of these emotional variables. Data from Experiment 3 offer the possibility to test this question directly. If the emotional variables account for the abstractnesss effect, then once the emotional variables (valence and arousal) are removed from the regression models of Experiment 3, there should be a significant advantage for abstract words. We reanalyzed the same data, with the same models, excluding valence and arousal. In this analysis, concreteness was a significant predictor of both RTs and accuracy: RTs: $F(1, 20793) = 13.76$, $p < .01$; accuracy: $\chi^2(1) = 12.38$, $p < 01$, indicating that more abstract words are faster and more accurate when emotional variables are not taken into account (see Figure 6).

Discussion. Experiment 3 shows that for a relatively large set of words spanning the entire range of concreteness and, crucially,

[6] The substantial decrease in amount of variance explained in Experiments 2 and 3 compared with Regression Analyses 1 and 2 arises because, in the regression analyses, RT data was averaged across participants. In the experiments the models were fit to trial-level data.

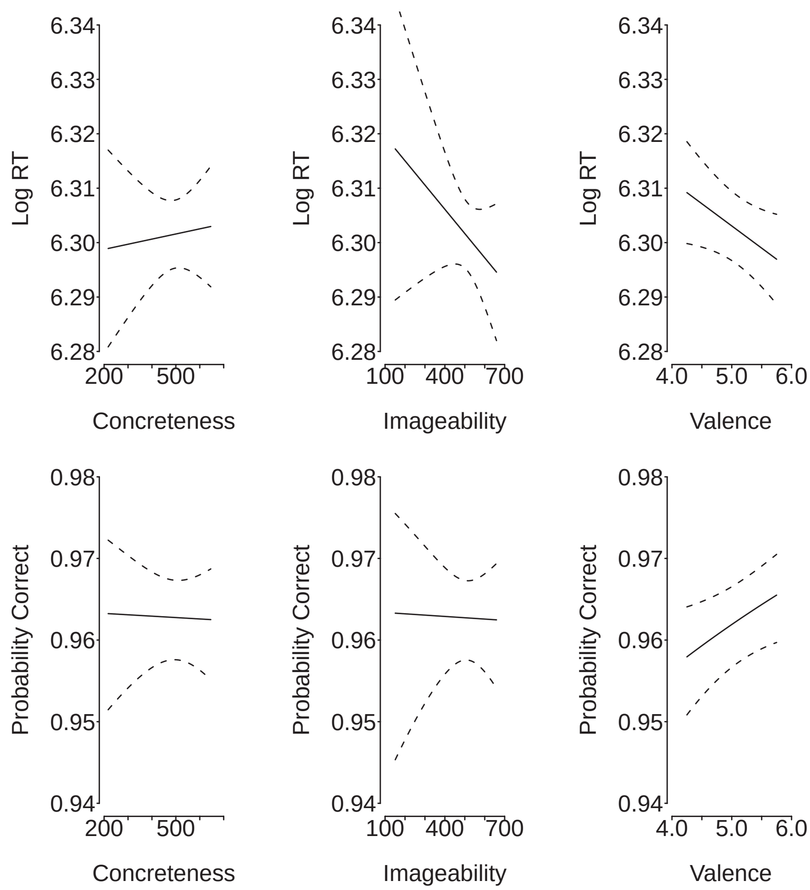

Figure 4. Plots of the partial effects of concreteness, imageability, and valence in Experiment 2 (upper panels indicate log reaction times [RTs]; lower panels indicate accuracy). Dashed lines represent 95% confidence intervals.

valence and arousal ratings, the abstractness effect is found in models that do not include ratings of affective associations, but the abstractness effect is eliminated in models that take affective associations into account, either by restricting the range of affective variables (as we did in Experiment 2) or by taking them into account statistically (as in Experiment 3). Thus, we conclude that the abstractness effect we reported earlier is accounted for by affective associations, results that provide the foundations for the embodied view of the semantic representation of abstract words that we spell out next.

General Discussion

An Embodied Theoretical View of Abstract Representation

In three experiments and two large-scale regression analyses we have shown that neither dual coding nor context availability can account for differences in processing of concrete and abstract words. Once imageability and context availability ratings (operationalizing dual coding and context availability hypotheses, respectively) are taken into account, abstract words are processed faster than concrete words. As we have shown, this advantage for abstract over concrete words can be accounted for in terms of greater degree of affective associations for abstract words. Next we present an embodied theory to account for these results.

In contrast to amodal theories of semantic and conceptual representation (perhaps best exemplified in the work of Fodor, 1983, and Jackendoff, 2002), embodied theories of cognition (an early example of which is dual coding theory) propose that cognition is grounded in bodily states, modal simulations, and situated action (Barsalou, 1999; Barsalou, Simmons, Barbey, & Wilson, 2003; Decety & Grezes, 2006; Gibbs, 2006; Rizzolatti & Craighero, 2004). Although embodied theories of semantic representation disagree about the directness of the link between semantic and experiential information (see, e.g., Gallese & Lakoff, 2005, vs. Vigliocco, Vinson, Lewis, & Garrett, 2004), they share the core assumption that the representation and processing of semantic information recruit the same neural systems that are engaged during perception and action. Recent work has provided evidence for such a link between semantic and sensorimotor information, by showing that either perception/action affects semantic computation (Kaschak et al., 2005; Kaschak, Zwaan, Aveyard, & Yaxley, 2006) or semantic computation affects perception/action (Meteyard, Bahrami, & Vigliocco, 2007; Meteyard, Zokaei, Bahrami, & Vigliocco, 2008; see Meteyard & Vigliocco, 2008, for a recent review of the evidence).

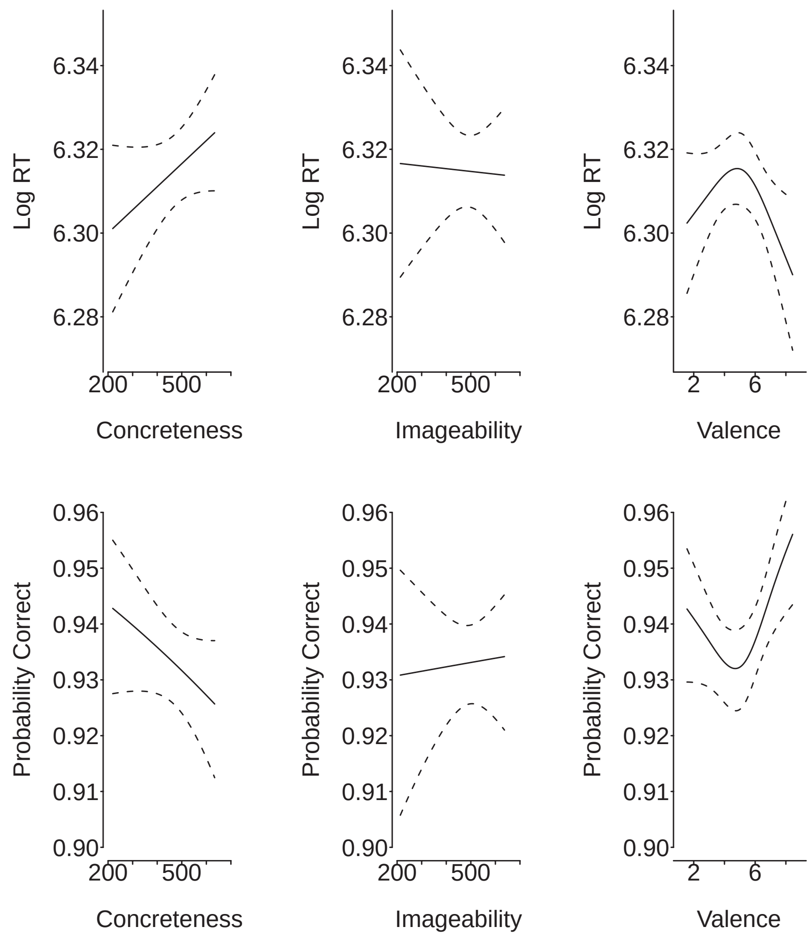

Figure 5. Plots of the partial effects of concreteness, imageability, and valence in Experiment 3 (upper panels indicate log reaction times [RTs]; lower panels indicate accuracy). Dashed lines represent 95% confidence intervals.

Although embodied approaches can be straightforwardly applied to the representation and processing of concrete word meanings, it is far less obvious how an embodied account can be valid for abstract word meanings, which have traditionally been considered to be within the purview of purely verbal systems. In one approach, which originates in work in cognitive linguistics, abstract concepts are grounded metaphorically in embodied and situated knowledge (Gibbs, 1994; Lakoff & Johnson, 1980, 1999). For example, communication of ideas can be understood in terms of goal-directed motion (e.g., throw an idea), and emotional states can be understood in terms of verticality (e.g., happy is up, and sad is down). Although there is increasing evidence that metaphors play a role in the conceptualization of some abstract domains (Boroditsky & Ramscar, 2002; Gibbs, 2006), it is a matter of controversy to what extent they are foundational in the development (and subsequent representation) of abstract concepts and word meanings or whether they provide structure to preexisting conceptual content (Barsalou, 1999; but see Glenberg, Sato, & Cattaneo, 2008).

One embodied account that offers the possibility of accounting for abstract words as well as concrete words has been put forward by Vigliocco, Meteyard, Andrews, and Kousta (2009). The main assumptions of this hypothesis are as follows:

1. Two classes of information contribute to the representation of all concepts (both concrete and abstract): experiential (sensory, motor, and affective) and linguistic (verbal associations arising through patterns of co-occurrence and syntactic information).

2. Differences between concrete and abstract word meanings, as well as differences within each domain (i.e., the domain of concrete words and the domain of abstract words) arise as a result of types and relative proportions of experiential and linguistic information they bind.

3. The apparent dichotomy between concrete and abstract word meanings arises because of a statistical preponderance of sensorimotor information to underlie concrete word meanings and a statistical preponderance of affective and linguistic information to underlie abstract word meanings.

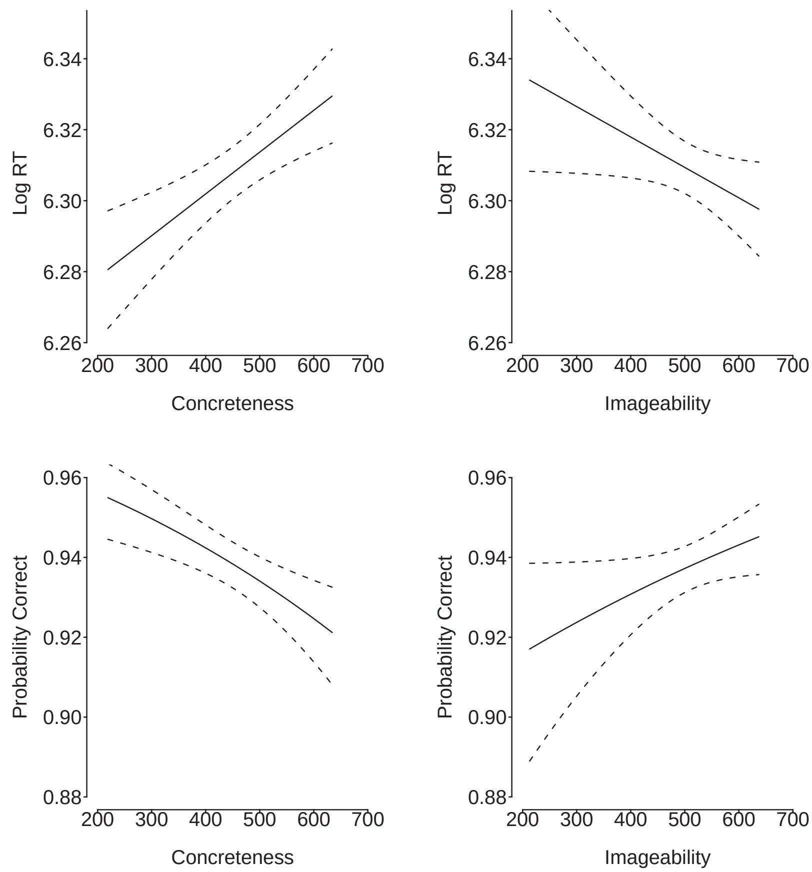

Figure 6. Plots of the partial effects of concreteness and imageability in Experiment 3 if emotional variables (valence and arousal) are not included in the models. Compare these plots with their counterparts in Figure 5, which illustrates these same effects when emotional variables are taken into account. Upper panels indicate log reaction times (RTs); lower panels indicate accuracy. Dashed lines represent 95% confidence intervals.

This approach is novel in that emotion is considered to be another type of experiential information (along with sensorimotor information) playing an important role in learning, representing, and processing, especially for abstract semantics (Vigliocco et al., 2009). The experiments we have reported here provide the critical evidence in favor of such an account.

It is interesting here that recent work by Havas, Glenberg, et al. (Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010; Havas, Glenberg, & Rinck, 2007) has shown a link between being able to express facial emotion and being able to comprehend emotion. An even more dramatic demonstration has been reported by Pistoia and colleagues (2010) for patients with locked-in syndrome who are tetraplegic and cannot command their facial muscles. Our results suggest that an inability to express facial emotion should also have consequences for abstract knowledge, if emotion plays a fundamental role in acquiring and representing abstract concepts.

Affective Associations and Semantic Representation

The idea that internal, and especially affective, states may play a role in the representation of abstract words and concepts is not new. In addition to the work by Altarriba and colleagues (Altarriba & Bauer, 2004; Altarriba et al., 1999), already discussed, Barsalou and Wiemer-Hastings (2005) also suggested that abstract concepts and word meanings are grounded in introspective states (mental and affective). In an exploratory study, Barsalou and Wiemer-Hastings asked speakers to generate features for words varying in concreteness (three highly abstract words: *truth, freedom*, and *invention*; three highly concrete words: *bird, car,* and *sofa*; three intermediate words: *cooking, farming,* and *carpeting*). They found that abstract concepts and word meanings focus on introspective content (as well as social and event content and, less centrally, content about physical settings). We take this idea further by proposing that differences between concrete and abstract words arise because of a general statistical preponderance of affective information for abstract words (and sensorimotor information for concrete words). Why would this be the case? We propose that emotion plays an important role during language acquisition, providing a bootstrapping mechanism for the acquisition of abstract lexical concepts and their labels at early stages.

Emotional development precedes the development of language in children (Bloom, 1998). Words that denote emotional states, moods, or feelings may provide crucial examples of how a word may refer to an entity that is not observable but resides within the organism. In this manner, the acquisition of words denoting emotions, moods, or feelings may actually be a crucial stepping-stone in the development of abstract semantic representations. According to Gleitman and colleagues (Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005), early word learning is effected by means of word-to-world mappings (i.e., by observing the situational contingencies of word usage), which is the case for a limited set of words that refer to concrete, basic-level concepts. Here we propose that abstract words denoting emotional states, moods, or feelings also fall in the same category of words for which a mapping from the word to the world (albeit the internal world) is possible. Consistent with this hypothesis, words denoting emotional states emerge early in language development, at around 20 months of age, and their rate of acquisition increases rapidly in the 3rd year of life (Bretherton & Beeghly, 1982; Wellman, Harris, Banerjee, & Sinclair, 1995). For instance, Ridgeway, Waters, and Kuczaj (1985) reported that 76.7% of children ages 18–23 months have acquired the meaning of the words *good* and *happy*.

Thus, according to our hypothesis, abstract words with affective associations should be acquired earlier than are neutral abstract words. To address this prediction, we took 2,120 words for which we have concreteness, age of acquisition, and valence ratings; partitioned the concreteness scale at the mean; and regressed age of acquisition ratings on valence ratings for abstract words using polynomial models. For abstract words, valence and age of acquisition are related by a U-shaped function—combined linear and quadratic components: $F(2, 1026) = 28.34$; $p < .001$; quadratic alone: $F(1, 1026) = 47.46$, $p < .001$ (see Figure 7). Higher order polynomial terms were not significant predictors of age of acquisition; emotionally significant abstract words, regardless of valence, are acquired earlier than are neutral abstract words. Although valence explains just under 8% of the variance in adult age of acquisition ratings for abstract words, these data are indicative of the possibility that emotion may provide a bootstrapping mechanism for the acquisition of abstract words.

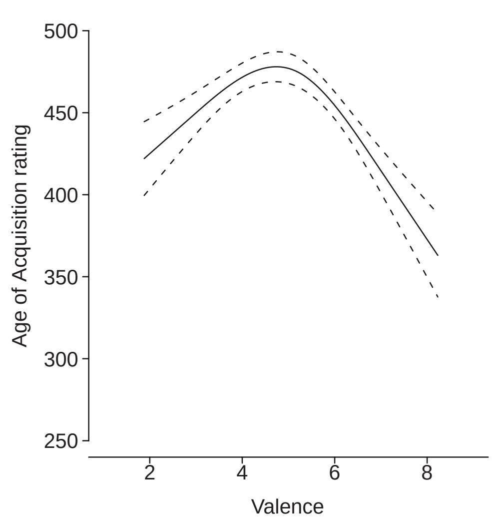

Figure 7. Plot of the effect of valence on age of acquisition for abstract words. Dashed lines indicate 95% confidence intervals.

The Necessity of Integrating Experiential and Linguistic Information

Although we have argued and provided evidence for a foundational role of experiential information in the semantic representation of abstract words, this may not be the whole story. First, it is intuitively clear that language provides vital information as well: After all, we learn a great many words from being told or reading about them. Second, many of the "nuisance variables" we have taken into account in our earlier analyses are not straightforwardly linked to experiential information. After all, more linguistic factors such as number of letters, orthographic neighborhood size, orthographic regularity, and frequency of occurrence also consistently predict lexical decision latencies and accuracy across the analyses we report here. Although these variables on their own do not account for the abstractness effect (otherwise it would have been eliminated by taking them into account statistically), it is important not to discount linguistic factors that may relate to processing of abstract and concrete words.

A role for linguistic information in semantic representation is emphasized by most theories (see Vigliocco & Vinson, 2007, for a discussion) and is supported by the imaging studies reviewed in the introduction, which reported greater activation for abstract/less imageable words in a left-lateralized language network, including LIFG, inferior frontal gyrus (IFG), and left superior temporal sulcus (Binder et al., 2009; Binder et al., 2005; Fiebach & Friederici, 2004; Jessen et al., 2000; Kiehl et al., 1999; Mellet et al., 1998; Noppeney & Price, 2004; Perani et al., 1999; Wise et al., 2000). Interestingly, in the Vigliocco et al. (2010) fMRI study we found that, once valence and arousal ratings were entered into a regression model to predict activation data, the significant greater activations reported in the rostral anterior cingulate cortex for abstract words (linked to emotion) were no longer present. However, abstract words were shown to engage the LIFG to a greater extent than did concrete words. Although the specific role of LIFG in language processing is still highly controversial, here these activations might simply indicate a greater reliance on linguistic information for abstract words.

Assuming a greater role for linguistic information for abstract words may also provide a way to account for the electrophysiological differences between concrete and abstract words that we discussed in the introduction. These studies have reported an amplified N400 in response to concrete versus abstract words, a finding that has been accounted for in terms of amodal theories of meaning such as the context availability model: Concrete words activate a larger amount of related contextual information in verbal memory, which makes integration of the appropriate featural representation into a wider contextual interpretation more difficult. However, the N400 concreteness effect has a different distribution to the classical N400 elicited in response to integration difficulties associated with texts (Kutas & Hillyard, 1980) and is more similar in distribution to the N400 elicited in single-word tasks manipulating lexical variables (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999). Recent evidence has also suggested that the N400 is modulated not only by postlexical controlled processes (e.g., integration processes) but also by automatic lexical processes (e.g., retrieval processes; see Barber & Kutas, 2007, for a review). It is therefore plausible that the N400 concreteness effect reflects difficulty in retrieving specifically linguistic information in response to concrete words. If this is correct, then we would expect to obtain an amplified N400 to concrete words with a similar distribution to that reported in earlier studies, even with materials such as those we used in Experiment 1, and this is precisely what Barber, Otten, Kousta and Vigliocco (2010) found.

Finally, the importance of integrating experiential and linguistic information in learning and representing semantic and conceptual knowledge is highlighted by computational work, using Bayesian probabilistic models, in which representations that combine these two types of information (for both concrete and abstract words) provide a better fit to semantic effects in behavioral tasks (Andrews, Vigliocco, & Vinson, 2009). Also relevant in this respect are approaches to lexical development, such as the syntactic bootstrapping account (Gleitman, 1990; Landau & Gleitman, 1985), which explore the role of syntactic information in acquiring the meaning of especially abstract words (Gleitman et al., 2005; see also Andrews & Vigliocco, 2010, for a computational demonstration of how sequential information may play an important role in learning semantic representations). In such accounts, at early stages, word acquisition relies on word-to-world mappings, in which situational contingencies of use enable the learning of new words. As knowledge of linguistic structure becomes more sophisticated, learners develop the ability to perform structure-to-world mappings that enable further learning of, especially abstract, words.

In closing, it is important to note that the abstractness effect we reported should be evaluated within the context of the concreteness effect: Zero-order correlations between concreteness and behavioral measures reveal an advantage for concrete words, and thus we are not invalidating this textbook finding. However, what may have created confusion in previous work is the attempt to specify a single process or type of information as responsible for differences between the two types of word meanings. Here instead we adopted a working hypothesis according to which concrete and abstract words differ along a number of dimensions, including differential recruitment of sensory, motoric, affective, and linguistic information. According to such an approach, the dimensions along which concrete and abstract words differ may not always point to an advantage for concrete words.

References

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language, 55,* 64–88. doi:10.1016/j.jml.2006.02.002

Altarriba, J. (2008). Expressions of emotion as mediated by context. *Bilingualism, Language and Cognition, 11,* 165–167. doi:10.1017/S1366728908003295

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology, 117,* 389–410. doi:10.2307/4149007

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, 31,* 578–602.

Andrews, M., & Vigliocco, G. (2010). The hidden Markov topic model: A probabilistic model of semantic representation. *Topics in Cognitive Science, 2,* 101–113. doi:10.1111/j.1756-8765.2009.01074.x

Andrews, M., Vigliocco, G., & Vinson, D. P. (2009). Integrating experiential and distributional data to learn semantic representations. *Psychological Review, 116,* 463–498. doi:10.1037/a0016261

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55,* 290–313. doi:10.1016/j.jml.2006.03.008

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Barber, H., Kousta, S.-T., Otten, L. J., & Vigliocco, G. (2010). *Word representation in the brain: Rethinking the ERP concreteness effect*. Manuscript in preparation.

Barber, H. A., & Kutas, M. (2007). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Research Reviews, 53,* 98–123. doi:10.1016/j.brainresrev.2006.07.002

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577–660.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7,* 84–91. doi:10.1016/S1364-6613(02)00029-3

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. A. Zwaan, *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 129–163). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511499968.007

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11,* 235–260. doi:10.1162/089892999563373

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767–2796. doi:10.1093/cercor/bhp055

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract words. *Journal of Cognitive Neuroscience, 17,* 905–917. doi:10.1162/0898929054021102

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 582–594. doi:10.1037/0278-7393.13.4.582

Bloom, L. (1998). Language acquisition in its developmental context. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology* (Vol. 2, pp. 309–370). New York, NY: Wiley.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley.

Boroditsky, L., & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science, 13,* 185–189. doi:10.1111/1467-9280.00434

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings* (Technical Report C-1). Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Breedin, S. D., Saffran, E. M., & Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology, 11,* 617–660. doi:10.1080/02643299408251987

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18,* 906–921. doi:10.1037/0012-1649.18.6.906

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222. doi:10.1016/S1364-6613(00)01483-2

Cartwright, D. S., & Nickerson, C. A. (1979). An empirical thesaurus:

Meaning norms for ninety common words. *Modern Language Journal, 63,* 442–447. doi:10.2307/326031

Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *Journal of the International Neuropsychological Society, 1,* 104–110. doi:10.1017/S1355617700000163

Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, 36,* 371–383.

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

Coltheart, M., Patterson, K. E., & Marshall, J. C. (1980). *Deep dyslexia.* London, England: Routledge.

Conger, A. J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement, 34,* 35–46. doi:10.1177/001316447403400105

Decety, J., & Grezes, J. (2006). The power of simulation: Imagining one's own and other's behaviour. *Brain Research, 1079,* 4–14. doi:10.1016/j.brainres.2005.12.115

de Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 824–845. doi:10.1037/0278-7393.15.5.824

D'Esposito, M. D., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., & Farah, M. J. (1997). A functional MRI study of mental image generation. *Neuropsychologia, 35,* 725–730. doi:10.1016/S0028-3932(96)00121-2

Duyck, W., Desmet, T., Verbeke, L., & Brysbaert, M. (2004). WordGen: A tool for word selection and non-word generation in Dutch, German, English, and French. *Behavior Research Methods, 36,* 488–499.

Estes, Z., & Adelman, J. S. (2008). Automatic vigilance for negative words in lexical decision and naming: Comment on Larsen, Mercer, and Balota (2006). *Emotion, 8,* 441–444. doi:10.1037/1528-3542.8.4.441

Fellbaum, C. (1998). *WordNet: An electronic lexical database.* Cambridge, MA: MIT Press.

Fiebach, C. J., & Friederici, A. D. (2004). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia, 42,* 62–70. doi:10.1016/S0028-3932(03)00145-3

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage, 32,* 1413–1421. doi:10.1016/j.neuroimage.2006.06.007

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press

Franklin, S., Howard, D., & Patterson, K. (1995). Abstract word anomia. *Cognitive Neuropsychology, 12,* 549–566. doi:10.1080/02643299508252007

Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex, 10,* 698–705. doi:10.1093/cercor/10.7.698

Friedman, L., & Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple linear regression. *American Statistician, 59,* 127–136. doi:10.1198/000313005X41337

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22,* 455–479. doi:10.1080/02643290442000310

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding.* Cambridge, England: Cambridge University Press.

Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind and Language, 21,* 434–458. doi:10.1111/j.1468-0017.2006.00285.x

Giesbrecht, B., Camblin, C. C., & Swaab, T. Y. (2004). Separable effects of semantic priming and imageability on word processing in human cortex. *Cerebral Cortex, 14,* 521–529. doi:10.1093/cercor/bhh014

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1,* 3–55. doi:10.1207/s15327817la0101_2

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1,* 23–64. doi:10.1207/s15473341lld0101_4

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language processing. *Current Biology, 18,* R290–R291. doi:10.1016/j.cub.2008.02.036

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., & Gee, J. (2002). The neural basis for category-specific knowledge: An fMRI study. *NeuroImage, 15,* 936–948. doi:10.1006/nimg.2001.1028

Hale, S. C. (1988). Spacetime and the abstract/concrete distinction. *Philosophical Studies, 53,* 85–102. doi:10.1007/BF00355677

Harrell, F. E. (2001). *Regression modeling strategies with applications to linear models, logistic regression and survival analysis.* New York, NY: Springer–Verlag.

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *Annals of Statistics, 13,* 70–84. doi:10.1214/aos/1176346577

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science,* doi:10.1177/0956797610374742.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14,* 436–441.

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 721–742. doi:10.1037/0278-7393.25.3.721

Horst, P. (1941). *The prediction of personal adjustment* (SSRC Bulletin No. 48). New York, NY: Social Science Research Council.

Horst, P. (1966). *Psychological measurement and prediction.* Belmont, CA: Wadsworth.

Howell, J. R., & Bryden, M. P. (1987). The effects of word orientation and imageability on visual half-field presentations with a lexical decision task. *Neuropsychologia, 25,* 527–538. doi:10.1016/0028-3932(87)90077-7

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar evolution.* Oxford, England: Oxford University Press.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 130–136. doi:10.1037/0096-1523.1.2.130

Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual-coding and context availability. *Brain and Language, 74,* 103–112. doi:10.1006/brln.2000.2340

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Research, 1148,* 138–148. doi:10.1016/j.brainres.2007.02.044

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition, 94,* B79–B89. doi:10.1016/j.cognition.2004.06.005

Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science, 30,* 733–744. doi:10.1207/s15516709cog0000_54

Katz, R. B., & Goodglass, H. (1990). Deep dysphasia: Analysis of a rare form of repetition disorder. *Brain and Language, 39,* 153–185. doi:10.1016/0093-934X(90)90009-6

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human Brain Mapping, 7,* 225–233. doi:10.1002/(SICI)1097-0193(1999)7:4<225::AID-HBM1>3.0.CO;2-P

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-encoding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 804–823. doi:10.1037/0278-7393.20.4.804

Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112,* 473–481. doi:10.1016/j.cognition.2009.06.007

Krause, B. J., Schmidt, D., Mottaghy, F. M., Taylor, J., Halsband, U., Herzog, H., . . . Muller-Gartner, H. W. (1999). Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates: A PET study. *Brain, 122,* 255–263. doi:10.1093/brain/122.2.255

Kroll, J., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 92–107. doi:10.1037/0278-7393.12.1.92

Kutas, M., & Hillyard, S. A. (1980, January 11). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205. doi:10.1126/science.7350657

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1994–2005. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). New York, NY: Elsevier. doi:10.1016/B978-012369374-7/50018-3

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* New York, NY: Basic Books.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child.* Cambridge, MA: Harvard University Press.

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex, 17,* 742–748. doi:10.1093/cercor/bhk024

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1,* 173–181. doi:10.1023/A:1026595011371

Marshall, J., Pring, T., Chiat, S., & Robson, J. (1996). Calling a salad a federation: An investigation of semantic jargon: Part 1—nouns. *Journal of Neurolinguistics, 9,* 237–250. doi:10.1016/S0911-6044(97)82796-0

Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language, 43,* 240–274. doi:10.1016/0093-934X(92)90130-7

McFatter, R. M. (1979). The use of structural equation models in interpreting regression equations including suppressor and enhancer variables. *Applied Psychological Measurement, 3,* 123–135. doi:10.1177/014662167900300113

Mellet, E., Tzourio, N., Denis, M., & Mazoyer, B. (1998). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *NeuroReport, 9,* 803–808. doi:10.1097/00001756-199803300-00007

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science, 18,* 1007–1013. doi:10.1111/j.1467-9280.2007.02016.x

Meteyard, L., & Vigliocco, G. (2008). The role of sensory and motor information in semantic representation: A review. In P. Calvo & T. Gomila (Eds.), *Handbook of cognitive science: An embodied approach* (pp. 293–312). Academic Press, Elsevier, London.

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology, 18,* R732–R733.

Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory & Cognition, 25,* 256–263.

Nelson, D. L., & Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language, 31,* 237–260. doi:10.1016/0749-596X(92)90013-N

Nittono, H., Suehiro, M., & Hori, T. (2002). Word imageability and N400 in an incidental memory paradigm. *International Journal of Psychophysiology, 44,* 219–229. doi:10.1016/S0167-8760(02)00002-8

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *NeuroImage, 22,* 164–170. doi:10.1016/j.neuroimage.2003.12.010

Osterhout, L., & Holcomb, P. J. (1995). Event-related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171–215). Oxford, England: Oxford University Press.

Paivio, A. (1971). *Imagery and verbal processes.* New York, NY: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental representations: A dual coding approach.* Oxford, England: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45,* 255–287. doi:10.1037/h0084295

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach.* Mahwah, NJ: Erlbaum.

Paivio, A., Yuille, J. C., & Smythe, P. C. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology, 20,* 362–377. doi:10.1037/h0082949

Papagno, C., Capasso, R., Zerboni, H., & Miceli, G. (2007). A reverse concreteness effect in a subject with semantic dementia. *Brain and Language, 103,* 90–91. doi:10.1016/j.bandl.2007.07.059

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., & Fazio, F. (1999). The neural correlates of verb and noun processing. *Brain, 122,* 2337–2344. doi:10.1093/brain/122.12.2337

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience, 19,* 1407–1419. doi:10.1162/jocn.2007.19.8.1407

Pistoia, F., Conson, M., Trojano, L., Grossi, D., Ponari, M., Colonnese, C., . . . Sara, M. (2010). Impaired conscious recognition of negative facial expressions in patients with locked-in syndrome. *Journal of Neuroscience, 30,* 7838–7844. doi:10.1523/JNEUROSCI.6300-09.2010

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10,* 377–500. doi:10.1080/02643299308253469

Reilly, J., & Kean, J. (2007). Formal distinctiveness of high- and low-imageability nouns: Analyses and theoretical implications. *Cognitive Science, 31,* 1–12.

Reilly, J., Peelle, J. E., & Grossman, M. (2007). A unitary semantics account of reverse concreteness effects in semantic dementia. *Brain and Language, 103,* 86–87. doi:10.1016/j.bandl.2007.07.057

Richardson, J. (2003). Dual coding versus relational processing in memory for concrete and abstract words. *European Journal of Cognitive Psychology, 15,* 481–509. doi:10.1080/09541440244000256

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21,* 901–908. doi:10.1037/0012-1649.21.5.901

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169–192. doi:10.1146/annurev.neuro.27.070203.144230

Roeltgen, D. P., Sevush, S., & Heilman, K. M. (1983). Phonological agraphia: Writing by the lexical-semantic route. *Neurology, 33,* 755–765.

Romani, C., McAlpine, S., & Martin, R. C. (2007). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology, 61,* 292–323.

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behaviour. *Journal of Verbal Learning and Verbal Behavior, 19,* 736–755.

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005).

Modulation of the semantic system by word imageability. *NeuroImage, 27,* 188–200. doi:10.1016/j.neuroimage.2005.04.012

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). Hillsdale, NJ: Erlbaum.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language, 27,* 499–520. doi:10.1016/0749-596X(88)90022-8

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82–102. doi:10.1037/0278-7393.9.1.82

Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly, 24,* 114–126. doi:10.2307/748013

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7,* 422–445. doi:10.1037/1082-989X.7.4.422

Sirigu, A., Duhamel, J. R., & Poncet, M. (1991). The role of sensorimotor experience in object recognition: A case of multimodal agnosia. *Brain, 114,* 2555–2573. doi:10.1093/brain/114.6.2555

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38,* 598–605.

Ter Doest, L., & Semin, G. R. (2005). Retrieval contexts and the concreteness effect: Dissociations in memory for concrete and abstract words. *European Journal of Cognitive Psychology, 17,* 859–881. doi:10.1080/09541440540000031

Tse, C. S., & Altarriba, J. (2009). The word concreteness effect occurs for positive, but not negative, emotion words in immediate serial recall. *British Journal of Psychology, 100,* 91–109.

Tyler, L. K., Russell, R., Fadili, J., & Moss, H. E. (2001). The neural representation of nouns and verbs: PET studies. *Brain, 124,* 1619–1634. doi:10.1093/brain/124.8.1619

Tzelgov, J., & Henik, A. (1991). Suppression situations in psychological research: Definitions, implications, and applications. *Psychological Bulletin, 109,* 524–536. doi:10.1037/0033-2909.109.3.524

van Schie, H. T., Wijers, A. A., Mars, R. B., Benjamins, J. S., & Stowe, L. A. (2005). Processing of visual semantic information to concrete words: Temporal dynamics and neural mechanisms indicated by event-related brain potentials. *Cognitive Neuropsychology, 22,* 364–386. doi:10.1080/02643290442000338

Vigliocco, G., della Rosa, P., Vinson, D. P., Kousta, S.-T., Devlin, J. T., & Cappa, S. F. (2010). *The neural representation of abstract words.* Manuscript in preparation.

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition, 1,* 219–248. doi:10.1515/LANGCOG.2009.011

Vigliocco, G., & Vinson, D. P. (2007). Semantic representation. In G. Gaskell & G. Altmann (Eds.), *Oxford handbook of psycholinguistics* (pp. 195–215). Oxford, England: Oxford University Press.

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology, 48,* 422–488. doi:10.1016/j.cogpsych.2003.09.001

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1256–1271. doi:10.1037/0278-7393.25.5.1256

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27,* 635–657. doi:10.1080/14640747508400525

Warrington, E. K., & Shallice, T. (1984). Category-specific semantic impairment. *Brain, 107,* 829–853. doi:10.1093/brain/107.3.829

Wellman, H. M., Harris, P. L., Banerjee, M., & Sinclair, A. (1995). Early understanding of emotion: Evidence from natural language. *Cognition & Emotion, 9,* 117–149. doi:10.1080/02699939508409005

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience, 12,* 1024–1037. doi:10.1162/08989290051137558

Whaley, C. P. (1978). Word-nonword classification times. *Journal of Verbal Learning and Verbal Behavior, 17,* 143–154. doi:10.1016/S0022-5371(78)90110-X

Wise, R. J. S., Howard, D., Mummery, C. J., Fletcher, P., Leff, A., Buchel, C., & Scott, S. K. (2000). Noun imageability and the temporal lobes. *Neuropsychologia, 38,* 985–994. doi:10.1016/S0028-3932(99)00152-9

Appendix A

Experiment 1 Items

Table A1
Concrete and Abstract Items Used in Experiment 1

Concrete	Abstract
office	horror
cancer	beauty
ounce	grief
relic	demon
trunk	spree
lamp	hell
estate	luxury
duke	fury
cousin	angel
rector	frenzy
leek	oath
gig	woe
ether[a]	havoc[a]
guest	crime
prong[a]	wealth[a]
creature	concert
oak	joy
date	love
stomach	romance
author	thrill
block	panic
asbestos	paradise
jersey	danger
channel	protest
column	temper
material	fashion
sound	minute
stick	ghost
plate	space
voice	dream
monsoon	slumber
belt	joke
freight	expanse
starch	burden
disease	number
weapon	dozen
manure	plunge
garment	bargain
lobby	quest
bureau	triumph

[a] Item excluded from analysis.

(*Appendices continue*)

Appendix B

Suppression and Enhancement in Linear Regression

Horst (1941, 1966) was the first to note that there are cases in which a predictor whose zero-order validity is 0 (i.e., a variable that is entirely uncorrelated with the criterion variable) but that correlates significantly with another predictor improves prediction when included in a regression model. As a substantive example, he described a World War II study aimed at predicting pilot success as a function of mechanical, numerical, spatial, and verbal ability—all measured in pen-and-paper tasks. The first three variables had a significant positive correlation with flying ability. Although verbal ability had a negligible correlation with pilot success, it correlated highly with the other three predictors. When verbal ability was included in the regression model, the amount of variance explained increased, despite the fact that verbal ability had a negligible zero-order correlation with flying ability. Verbal ability, however, was needed to read the instructions and items on the pen-and-paper tests. Thus, the way mechanical, numerical, and spatial ability were measured introduced measurement error variance into the scores. Including verbal ability scores as a predictor in the model improves overall prediction by removing artifactual measurement error from the other predictors. To illustrate, using for simplicity a two-predictor model with mechanical ability (X_1) and verbal ability (X_2) as predictors of flying ability (X_0), we assumed the following sample correlations between the (standardized) variables: $r_{01} = .8$, $r_{02} = 0$, and $r_{12} = .4$.

We used Equation B1 to calculate the least squares estimate of the regression coefficients

$$\beta_1 = (r_{01} - r_{02}r_{12})/(1 - r_{12}^2) \quad \text{(B1)}$$

and found that the coefficient for X_1 was higher than r_{01},

$$\beta_1 = (.8 - 0 \times .4)/(1 - .4^2) = .95 > .8, \quad \text{(B2)}$$

whereas the coefficient for X_2 received a negative weight despite the fact that $r_{02} = 0$, as seen in

$$\beta_2 = (0 - .8 \times .4)/(1 - .4^2) = -.38 \neq 0. \quad \text{(B3)}$$

The amount of variance explained in the criterion is given by Equation B4 and is calculated in Equation B5:

$$R^2 = (r_{01}^2 + r_{02}^2 - 2r_{01}r_0r_{12})/(1 - r_{12}^2) \quad \text{(B4)}$$

$$R^2 = [.8^2 + 0^2 - (2 \times .8 \times 0 \times .4)]/(1 - .4^2)$$
$$= .76 > .8^2 + 0^2. \quad \text{(B5)}$$

In other words, the squared multiple correlation coefficient exceeds the sum of the two squared simple correlation coefficients with X_0: $R^2 > {r_{01}}^2 + {r_{02}}^2$. Including verbal ability with a negative weight in the regression equation serves to penalize those participants whose mechanical ability scores were high purely because of verbal ability and to compensate those participants whose mechanical ability scores were low purely because of low verbal ability. This improves the predictive validity of mechanical ability and hence the amount of variance explained by the model. It is not necessary for X_2 to be 0 for predictive validity to be improved; enhancement/suppression is possible even when $X_2 \neq 0$ (termed net enhancement/suppression; see Friedman & Wall, 2005, for intervals of possibilities for r_{01}, r_{02}, and r_{12} for net enhancement/suppression to obtain).[B1]

The explanation of enhancement as suppression of irrelevant variance in another predictor implies a specific underlying model, the two-factor model (Conger, 1974). Conger (1974, p. 37) introduced the two-factor model in this way:

> Discussions of suppressor variables . . . suggest an underlying model in which there is nonerror variance in the predictor which is unrelated to the criterion (factor S) as well as an uncorrelated common factor of that which the criterion is measuring (factor T). The suppressor either has no relation to the criterion (a loading of zero for factor T) or is measuring the criterion less than it is measuring the irrelevant variance.

So in the previous example, verbal ability scores measure true verbal ability (T), whereas mechanical ability scores measure both true flying ability (S) and true verbal ability (T)—the latter because the mechanical ability test needed verbal ability in order to be carried out. Verbal ability and flying ability are uncorrelated.

Although the misconception that enhancement can be interpreted only as suppression of irrelevant variance continues in some settings until today, McFatter (1979; see also Bollen, 1989; MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002, for more recent treatments) pointed out that the two-factor model is just one type of underlying model that can give rise to enhancement and that there are several cases in which interpreting enhancement as suppression of irrelevant variance is not warranted. He provided the following substantive example of a case in which interpreting enhancement as suppression of irrelevant variance is meaningless:

> Suppose one were interested in predicting the number of errors made by assembly-line workers as a function of IQ and Intolerance of Boredom scores. Let X_0 be number of errors, X_1 be Intolerance of Boredom score, and X_2 be IQ score. If one obtained sample correlations of $r_{01} = .3535$, $r_{02} = 0$, and $r_{12} = .707$; noted that this was a case of classical enhancement; and relied on the usual discussions of suppression, one might be tempted to conclude that IQ was totally irrelevant to the number of assembly-line errors made, but did measure precisely that aspect of Intolerance of Boredom which is also irrelevant to the number of errors made. This is the interpretation one would make were the two-factor model the structure underlying these variables. (p. 128)

[B1] There is a third type of enhancement/suppression (cooperative enhancement/suppression), first introduced by Conger (1974), but it is beyond the scope of this appendix.

(*Appendices continue*)

McFatter (1979) formally demonstrated that a number of completely different underlying models can give rise to enhancement and that the only case in which enhancement can be interpreted as suppression of irrelevant variance is if the underlying model is assumed to be the two-factor model. For instance, he showed how enhancement can arise out of both the two-factor model and an alternative model in which a predictor has both a direct and indirect effect on the criterion. We provide an explanation of his specifications for the two alternative models next. In both cases it is assumed that variables are standardized and that r_{01}, $r_{12} > 0$.

Two-Factor Model

McFatter (1979) specifies the path diagram depicted in Figure B1 for the two-factor model. The structural equations corresponding to the path diagram in Figure B1 are as follows:

$$X_0 = a_0T + e_0$$

$$X_1 = a_1T + b_1S + e_1$$

$$X_2 = a_2T + b_2S + e_2.$$

Given those equations, the population correlation matrix for this model would be given as follows:

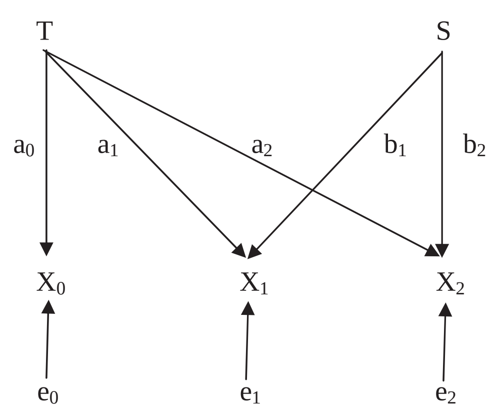

Figure B1. Path diagram for the two-factor model. X_1 and X_2 are possible predictors of dependent variable X_0. Underlying factors are expressed by S and T: S indicates a predictor unrelated to the criterion; T indicates an uncorrelated common factor of that which the criterion is measuring. Structural equation coefficients are expressed by a and b, with error terms e (see text for details).

$$\rho_{01} = a_0a_1$$

$$\rho_{02} = a_0a_2$$

$$\rho_{12} = a_1a_2 + b_1b_2.$$

McFatter shows that classical enhancement obtains in this model if $a_2 = 0$ and $b_1b_2 > 0$; net enhancement obtains if $a_2 < a_1(a_1a_2 + b_1b_2)$.

Direct and Indirect Effects

McFatter (1979) then presents the path diagram depicted in Figure B2 for a model in which a predictor has both a direct and an indirect effect (through another predictor) on a dependent variable. In this model, stochastic disturbance terms u and v are included in order to represent all sources of variation that are not included in the model—with E(u) = E(v) = E(uv) = 0. The structural equations for this model are as follows:

$$X_0 = b_1X_1 + b_3X_2 + u$$

$$X_1 = b_2X_2 + v.$$

These equations generate the following correlation matrix for this model:

$$\rho_{01} = b_1 + b_2b_3$$

$$\rho_{02} = b_1 + b_1b_2$$

$$\rho_{12} = b_2.$$

When b_1 and b_2 are positive and $b_3 = -b_1b_2$, classical enhancement will obtain in this model; when $-b_1b_2 < b_3$ and b_1, $b_2 < 0$, net enhancement will obtain. Returning to the IQ–boredom errors example mentioned earlier, if, for instance, $b_1 = b_2 = .707$ and $b_3 = -.50$, then the population correlations are the same as those in the IQ example ($r_{01} = .3535$, $r_{02} = 0$, and $r_{12} = .707$).

As McFatter (1979) noted, in this model the interpretation of the effect of X_2 subtracting or suppressing irrelevant or invalid variance in X_1 is nonsensical. The negative weight of the coefficient for X_2 represents the fact that X_2 in this model has a direct negative effect on X_0 and a compensating positive influence on X_0 through X_1.

It is important to note that there are no objective criteria for deciding between Model 1 and Model 2 as presented earlier—there is nothing in the statistical computations that forces the adoption of either model. Both models can generate the correlation matrix that produces enhancement, and the choice between the models is a matter of theoretical consideration and a priori hypotheses.

(*Appendices continue*)

In our multiple regression models we found an indication that concreteness was functioning as an enhancer variable. Although enhancement/suppression has been extended to models involving more than two predictors (see Tzelgov & Henik, 1991), the situation becomes much more complex. In our case, we were interested in the relationship between the distinction between concrete and abstract concepts and one of its reflexes (i.e., the ability to form a mental image to the referents of words). For this reason, we decided to look at a two-predictor model, in which concreteness and imageability are used as predictors of lexical decision RTs, in order to confirm our intuition that in our large-scale model concreteness functions as an enhancer with respect to imageability. For these analyses, we used imageability (X_1) and concreteness (X_2) ratings for 4,075 words from the MRC Psycholinguistic Database (Coltheart, 1981) and lexical decision latencies (X_0) from the ELP (Balota et al., 2007). The correlation between imageability and latency was $r_{01} = -.261$, the correlation between concreteness and latency was $r_{02} = -.137$, and the correlation between concreteness and imageability was $r_{12} = .849$. For simplicity, we assumed that the relationship between each of the two predictors and the dependent variable was linear. The beta weight for imageability in the two-predictor model was $\beta = -.520$ ($>r_{01}$), the beta weight for concreteness was $\beta = .305$ (different in sign from r_{02}), and the squared multiple correlation coefficient was $R^2 = .094$, which is higher than the sum of the two squared simple correlation coefficients ($r_{01}^2 + r_{02}^2 = .087$). This is indeed a case of net enhancement.

As we saw earlier, the sample simple correlations between concreteness, imageability, and latency can be generated by both the two-factor model and a model in which concreteness has a direct effect on RTs and an indirect effect through imageability (although there are other possibilities, we do not discuss them here because we consider them theoretically irrelevant). We argued in the main text of this article that by hypothesis the latter model fits the data better than does the former. However, one could consider the possibility that the two-factor model underlies our data and try to construct latent variables for this model. For the two-factor model to work one would need to assume that concreteness is a measure of true concreteness (e.g., the ontological distinction between concrete and abstract concepts) and that true concreteness is a poor measure of word-recognition speed. Imageability measures both word-recognition speed and true concreteness. Including concreteness in the model suppresses variance in imageability scores that is due to concreteness and that is irrelevant to measurement of word-recognition speed. Such an interpretation goes against everything anybody has ever claimed about the relationship between concreteness and imageability. The assumption in the concreteness literature is that, theoretically, concreteness reflects the directness of connections between verbal representations and modality-specific imagery (Paivio, 2007). In other words, the assumption is that there is nothing in concreteness that is not explained by the perceived ability (or otherwise) to evoke modality-specific imagery. We are claiming instead that, although the image-evoking aspect of word meanings is one of the reflexes of concreteness, variation in concreteness is not exhausted by the extent to which visual imagery is evoked by the referents of words but that other critical variables are involved. Our point here is that there is no theoretical account of concreteness advanced up to the present to support the proposal that the two-factor model is an appropriate underlying model for the trivariate relationship we are considering. The earlier literature would assume that the effect of concreteness is completely mediated by the effect of imageability; schematically, the model assumed is illustrated in Figure B3 (Panel A).

We are instead proposing that, apart from the indirect link between concreteness and imageability, there is also a direct link from concreteness to word recognition (and we attempt to specify this relationship further by identifying other variables that mediate the relationship between concreteness and behavior). This is illustrated schematically in Figure B3 (Panel B). The two-factor model is inappropriate in either case.

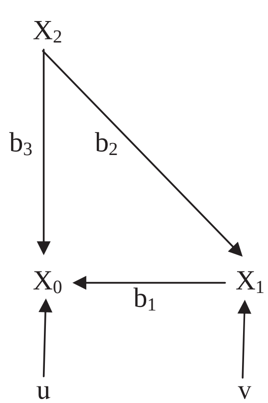

Figure B2. Path diagram for a model I which predictor X_2 has both a direct effect on dependent variable X_0 and an indirect effect via predictor X_1. Structural equation coeffiecients are expressed by b, with stochastic disturbance terms u and v (see text for details).

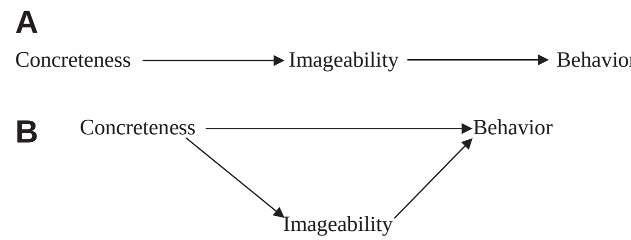

Figure B3. A: Path diagram reflecting prior literature, in which the effect of concreteness on behavior (word recognition) is assumed to be completely mediated by the effect of imageability. B: Path diagram reflecting a direct link from concreteness to behavior (word recognition) in addition to an indirect link via imageability.

Received August 7, 2008
Revision received June 23, 2010
Accepted June 24, 2010 ■