# Sensorimotor dominance in semantic feature listings: When I say ‘dog’, will you say ‘tail’?

**Avery Pattullo[1,2] (avery.pattullo@temple.edu)**

**Ana Matić Škorić [3](ana.matic@erf.unizg.hr)**

**Luca M Scott[2] (luca.scott@temple.edu)**

**Erin Buchanan[4] (ebuchanan@harrisburgu.edu)**

**Jamie Reilly[1,2] (jamie_reilly@temple.edu)**
[1]Department of Communication Sciences and Disorders
Temple University, Philadelphia, Pennsylvania USA

[2]Department of Psychology and Neuroscience
Temple University, Philadelphia, Pennsylvania USA

[3]Department of Speech and Language Pathology, University of Zagreb
Zagreb, Croatia

[4]Harrisburg University of Science and Technology
Harrisburg, Pennsylvania USA

## Abstract

Object knowledge comprises a virtually limitless feature space. When cued to generate attributes for an exemplar (*dog*), any feature is possible (*has molecules*). Yet, some features are definitively favored (*tail*). We hypothesize a sensorimotor bias in feature generation wherein perceptually salient features upon first pass are more cognitively accessible than abstract, verbally mediated knowledge. We examined the role of sensorimotor dominance in semantic feature generation by yoking concreteness values to cues (N=4436) and features (N=69,284) within the Buchanan et al. (2019) norms. We predict that cues regardless of their concreteness evoke relatively more concrete features (e.g., *dogs* evoke *tails*, *justice* evokes *lawyers*). The data moderately supported this hypothesis. Feature concreteness increased linearly with cue concreteness (R=.83) but the y-intercept (2.78) indicates that overall, features were more concrete than their cues. We discuss alternate factors (e.g., frequency, familiarity) that may moderate the likelihood that people retrieve *tail* when cued with *dog*.

**Keywords:** Semantic Feature; Semantic Memory; Lexical Retrieval; Concreteness

## Introduction

In a typical semantic feature listing paradigm, a participant is presented with a written or spoken cue word and must generate as many features (has a...) or properties (is a...) as possible within a fixed time interval (McRae et al., 1997, 2005). Semantic feature generation can tell us about the structure of semantic memory both in terms of underlying structure relations and the strength of association between concepts. One possibility is that some features inherently have a higher activation strength, and that differential activation at baseline is an emergent property of the semantic system reflected in similar response patterns across many individuals (Maki, 2007). For example, dogs are composed of molecules, cells, livers, eyes, and tails (among many other things). Yet, most of us agree that *has a tail* is a more plausible and/or informative feature than *has molecules*. Similarly, justice is an abstract principle that inherently lacks tails, color, or other tangible features. Yet, *justice* can be reified through contextual, symbolic, and verbal associations (e.g., *courtrooms*, *judges*, *scales*).

Depth of processing differentiates semantic feature generation from naturalistic semantic tasks (e.g., identifying an object, retrieving a word). Semantic feature generation is intrinsically a search task where one must access a target concept and activate a constellation of attributes: sensorimotor representations, word associations, polysemous meanings, memories, encyclopedic knowledge, etc. Some of these attributes are more diagnostic and informative than others (e.g., distinctive features). Yet, it is unclear how or indeed whether features compete against one another to attain dominance. We hypothesize that in the battle for salience during deep semantic processing, sensorimotor information wins – that is, sensorimotor information is more readily accessible and therefore more salient. We predict that this



In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* 

sensorimotor bias drives feature generation in a semantic feature task.

Several theories support sensorimotor dominance in semantic processing. In dual coding theory, sensorimotor information causes concrete concepts to have more semantic richness due to being dually represented in both lexical and sensorimotor systems (Paivo, 1999). Under dual coding theory, we would expect concrete features to be more salient (and thus more likely to be selected during a feature generation task) than abstract ones. Similarly, the context availability theory posits that concrete words have more contextual information and therefore processed more effectively (Schwanenflugel et al., 1992); in semantic feature generation a concrete feature with more context availability would also be more salient. Under Barsalou's perceptual symbol systems theory, both concrete and abstract concepts are acquired through sensorimotor experience (including introspective/affective experiences) and during semantic processing, we simulate that experience (Barsalou, 1999). Thus, during a feature generation task, cue words would evoke sensorimotor representations regardless of cue type.

Sensorimotor bias in feature generation could be task- rather than theory-driven: when asked to generate semantic features of a concept, people may (either intentionally or unconsciously) construct a mental image of or create a context for that concept. That is, regardless of the perceptual salience of the cue, people will attempt to situate that target concept within a meaningful event schema (e.g., *forks* in kitchens, *mourning* at a funeral), activating a feature space composed of parts and associated objects and agents that can be experienced through the senses. In turn, people 'read out' what they imagine experiencing in these personally relevant scenes, mentally scanning the cued concept and identifying salient/diagnostic features (Barsalou et al., 2003; Kellenbach et al., 2000; Solomon & Barsalou, 2007). Although such a strategy might be effective in some cases, other target concepts are not amenable to an imagery-based strategy. In such cases, people would need to flexibly pivot to a different cognitive search strategy (e.g., verbal encyclopedic knowledge) not unlike optimal foraging behaviors in category fluency tasks (e.g., when exhausted naming house pets, switch to naming jungle animals; Hills et al., 2012). At some point during a perceptual read out process, people will exhaust the possible range of informative features. At this point, they must disengage from salient sensorimotor features and retrieve more abstract, verbally learned features (e.g., *is loyal* for *dog*).

This default search strategy would apply to any cue, regardless of its own inherent sensorimotor salience. For example, a readout strategy for *beaver* might yield a mental image of a beaver, with a *flat tail*, *buck teeth*, and *brown fur* before pivoting to verbally knowledge like *builds dams* or *eager*. In contrast, a read-out strategy for *persistent* might involve imagining a canonical scene of something behaving persistently (e.g., *tortoise)*, or activation of the physiological, affective and introspective sensations involved in persistence (*blood*, *sweat, tears*).

We argue that semantic feature generation may be biased toward concrete features not only because sensorimotor salience confers a higher baseline activation strength but because semantic feature generation encourages a default strategy of sensorimotor "read out" due to the acontextual nature of the task.

## Do Abstract Words Have Semantic Features?

Wiemer-Hastings and Xu (2005) showed that the types of features produced by participants in the semantic feature production task vary in response to the abstractness or concreteness of the cue. People tend to provide more entity properties (i.e., *tail* and *fur* for *dog*) for concrete words and more experiential properties (i.e., features that are relational, evaluative, or emotional) for abstract words. This dissociation suggests that semantic feature generation engages mental imagery at least for concrete words. In contrast, abstract concepts that lack entity properties have meanings that are more grounded in relational and associative knowledge with other words (e.g., *win* when cued for *glory;* Borghi et al. 2022; Borghi et al., 2014; Dove, 2022; Villani et al., 2021) and in interoceptive and emotional experiences within the body (e.g., *shame* for the cue *disgrace;* Troche et al., 2013; Crutch & Warrington, 2005; Kousta et al., 2011; Reilly et al., 2016; Troche et al., 2014; Vigliocco et al., 2014). Additionally, there is fMRI evidence for activation of visual processing areas during semantic decision-making for abstract concepts, suggesting some perceptual enactment of abstract concepts (Pexman et al., 2007).

## Predictions And Aims

We leveraged the largest known set of publicly available norms (Buchanan et al., 2019) to elucidate the role of sensorimotor salience in semantic feature generation. This aggregated corpus includes 4436 cues paired with 69,284 corresponding features. Our aim was to evaluate the impact of sensorimotor salience on numerous cue-feature relationships. Since the feature listings involved words, this necessitated a deliberate choice about what might be the best psycholinguistic indicator variable at the lexical level (e.g., concreteness, imageability, perceptual strength). This decision was informed both by alignment with the construct of interest and maximizing the amount of coverage across almost 70,000 cue-feature pairs. We elected to yoke concreteness values to each cue and feature. Thus, our predictions to follow reflect the assumption that concreteness is a viable proxy measure for sensorimotor salience.

Our hypothesis is that semantic feature generation will be biased toward concrete attributes (i.e., things that can be seen, heard, touched, felt) with a pivot to verbally learned features once sensorimotor attributes are exhausted. That is, all cues regardless of their concreteness level (e.g., *beach* vs. *judgment*) initially evoke features with sensorimotor salience. It is important to note, however, that our hypothesis of sensorimotor bias does not mean that we predict equally high feature concreteness for abstract vs. concrete cues, but rather that, across the concreteness spectrum, features will tend to be as or more concrete than their cues.

We further predict that as concreteness of the cue rises, people will produce more features and show higher intersubject agreement on the features they do in turn generate. That is, concreteness facilitates feature production, allowing participants to produce more features, and concreteness also provides a shared sensorimotor framework that acts to reduce diffusion of features across participants.

## Methods

### Overview

We executed numerous data transformations on the semantic feature production norms reported by Buchanan and colleagues (2019). These transformations yielded a summary dataframe grouped by cue with all subsequent information about the cue's constituent features aggregated (e.g., arithmetic weighted mean of concreteness values for all features of a cue). We then indexed this summary dataframe to investigate relations between the cues and features (see measures). Data and R scripts are available for download via the OSF (link anonymized) at https://osf.io/juhnf/?view_only=5e627980fc124424975dad454f3a631f.

### Materials and Measures

We analyzed the semantic feature norms reported by Buchanan and colleagues (2019). This dataset consists of almost 70k English cue-feature pairings. Cues varied in semantic complexity and part-of-speech with an approximate distribution that includes 70% nouns, 14.9% adjectives, 12.4% verbs, and 2.3% other forms (Buchanan et al., 2019). These norms were obtained from instructions for participants to list physical, functional, and categorical features of the cue words. The data included both raw agreement and normalized agreement values that are adjusted for different numbers of participants across studies (e.g., 60 per cue vs. 45 per cue). The original transcriptions were lemmatized and stopwords were omitted (e.g., has a, is a). The feature data were then high pass filtered to retain only features that elicited agreement across more than two participants.

We first imported the raw dataset into R (R Core Team, 2021) and then joined each cue and feature to their corresponding lexical norms in an offline lookup database using *dplyr* (Wickham et al., 2023). We appended norms for concreteness (Brysbaert et al., 2014), word frequency (Brysbaert & New, 2009), familiarity (Brysbaert et al., 2025), and semantic neighborhood density (rescaled from 0-10) (Gao et al., 2022; Shaoul & Westbury, 2010) to any word with a match in the lookup database.

After appending lexical norms, we transformed the original dataset (4436 cues, 69,284 features) into a summary dataframe grouped by cue. This process involved collapsing all information across features to a single datapoint corresponding to each cue. We computed weighted arithmetic means of concreteness, familiarity, frequency, and neighborhood density. These means were derived for each dimension by multiplying the respective value for each feature (e.g., concreteness for tail) by the total number of participants who produced that feature then dividing by the total number of features for each cue word. Table 1 reflects the final set of derived variables in this summary dataframe.

**Table 1. Variables for Cue-Feature Contrasts**

| Variable | Description |
|---|---|
| cue | Cue word (lemmatized) |
| cue_cnc | Concreteness value of the cue word |
| cue_familiar | Familiarity value of the cue word |
| cue_frequency | Frequency (log10) per million words of the cue word |
| cue_neighbors | Semantic neighborhood density of the cue word |
| feat_agree_m | Mean percent agreement for all features associated with a given cue |
| feat_cnc_m | Arithmetic weighted mean of concreteness for all features of a given cue |
| feat_count_m | Mean number of features produced by participant per cue word |
| feat_familiar_m | Arithmetic weighted mean of familiarity for all features of a given cue |
| feat_freq_m | Arithmetic weighted mean lexical frequency all features of a given cue |
| feat_neighbor_m | Arithmetic weighted mean semantic density all features of a given cue |

### Data Analysis Procedures

We first produced a Pearson correlation matrix to examine bivariate relations between cue and feature attributes. We used these data both for descriptive purposes and to exclude highly correlated variables that might produce multicollinearity problems in regression models. We used a threshold of R>.70 and excluded one variable within a highly correlated pair (see below). After isolating a set of predictors with low to moderate strength correlations, we executed the following models:

1) feat_cnc_m ~ cue_cnc: Average concreteness of the features as explained by concreteness of the cue

2) feat_count_m ~ cue_cnc: Mean number of features produced by a participant per cue as explained by concreteness of the cue
3) feat_agree_m ~ cue_cnc: Mean percent agreement for all features associated with a particular cue as explained by concreteness of the cue

## Results

### Bivariate Correlations Between Cues and Features

Figure 1 represents a correlation matrix detailing all statistically significant ($p<.05$) bivariate relationships across cue and predictor variables. Figure 1 is noteworthy for several associations ordered in descending strength from: 1) cue_concreteness ~ feature_concreteness (R=.83); 2) cue_freq ~ cue_neighbors (R=.72); 3) feature_freq ~ feature_neighbors (R=.69); and 4) cue_freq ~ cue_familiar (R=.64). Based on the strength of these relationships, we excluded the semantic neighborhood density of the cue (cue_neighbors) and concreteness of the feature (feat_cnc_m) from the multiple regression analyses to follow.

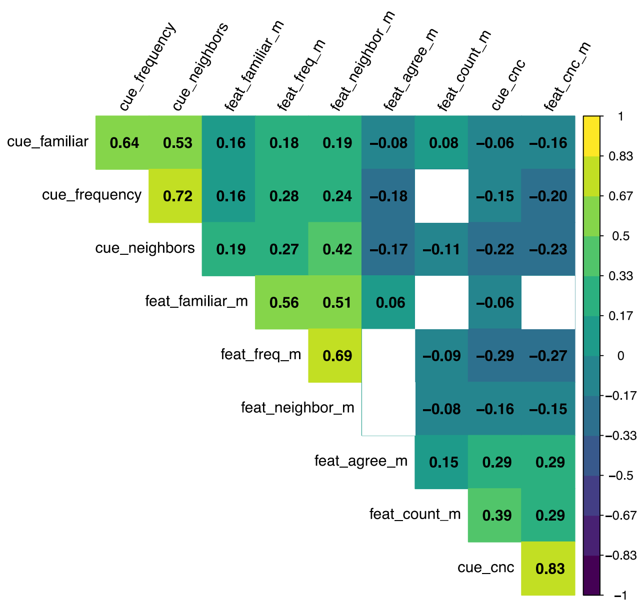


Figure 1: Correlation matrix of cue-feature predictor relationships. **Note**: Values reflect Pearson correlation coefficients. Variables are clustered by similarity using the hierarchical clustering algorithm within the corrplot R package (Wei & Simko, 2016). Blank cells represent non-statistically significant correlation coefficients.

## Individual Predictor Regression Models

Table 2 represents output of the three separate single predictor regression models described to follow.

### Model 1: Feature Concreteness ~ Cue Concreteness

Model 1 tested prediction of feature concreteness (i.e., arithmetic weighted mean of the concreteness values of all features within a given cue) by cue concreteness. The overall model was statistically significant with a large proportion of variance explained $R^{2=}.68$ [F (1,4188) = 9039, $p < .001$]. The relation between cue and feature concreteness was linear as confirmed by a rainbow test [rain stat: 0.94, df1 = 2095, df2 = 2093, p = .90], and the correlation between cue and feature concreteness was strongly positive (R=.83). The y intercept of this model was 2.78 ($p<0.001$).

### Model 2: Mean Feature Count by Participant ~ Cue Concreteness

Model 2 tested prediction of the average feature count across participants as a linear function of the concreteness of the cue. The Pearson correlation between the mean number of features produced per participant and concreteness of the cue was R=.39 (small to moderate effect). The overall model was statistically significant [F(1,4188) = 368.71, $p < .001$], but cue concreteness explained little of the variance of feature counts R2=.15.

### Model 3: Mean % Feature Agreement ~ Cue Concreteness

Model 3 tested prediction of average feature agreement as a function of the concreteness of the cue (e.g., if 50 of 60 participants converged on 'tail' when cued with dog, then tail would have elicited 83% agreement). The Pearson correlation between average % feature agreement and cue concreteness was R=.29 (a small effect). Although the overall model was statistically significant [F(1,4188) = 741.70, $p < .01$], cue concreteness explained only a very small proportion of the variance $R^{2=}.08$ of feature agreement/convergence across participants.

Table 2. Single Predictor OLS Regression Results

| | Model 1 Feature Concreteness | Model 2 Feature Count | Model 3 Feature Agreement |
|---|---|---|---|
| (Intercept) | 2.78 *** | 2.75 *** | 28.82 *** |
| | (0.04) | (0.04) | (0.13) |
| Cue Concreteness | 1.49 *** | 1.01 *** | 2.42 *** |
| | (0.02) | (0.04) | (0.13) |
| N | 4190 | 4190 | 4190 |
| $R^2$ | 0.68 | 0.15 | 0.08 |

All continuous predictors are mean-centered and scaled by 1 standard deviation. The outcome variable is in its original units. (SE) in parentheses. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

**Model 4 Simultaneous Multiple Regression: Feature Concreteness ~ Predictors**

In Model 4, we examined prediction of the concreteness of features produced for each cue as a function of seven variables using the following linear model specification:lm(feature_cnc~feature_neighborhood_density + feature_word_frequency + feature_familiarity + feature_%agreement + feature_count + cue_word_frequency + cue_familiarity + cue_concreteness). Table 3 reflects output of the regression.

Model 4 was statistically significant [F(8,4138) = 1236.95, p < .0001] accounting for 70% of the variance of concreteness of the features [adjusted $R^2$ =.70].

Table 3. Multiple regression predicting feature concreteness

| | feat_cnc ~ predictors |
|---|---|
| (Intercept) | 6.45 (0.02) *** |
| Feature neighborhood density (mean) | -0.00 (0.02) |
| Feat word frequency (mean) | -0.12 (0.02) *** |
| Feature familiarity (mean) | 0.16 (0.02) *** |
| Feature count (mean by participant) | -0.05 (0.02) ** |
| Feature % agreement (mean by participant) | 0.09 (0.02) *** |
| Cue word frequency | 0.03 (0.02) |
| Cue word familiarity | -0.22 (0.02) *** |
| Cue concreteness | 1.45 (0.02) *** |
| N | 4147 |
| $R^2$ | 0.71 |
| Log likelihood | -5806.45 |
| AIC | 11632.91 |

All continuous predictors are mean-centered and scaled by 1 standard deviation. Coefficients standardized, and SE is in parentheses. The outcome variable is in its original units. *** p < 0.001; ** p < 0.01; * p < 0.05.

## Discussion

Strong sensorimotor dominance in cue-feature relations would be evident as independence of their respective concreteness values. If semantic feature generation were consistently biased toward concrete attributes, the correlation between cue concreteness and feature concreteness would be negligible. That is, feature concreteness would be uniformly high across the spectrum of cue concreteness. The data support the opposite effect: concreteness of the cue was strongly predictive of feature concreteness (R = .83, $R^2$ = .68). Moreover, when we ran a multiple regression that included other lexical variables to which concreteness effects are often attributed (familiarity, frequency, semantic neighborhood density), they only accounted for an additional 2% of the variance (R2 = .70).

However, this is a situation in which the intercept of the linear model is interpretable. That is, for cue words with a value of zero (i.e., highly abstract concepts), the y intercept is 2.78 (p<0.001). This indicates that even for the most abstract words, the features produced are more concrete than the cue itself, and the trend increases linearly from there – at each point along the x-axis (cue concreteness), the y values (feature concreteness) are higher than the x values.

Taken together, these results suggest that there is both a sensorimotor bias in feature generation as evidenced by the intercept, but also that salience in a feature generation task may be driven in part by equivalent concreteness such that when we produce features for *dog,* we are drawn to *tail* rather than *loyal* and when we think of features for *liberty,* we are drawn to *freedom* rather than *bell*.

Sensorimotor bias in cue-feature relations would predict that participants produce more features for concrete than abstract words, due to the comparative abundance of informative sensorimotor features for concrete concepts.

We found a low-moderate positive correlation between cue concreteness and the average number of features produced per participant (R= 0.39, $R^2$= 0.15), suggesting that factors besides sensorimotor salience better predict ease of feature generation. Individual differences in overall task performance could account for some of the missing variance; some participants may produce only a few features even for frequent, highly familiar, highly concrete cues like *dog*, while others may be garrulous for all cue types. It is also possible that differences in experimental design across data sources for the original database could cause variation in how many features were produced. For example, McRae and colleagues provided participants with a form with ten lines per cue, while Buchanan and colleagues' procedure was open-ended (McRae et al., 2005; Buchanan et al., 2019). Finally, data cleaning likely removed some valid but idiosyncratic features, as only those features that had at least 16% consensus across participants or were in the top five most frequent features listed for a cue were retained (Buchanan et al., 2019). It is conceivable that many concrete but low-frequency features were discarded, especially for cues that generated a large number of features across participants.

None of the above explanations, however, address underlying cognitive explanations for the remaining variance. Semantic feature production is a task that indexes semantic memory and underlying semantic structure, so it is perhaps not surprising that ease of feature production (as indexed by number of features produced) is not strongly predicted by a single semantic variable. In future work, we plan to expand the model to include semantic variables such as bigram semantic distance and type of relationship (thematic vs. taxonomic) both between cues and features as well as among features as predictors of number of features produced.

We hypothesized that concreteness would predict interparticipant agreement for features: if people employ an imagistic "read off" strategy when generating features, there would likely be more consensus for concrete words than for abstract ones. In fact, the data did not support this: there was only a weak correlation between interparticipant feature agreement and concreteness (R = 0.29). This could be caused by lower-than-expected agreement for concrete cues, perhaps because there is more modality-specific information from which to choose: *dogs* have *tails*, but also *paws*, *fur*, *eyes*, *ears*, *legs*, *teeth*, *tongues*, *drool*, and *collars* (Schwanenflugel et al., 1992). A given participant may only produce a subset of these available sensorimotor features for each cue, resulting in a more diffuse feature pool. Alternately, agreement may be easier to attain for abstract cues, perhaps due to strong contextual embedding (*justice* is strongly associated with *freedom* in the context of the legal system) or relatively low competition in semantic space during feature retrieval resulting in less diffusion and higher consensus among participants.

Our prediction of a sensorimotor-biased cognitive strategy for semantic feature generation was borne out by the data, as evidenced by the y-intercept of the initial linear model, but we also found strong cohesion of concreteness during feature production such that feature concreteness increases linearly with cue concreteness. Whether this is due to equivalent concreteness conferring a higher baseline activation strength or to equivalent concreteness being used as an evaluative strategy when generating a feature list is unknown. One limitation of the cue-feature data set we used in these analyses is the lack of information about the order in which features were produced. Including order information in future analyses could be used to investigate generation strategies with more granularity. Future work will also include more semantic variables such as bigram semantic distance and type of semantic relationship to further characterize semantic feature generation both as a commonly used behavioral task in psycho-/neurolinguistics and as an index of underlying semantic representations.

## References


Barsalou, L. W., Kyle Simmons, W., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in cognitive sciences*, *7*(2), 84–91. https://doi.org/10.1016/s1364-6613(02)00029-3

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577–609; discussion 610-660. https://doi.org/10.1017/s0140525x99002149

Borghi, A. M. (2022). Concepts for which we need others more: The case of abstract concepts. *Current Directions in Psychological Science*, *31*(3), 238–246. https://doi.org/10.1177/09637214221079625

Borghi, A. M., Capirci, O., Gianfreda, G., & Volterra, V. (2014). The body and the fading away of abstract concepts and words: A sign language analysis. *Frontiers in Psychology*, *5*(July), 811. https://doi.org/10.3389/fpsyg.2014.00811

Brysbaert, M., Martínez, G., & Reviriego, P. (2025). Moving beyond word frequency based on tally counting: AI-generated familiarity estimates of words and phrases are an interesting additional index of language knowledge. *Behavior Research Methods*, *57*(1), 1–15. https://doi.org/10.3758/s13428-020-01497-y

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/41/4/977 [pii] 10.3758/BRM.41.4.977 [doi]

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911.

Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods*, *51*(4), 1849–1863. https://doi.org/10.3758/s13428-019-01243-z

Crutch, S. J., Troche, J., Reilly, J., & Ridgway, G. R. (2013). Abstract conceptual feature ratings: The role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Frontiers in Human Neuroscience*, *7*, 186. https://doi.org/10.3389/fnhum.2013.00186

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627.

Dove, G. (2022). *Abstract concepts and the embodied mind: Rethinking grounded cognition*. Oxford University Press. https://books.google.com/books?hl=en&lr=&id=4dV2EAAAQBAJ&oi=fnd&pg=PP1&dq=guy+dove+2022&ots=dBz7O7hf8X&sig=kSR_SjtCScPBE6Td89Ayw15T1_c

Gao, C., Shinkareva, S. V., & Desai, R. H. (2022). SCOPE: The South Carolina psycholinguistic metabase. *Behavior Research Methods*, *55*(6), 2853–2884. https://doi.org/10.3758/s13428-022-01934-0

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. https://doi.org/10.1037/a0027373

Kousta, S.-T. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology General*, *140*(1), 14–34. https://doi.org/2010-26153-001 [pii] 10.1037/a0021446 [doi]

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*(4), 547–559.

McRae, K., De Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*(2), 99.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology / Revue canadienne de psychologie, 45*(3), 255–287. https://doi.org/10.1037/h0084295

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*(8), 1407–1419. https://doi.org/10.1162/jocn.2007.19.8.1407

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: The dynamic multilevel reactivation framework. *Psychonomic Bulletin & Review*, *23*(4), 1002–1014. https://doi.org/10.3758/s13423-015-0824-5

Schwanenflugel, P. J., Akin, C., & Luh, W.-M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, *20*(1), 96–104. https://doi.org/10.3758/BF03208259

Shaoul, C., & Westbury, C. (2010). Exploring lexical co-occurrence space using HiDEx. *Behavior Research Methods*, *42*(2), 393–413.

Troche, J., Crutch, S., & Reilly, J. (2014). Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Frontiers in Psychology*, *5*, 360. https://doi.org/10.3389/fpsyg.2014.00360

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex (New York, N.Y. : 1991)*, *24*(7), 1767–1777. https://doi.org/10.1093/cercor/bht025

Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., & Borghi, A. M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *Journal of Memory and Language*, *116*, 104173. https://doi.org/10.1016/j.jml.2020.104173

Wei, T., & Simko, V. (2016). *Corrplot: Visualization of a Correlation Matrix* (Version R package version 0.77) [Computer software]. https://CRAN.R-project.org/package=corrplot

Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D., Software, P., & PBC. (2023). *dplyr: A Grammar of Data Manipulation* (Version 1.1.2) [Computer software]. https://cran.r-project.org/web/packages/dplyr/index.html

Wiemer-Hastings, K., & Xu, X. (2005). Content Differences for Abstract and Concrete Concepts. *Cognitive Science*, *29*(5), 719–736. https://doi.org/10.1207/s15516709cog0000_33