COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 38 (2014) 162–177

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12076

A Quantitative Empirical Analysis of the Abstract/Concrete Distinction

Felix Hill,[a] Anna Korhonen,[a] Christian Bentz[b]

[a]*Computer Laboratory, University of Cambridge*
[b]*Department of Theoretical and Applied Linguistics, University of Cambridge*

Received 3 August 2012; received in revised form 25 January 2013; accepted 31 January 2013

Abstract

This study presents original evidence that abstract and concrete concepts are organized and represented differently in the mind, based on analyses of thousands of concepts in publicly available data sets and computational resources. First, we show that abstract and concrete concepts have differing patterns of association with other concepts. Second, we test recent hypotheses that abstract concepts are organized according to association, whereas concrete concepts are organized according to (semantic) similarity. Third, we present evidence suggesting that concrete representations are more strongly feature-based than abstract concepts. We argue that degree of feature-based structure may fundamentally determine concreteness, and we discuss implications for cognitive and computational models of meaning.

Keywords: Psychology; Computer science; Cognitive architecture; Concepts; Representation; Semantics; Concreteness

All abstract sciences are nothing but the study of relations between signs.
—Denis Diderot, *D'Alembert's Dream* (1769).

1. Introduction

A large body of empirical evidence indicates important cognitive differences between abstract concepts, such as *guilt* or *obesity,* and concrete concepts, such as *chocolate* or *cheeseburger*. It has been shown that concrete concepts are more easily learned than abstract concepts (Caramelli, Setti, & Maurizzi, 2004; Gentner, 1982; Paivio, 1971), more easily remembered (Begg & Paivio, 1969), and that language referring to concrete versus

Correspondence should be sent to Felix Hill, Computer Laboratory, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, CB3 0FD Cambridge, UK. E-mail: fh295@cam.ac.uk

abstract concepts is more easily processed (De Groot, 1992; Holmes & Langford, 1976; James, 1975; Moeser, 1974; Schwanenflugel, 1991; Schwanenflugel & Shoben, 1983). Moreover, there are cases of brain damage in which either abstract or concrete concepts appear to be specifically impaired (Breedin, Saffran, & Coslett, 1994; Tyler, Moss, & Jennings, 1995; Warrington, 1975). In addition, functional magnetic resonance imaging studies implicate overlapping but partly distinct neural systems in the processing of the two concept types (Binder, Westbury, McKiernan, Possing, & Medler, 2005; Wilson-Mendenhall, Simmons, Martin, & Barsalou, 2013), and topographical differences between abstract and concrete concepts in event-related potential components have been reported (Adorni & Proverbio, 2012; Huang, Lee, & Federmeier, 2010). Despite these widely known findings, however, there is little consensus on the cognitive basis of the observed differences (Paivio, 1986; Schwanenflugel, 1991). Indeed, while many studies of conceptual representation and organization focus on concrete domains (Gopnik & Schulz, 2004; Medin, 1989; Taylor, Devereux, & Tyler, 2011; Taylor, Moss, & Tyler, 2007), comparatively little has been established empirically about abstract concepts (Barsalou, 1999; Barsalou & Wiemer-Hastings, 2005).[1]

In this study, we test various theoretical claims concerning the abstract/concrete distinction by exploiting large publicly available experimental data sets and computational resources. By analyzing thousands of abstract and concrete concepts, our approach marginalizes potential confounds more robustly than in smaller scale behavioral studies. In Analysis 1, we show that abstract concepts are associated in the mind to a wider range of other concepts, although the degree of this association is typically weaker than for concrete concepts. In Analysis 2, we explore the basis of these associations by testing the hypothesis that similarity predicts association for concrete concepts to a greater extent than for abstract concepts. In Analysis 3, we show that free association is a more symmetric relation for abstract concepts than for concrete concepts. The findings together suggest contrasts in both the organization and representation of abstract and concrete concepts. We conclude by discussing the implications of the findings for existing theories and models of conceptual representation.

2. Data

2.1. University of South Florida (USF) norms

All three experimental analyses use the USF Free-association Norms (Nelson & McEvoy, 2012). The USF data consist of over 5,000 words and their associates. In compiling the data, more than 6,000 participants were presented with cue words and asked to

> write the first word that comes to mind that is meaningfully related or strongly associated to the presented word.

For a cue word c and an associate a, the Forward Association Probability (*FAP*) from c to a is the proportion of participants who produced a when presented with c. *FAP* is thus a measure of the strength of an associate *relative to other associates of that cue*.

Many of the cues and associates in the USF data have a concreteness score, derived from either the norms of Paivio, Yuille, and Madigan (1968) or Toglia and Battig (1978). In both cases, contributors were asked to rate words based on a scale of 1 (very abstract) to 7 (very concrete).[2]

2.2. WordNet

WordNet is a tree-based lexical ontology containing over 155,000 words produced manually by researchers at Princeton University (Felbaum, 1998). The present work used WordNet version 3.0.

2.3. Brown corpus

Word frequencies were extracted from the 1-million-word Brown Corpus (Kucera & Francis, 1967), chosen because it is an American corpus compiled at a similar time to the USF data. Word tokens in the Brown Corpus are tagged for their part of speech (POS). For any word type, it is therefore possible to extract the *majority POS* (the POS with which the type is most frequently tagged).

3. Analyses

3.1. Analysis 1: Patterns of association

3.1.1. Motivation

The *Context Availability Model* (1991, Schwanenflugel & Shoben, 1983) is designed to provide a theoretical basis for the aforementioned empirical differences between abstract and concrete concepts. Specifically, the model proposes that abstract concepts are more difficult to process because it is more difficult to retrieve information (prior knowledge) associated with such concepts than information associated with concrete concepts. Schwanenflugel's exposition of the model relies on the following hypothesis.[3]

(H1) Abstract concepts have more but weaker connections (to other concepts) than concrete concepts.

Schwanenflugel (1991) presents only small-scale behavioral experiments in support of H1, testing between 40 and 67 participants with 64 concepts (see also Schwanenflugel & Shoben, 1983). In Analysis 1 we test H1 on a far larger data set.

3.1.2. Method

We extracted those 3,255 pairs in the USF data for which the concreteness of the cue word was known. Since cue words are connected to a finite set of associates by *FAP* values, we can isolate a probability distribution over associates for each cue. Since our measure of association strength (*FAP*) is relative, it is not possible to compare these strengths directly across cue words. Nonetheless, we can make certain inferences about absolute cue-associate strength from properties of the *FAP* distributions. If a cue has many associates with little variance in the *FAP* distribution, each *FAP* value must necessarily be low (and absolute association strength intuitively weak). In contrast, for a given number of associates, higher variance implies that some *FAP* values are notably higher than the mean, and thus likely to be strong in an absolute sense. Therefore, to address H1, we considered both the dimension (number of associates) and the variance of the *FAP* distribution for each cue word.

In an initial analysis of the data, we noted a moderate but significant negative correlation between concreteness and frequency, $r(3255) = -.16$, $p < .001$. Therefore, a multiple regression analysis was conducted with *log*(Frequency), Number of Associates and Variance of *FAP* as predictors, and Concreteness as dependent variables. Because the Concreteness/Frequency multicollinearity was exacerbated by high-frequency abstract prepositions and verbs, a second analysis was conducted solely over those cue words with "noun" as majority POS ($n = 2{,}320$).

3.1.3. Results and discussion

Both regression models were statistically significant, explaining 17% of the variance of Concreteness in each case. The beta coefficients in Table 1 indicate that the concreteness of a cue word correlates negatively with both the number of associates and the variance in *FAP* over these associates. Both variables are highly significant predictors of concreteness even when controlling for frequency as an independent predictor.

The analysis confirms that abstract words have both more associates than concrete words and lower variance in *FAP* distributions, which is consistent with the idea that the strength of their associates is on average weaker than for concrete words. Fig. 1 represents these two effects visually. While this confirmation of H1 is consistent with Schwanenflugel's Context Availability model, it is also compatible with other theoretical

Table 1
Multiple regression analysis of Concreteness

	All Words		Nouns Only	
	Coeff. (β)	t	Coeff. (β)	t
# Associates	−0.04***	−16.70	−0.04***	−15.97
Variance	18.01***	5.85	15.64***	4.41
log(Freq)	−0.18***	−14.21	−0.12***	−7.87
	$R^2 = .17$, $F(3, 3196) = 211.82$***		$R^2 = .17$, $F(3, 2319) = 157.51$***	

***$p < 0.001$.

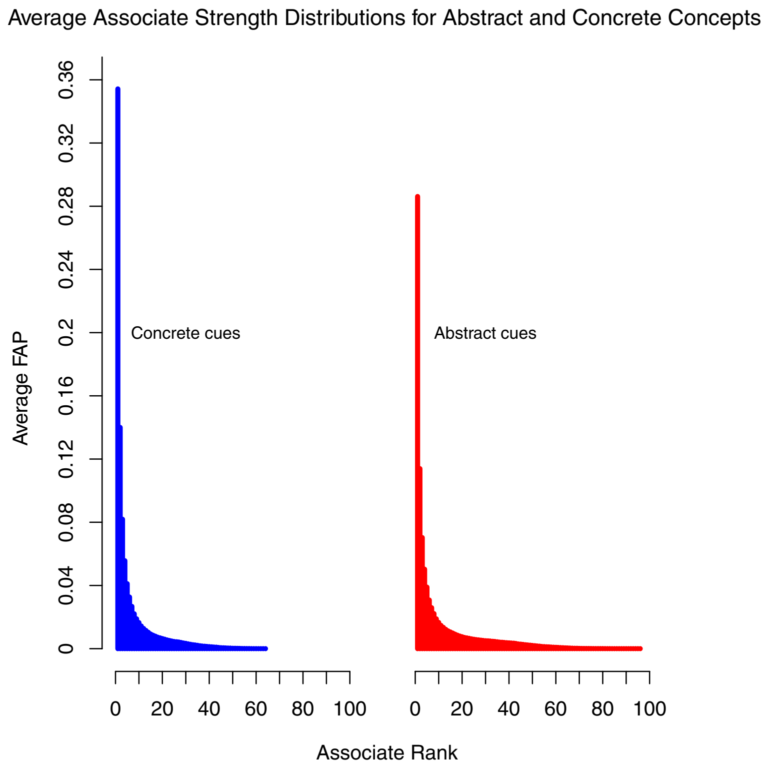

Fig. 1. Average *FAP* mass at associate ranks 1–100 for the 500 most abstract and concrete cue words in the USF data. Note the stronger initial associates in the concrete case and the longer tail of weak associates in the abstract case.

characterizations of abstract/concrete differences. For instance, the *Dual Coding Theory* (Paivio, 1986) posits that concrete concepts are represented via two distinct knowledge encoding systems, visual and verbal, whereas abstract concepts are represented only in the verbal system. Under this assumption, the associates of concrete concepts should be particularly strong, since associations can be reinforced via connections in two independent representational systems. Markman and Stilwell's (2001) distinction between *feature-based categories* and *relational categories* also provides a framework by which to explain H3. According to Markman and Stilwell, feature-based categories, including those noun concepts typically considered highly concrete, are represented by (featural) information "*subordinate to*" or "*contained within*" that representation (2001, p. 330), whereas relational categories, which include abstract noun, preposition, verb, and event categories, are defined by external information such as the position of the representation in a relational structure. By this account, concrete associations would be particularly strong because they are reinforced by the salience of common features. On the other hand, the associates of a given abstract concept would include all potential neighbors in the relational structure encoded by that concept's representation and would thus be particularly numerous. Since the findings of Analysis 1 appear to be consistent with each of these theoretical perspectives, we investigate the distinction in more detail in Analyses 2 and 3.

3.2. Analysis 2: Distinct conceptual organization?

3.2.1. Motivation

On the basis of recent behavioral studies of healthy and brain-damaged subjects, (Crutch, Connell, & Warrington, 2009; Crutch, Ridha, & Warrington, 2006; Crutch & Warrington, 2005, 2010), Crutch and colleagues argue that abstract and concrete concepts differ "*qualitatively*" in how they relate to other concepts. More specifically, they propose the following:

> (H2) Concrete concepts are organized in the mind according to similarity, whereas abstract concepts are organized according to association.[4]

The terms *association* and *similarity* refer to the ways the concept pairs [*car, bike*] and [*car, petrol*] are related: *Car* is said to be (semantically) similar to *bike* and associated with (but not similar to) *petrol*. Intuitively, *car* and *bike* may be understood as similar because of their common physical features (wheels), their common function (transport), or because they fall within a clearly definable category (modes of transport). In contrast, *car* and *petrol* may be associated because they often occur together or because of the functional relationship between them (McRae, Khalkhali, & Hare, 2012; Plaut, 1995).[5] The two relations are neither mutually exclusive nor independent; *bike* and *car* are related to some degree by both association and similarity.

In support of H2, Crutch et al. (2009) asked 20 participants to select the odd-one-out when lists of five words appeared on a screen. The lists comprised either concrete or abstract words (based on ratings of six informants) connected either by similarity (e.g., *dog*, *wolf*, *fox*; *theft*, *robbery*, *stealing,* etc.) or association (*dog, bone, collar,* etc.; *theft, law, victim,* etc.), with an unrelated odd-one-out item in each list. Controlling for frequency and position, subjects were both significantly faster and more accurate if the related words were either abstract and associated or concrete and similar. These results support H2 on the basis that decision times are faster when the related items form a more coherent group, rendering the odd-one-out more salient. Related experiments on brain-damaged subjects produced similar findings (Crutch & Warrington, 2010; Crutch et al., 2006)

Despite the consistency in these findings, each of Crutch et al.'s experiments tests a small sample of subjects (<20) with a small (<20) number of concepts. It is therefore possible that the observed differences resulted from semantic factors particular to the subjects and items but independent of concreteness. Analysis 2 exploits the USF data and WordNet to investigate H2 more thoroughly.

3.2.2. Method

Because similarity and association are not mutually exclusive, H2 can be interpreted in terms of differing interactions between these two relation types. If concrete concepts are organized in the mind to a greater extent than abstract concepts according to similarity, then the associates of a given concrete concept should be more similar to that concept than the associates of a given abstract concept. In other words, there should be greater correlation between similarity and association over concrete concepts than abstract. We test for this effect with a multiple regression over cue-associate pairs, with *FAP* as dependent variable (representing strength of association) and Concreteness, *SIM*, and their interaction as predictors. Relevant to H2 is the presence or absence of a positive interaction between concreteness and similarity.

Following other studies of conceptual structure (Markman & Wisniewski, 1997), we model similarity as proximity in a conceptual taxonomy, in this case, WordNet. Various measures of similarity have been developed for WordNet (e.g., Resnik, 1999). *PathSim*, based on the shortest path between two senses, is perhaps the simplest, and it mirrors the manual approach taken by Markman and Wisniewski (1997).[6] For this experiment, *SIM*, a measure of the similarity of two words w_1 and w_2 on the range [0, 1], was defined as the maximum *PathSim* between all senses of w_1 and all senses of w_2. Since verbs, adjectives and nouns occupy separate taxonomic structures in WordNet, PathSim does not effectively quantify similarity across these categories. We therefore restrict our analysis to those 18,668 pairs in the *FAP* data for which cue concreteness and *FAP* are known and in which the majority POS for both words is "noun."

As a pre-test, *SIM* was evaluated on Rubenstein and Goodenough's (1965) similarity data for 65 word pairs,[7] previously used as a benchmark for automatic similarity measures. The correlation between these judgments and *SIM*, $r(63) = .77$, $p < .05$, was comparable to other more complex WordNet metrics such as Resnik's (1995) Information Content, $r(63) = .79$, $p < .05$, and approaching the human replication baseline, $r(63) = .90$ (Resnik, 1995).

3.2.3. Results and discussion

As detailed in Table 2, the regression model was significant, $F(2, 3252) = 194.53$, and, as expected, *SIM* was a significant predictor of *FAP*. The interaction term *SIM*:Concreteness was positive, as predicted by H2, and a significant predictor of *FAP*.

The positive interaction between similarity and concreteness in our model lends some support to H2. However, the size of this effect is small: The model explains less than 0.1 of a percentage point more variance in *FAP* than a model with no interaction term. While statistically significant, this difference is not consistent with a "*qualitative difference*" in

Table 2
Multiple regression analysis of *FAP* over cue (noun)—associate (noun) pairs

	Coeff. (β)	*t*-value
SIM	0.048	3.66***
Concreteness	0.003	1.64
SIM:Conc	0.005	2.07*
	$R^2 = .03$, $F(3, 18665) = 194.53$	

*$p < 0.05$; ***$p < 0.001$.

conceptual organization between abstract and concrete concepts, as Crutch and Warrington (2005) propose. Rather, our analysis supports a gradual contrast in patterns of organization along a continuum from concrete to abstract. Of course, qualitative or categorical differences may exist that are too subtle to be detected by the current method. We intend to examine this possibility in future work, using the USF data and WordNet to generate appropriate items for larger scale behavioral experiments.

3.3. Analysis 3: Distinct conceptual representation?

3.3.1. Motivation

Hypothesis H2 (Analysis 2) characterizes the abstract/concrete distinction in terms of conceptual organization. With respect to the differences in representation that cause the H2 effect, Crutch and Warrington offer only speculative hypotheses. For instance, they suggest that "*abstract concepts are represented in associative neural networks*," while "*concrete concepts have a categorical organization*" (Crutch & Warrington, 2005; p. 624). Wiemer-Hastings and Xu (2005) address this question empirically, by asking subjects to write down properties or features of both abstract and concrete words. Nevertheless, given the untimed, conscious nature of their feature-generation task, and the fact they test only 31 subjects with 36 concepts, the strength of their findings is limited in a similar way to those of Crutch et al. In Analysis 3, we test for evidence of specific representational differences that could explain H2 and the other concreteness effects detailed in the Introduction.

Although the limitations of classical theories of representation based on strict binary property specifications are well known (Prinz, 2002), many recent theories characterize representations as *feature-based* in a more dynamic sense (McRae, de Sa, & Seidenberg, 1997; Plaut & Shallice, 1993; Tyler & Moss, 2001; Wu & Barsalou, 2009). Indeed, the idea of concepts as complexes of conceptually basic features underlines explanations of various empirical observations, including typicality effects (Rosch, 1975), category learning (Rogers & McLelland, 2003), and category-specific semantic impairments (Tyler et al., 1995).

Feature-based models are not ubiquitous. Competing approaches such as spatial models (Landauer & Dumais, 1997; Shepard, 1957) or associative networks (Quillian, 1968; Steyvers & Tennembaum, 2005) have also captured various established cognitive phenomena. One criticism of such models, however, is that they naturally model relatedness with a symmetric operation: for all concepts x and y, $relatedness(x, y) = relatedness(y, x)$. As often observed (Griffiths, Steyvers, & Tenembaum, 2007; Steyvers & Griffiths, 2007; Tversky, 1977), empirical measures of conceptual proximity are in general asymmetric. For instance, it is common to find concept pairs X and Y for which subjects judge the statement "X is like Y" to be more acceptable than "Y is like X." This effect can be particularly evident when one concept is more salient than the other ("Justin Bieber is like Elvis" vs. "Elvis is like Justin Bieber?") or more prototypical ("an ellipse is like a circle" vs. "a circle is like an ellipse?"). Asymmetries are also observed in priming effects and free association, for instance with category name/member or whole/part pairs ("Alsatian" primes "dog" more than "dog" primes "Alsatian").

A noted strength of feature-based models is that they naturally capture the asymmetry of semantic relations. In the *Contrast Model*, Tversky (1977) proposes that the similarity of conceptual representations is computed as some continuous function of their common and distinctive features. Such operations are generally asymmetric, particularly given a disparity in the number of features. For instance, suppose the concept *jackal* is represented with the features {*4LEGS, FUR, HOWLS*} and the concept *dog* with the features {*4LEGS, FUR, TAIL, COLLAR, LOYAL, DOMESTIC*}. According to Tversky, the fact that it is more natural to say that *jackals* are like *dogs* than *vice versa* derives from the fact that two thirds of *jackal* features are shared by *dog*, whereas only one third of *dog* features are shared by *jackal*. As with other theories of representation mentioned previously, Tversky's demonstrations are typically confined to concrete words. Nevertheless, his conclusions could be aligned with H2 (Analysis 2) if the following hypothesis held:

> (H3) Concrete representations have a high degree of feature-based structure, whereas abstract representations do not.

Indeed, the soundness of H3 could point to a causal explanation of the H2 effect. By H3, computing similarity between abstract concepts by mapping features would be harder than computing similarity between concrete concepts in this way. Alternative types of semantic relation would therefore be required to group collections of abstract concepts in the mind.

Proposals similar to H3 have been made by several researchers. Plaut and Shallice (1993) showed that integrating differential degrees of feature-based structure into connectionist simulations of dyslexia leads to more accurate replication of established concrete word advantages. In addition, Markman and Stilwell's (2001) distinction between feature-based (concrete noun) and relational (verb and abstract noun) categories is entirely consistent with H3. Finally, H3 is also compatible with the feature-generation study of Wiemer-Hastings and Xu (2005), in which subjects tended to generate fewer "*intrinsic*" and proportionally more "*relational*" properties for abstract concepts.

In Analysis 3, we exploit the USF data to test a prediction of H3. If Tversky's demonstration that asymmetry derives from features is sound, there should be greater asymmetry between concrete concepts than between abstract concepts.

3.3.2. Method

Although Tversky's reasoning pertains to a similarity relation, we use the USF data to explore asymmetries in association. Similarity is an important factor in association in general, as evidenced by the high *SIM*/*FAP* correlation (Analysis 2). We therefore expect asymmetries deriving from similarity to be reflected in *FAP* values, noting free-association has been shown to be an asymmetric relation in previous studies (Michelbaker, Evert, & Schütze, 2011).

For each of the 18,668 unordered cue-associate pairs [c, a] for which the concreteness of c and a is known, we calculate the (additive) asymmetry $|FAP[c, a] - FAP[a,c]|$. We define the total cue asymmetry, CueAsymm(c), as the sum of the additive asymmetries over all associates of that cue. For a given cue item in our analysis, we experiment with three different measures of concreteness. The first is the cue concreteness Conc(c). Since Tversky's explanation of asymmetry relies on both concepts in a given pair having a feature-based representation, for each pair [c,a] we also calculate both the sum and the product of concreteness scores. We then define ConcSum(c) as the sum of the sums over all associates, $ConcSum(c) = \sum_{\alpha}$ Conc(c) + Conc(a), and ConcProd(c) as the sum of products $ConcProd(c) = \sum_{\alpha}$ Conc(c) + Conc(a). To control for the possibility that *FAP* asymmetries are caused exclusively by a disparity in frequency between cue and associate, we define the measure FreqDisp(c); the sum of the absolute differences between the frequency of a cue word and that of each of its associates, $FreqDisp(c) = \sum_{a} |\mathrm{Freq(a)} - \mathrm{Freq(c)}|$. We analyze the relationship between CueAsymm (dependent variable) and the three measures of concreteness (predictors) in separate multiple regression models, with FreqDisp as an independent predictor in each.

3.3.3. Results and discussion

The results in Table 3 reveal a significant positive correlation between the concreteness measure and CueAsymm in all three models, confirming the prediction of H3. Moreover, the model with ConcProd (R^2 = 13.73) accounts for more of the CueAsymm variance than the model with ConcSum (R^2 = .12), which in turn accounts for more than Conc (R^2 = .08). These two comparisons show that information about the concreteness of both cue and associate is important for predicting asymmetry, consistent with Tversky's explanation of how asymmetry arises from a feature-mapping comparison. It is also notable that FreqDisp is a (marginally) significant predictor in only one of the three models. Therefore, the predictive relationship between concreteness and asymmetry does not derive from discrepancies in frequency between the words in a pair. The strength of this predictive relationship is illustrated in Fig. 2.

Table 3
Multiple regression analyses of CueAsymm

	Coeff. (β)	*t*
Conc	0.001***	16.28
FreqDisp	−0.000	−1.44
	R^2 = .08, F(2, 3252) = 135.60***	
ConcSum	0.003***	21.33
FreqDisp	−0.000*	−2.43
	R^2 = .12, F(2, 3252) = 230.92***	
ConcProd	0.001***	22.60
FreqDisp	−0.000	−0.39
	R^2 = .14, F(2, 3252) = 258.81***	

Note. *p < .05; ***p < .001.

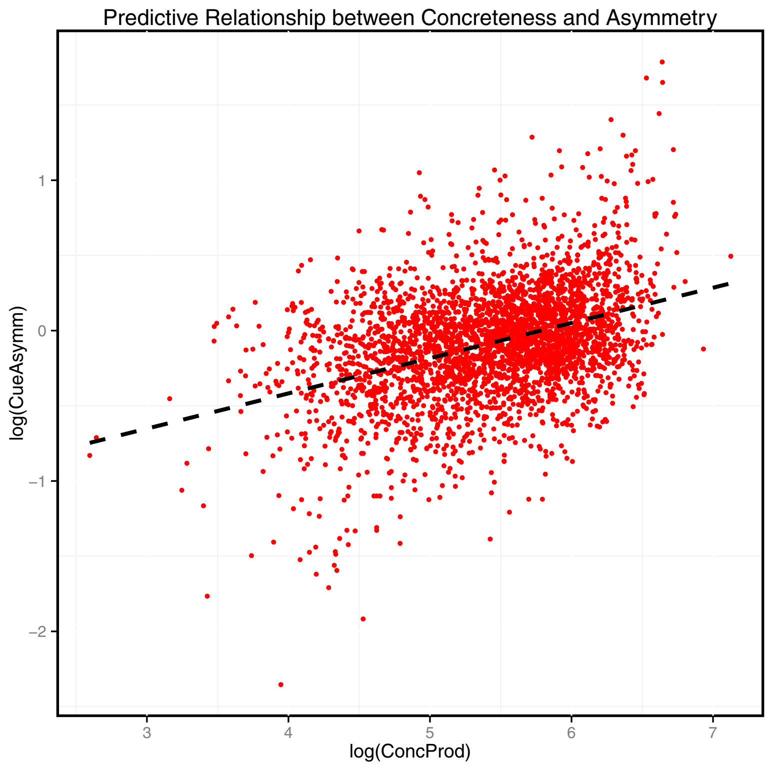

Fig. 2. Scatterplot of CueAsymm versus ConcProd, illustrating the significant positive correlation between concreteness and asymmetry as predicted by H3.

In a separate analysis, we observed that the ConcProd model restricted to noun cues (R^2 = .13) is better than the model restricted to all non-nouns ($\mathbf{R^2}$ = 0.10) or just verbs (R^2 = .11). Indeed, across all 18,668 pairs, the mean additive asymmetry when both cue and associate are nouns (.071) is significantly greater than when both are not (.066), $t(9351.3)$ = 2.78, $p < .01$. Together with Tversky's analysis, these observations support Markman and Stilwell's proposal that many noun representations are feature-based, whereas representations of verbs and prepositions rely on features to a lesser extent.

4. Conclusion

In this study, we have reported the following effects of increasing conceptual concreteness:

1. Fewer but stronger associates (Analysis 1).
2. A stronger correlation between the similarity of concepts and the strength of association between them (Analysis 2).
3. Greater asymmetry of association between concepts (Analysis 3).

These findings derive from analyses of thousands of concepts and data from thousands of subjects, an approach that significantly increases their robustness in comparison with previous behavioral experiments.

Finding 3 is consistent with, and arguably suggestive of, the view that concrete representations are more strongly feature-based than abstract representations. Instead of a strong feature-based structure, abstract representations encode a pattern of relations with other concepts (both abstract and concrete). We hypothesize that the degree of feature-based structure is the fundamental cognitive correlate of what is intuitively understood as concreteness.

By this account, computing the similarity of two concrete concepts would involve a (asymmetric) feature comparison of the sort described by Tversky. In contrast, computing the similarity of abstract concepts would require a (more symmetric) comparison of relational predicates such as analogy (Gentner & Markman, 1997; Markman & Gentner, 1993). Because of their representational structure, the feature-based operation would be simple and intuitive for concrete concepts, so that similar objects (of close taxonomic categories) come to be associated. On the other hand, for abstract concepts, perhaps because structure mapping is more complex or demanding, the items that come to be associated are instead those that fill neighboring positions in the relational structure specified by that concept (such as arguments of verbs or prepositions).[8] Intuitively, this would result in a larger set of associates than for concrete concepts, as confirmed by Finding 1. Moreover, such associates would not in general be similar, as supported by Finding 2.

If this is correct, it is likely that computational models of meaning could be improved by integrating a dimension of concreteness. For instance, distributional models that connect words via syntagmatic co-occurrence (e.g., Turney & Pantel, 2010) would be particularly appropriate for modeling free association in abstract domains. WordNet-based measures, or those based on paradigmactic co-occurrence (Ruiz-Casado, Alfonseca, & Castells, 2005), would better reflect similarity and be more apt for concrete domains. These observations could be applicable to a range of natural language processing tasks, including word-sense disambiguation, semantic role labeling, and metaphor analysis. Moreover, to effectively model human cognition across domains, a combination of both approaches may be required. Indeed, it has already been demonstrated that cognitive models that learn from a combination of "experiential" (featural) and distributional (co-occurrence) data produce more realistic representations than those that learn from either data type in isolation (Andrews, Vigliocco, & Vinson, 2009). Our conclusions suggest that concreteness could provide a principled basis for determining the optimum balance between these two information sources.

Linguists and psychologists have long sought theories that exhaustively capture the empirical facts of conceptual meaning. Approaches that fundamentally reflect association, such as semantic networks and distributional models, struggle to account for the reality of categories or prototypes. On the other hand, certain concepts evade satisfactory characterization within the framework of prototypes and features, the concept *game* being a prime example, as Wittgenstein (1953) famously noted. The differences between abstract and concrete concepts highlighted in this and other recent work might indicate why a general theory of concepts has proved so elusive. Perhaps we have been guilty of trying to find a single solution to two different problems.

Acknowledgments

Thank you to Barry Devereux, Bodo Winter, and John Williams for helpful discussions and comments. The study was supported by St John's College, the Royal Society, and the Arts and Humanities Research Council.

Notes

1. Notwithstanding a large body of theoretical work (see e.g., Lakoff & Johnson, 1980; Markman & Stilwell, 2001).
2. Although concreteness is well understood intuitively, it lacks a universally accepted definition. In empirical work, it is usually described in terms of reference to sensory experience (Paivio et al., 1968; Toglia & Battig, 1978). However, intuitively it also connected to specificity; *rose* is often said to be more concrete than *flora*. In this study, we follow the existing empirical characterization but note the need for a clearly defined construct in future studies.
3. For example, she states, "what is important to this view is not how abstract words come to have weaker connections [to associated information]…only that they generally do" (Schwanenflugel, 1991, p. 243).
4. This hypothesis is referred to as the *Qualitatively Different Representation Hypothesis* by Duñabeitia, Avilés, Afonso, and Scheeper (2009).
5. We do not prescribe formal definitions of association or similarity, but rather work with empirical characterizations, identifying association with the free-association task and similarity with proximity in a taxonomy.
6. Empirical measures of similarity are not in general symmetric (see Analysis 3), whereas the predominant WordNet-based measures, including PathSim, are symmetric (Lin, 1998; Resnik, 1999). We use PathSim in the present work because asymmetric measures are not widely implemented or subjected to the same empirical validation (e.g., Schickel-Zuber & Faltings, 2007). This does not affect the conclusion that associates of concrete concepts are more likely to be taxonomically close than associates of abstract concepts.
7. Subjects were asked to consider their idea of synonymy and then rate the "similarity of meaning" of word pairs (Rubenstein & Goodenough, 1965, p. 628).
8. According to Markman and Gentner, "in analogy, only relational predicates are shared, whereas in literal similarity, both relational predicates and object attributes are shared" (1997, p. 48). In these terms, our proposal is that concreteness corresponds with the degree to which concepts are compared literally versus by analogy.

References

Adorni, R., & Proverbio, A. (2012). The neural manifestation of the word concreteness effect: An electrical neuroimaging study. *Neuropsychologia*, *50*(5), 880–891.

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating experiential and distributional data to learn semantic represenations. *Psychological Review*, *116*(3), 463–498.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–609.

Barsalou, L., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129–163). New York: Cambridge University Press.

Begg, I., & Paivio, A. (1969). Concreteness and imagery in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 821–827.

Binder, J., Westbury, C., McKiernan, K., Possing, E., & Medler, D. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*(6), 905–917.

Breedin, S., Saffran, E., & Coslett, H. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology*, *11*, 617–660.

Caramelli, N., Setti, A., & Maurizzi, D. (2004). Concrete and abstract concepts in school age children. *Psychology of Language and Communication*, *8*(2), 6–9.

Crutch, S., Connell, S., & Warrington, E. (2009). The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgments. *Quarterly Journal of Experimental Psychology*, *62*(7), 1377–1388.

Crutch, S., Ridha, B., & Warrington, E. (2006). The different frameworks underlying abstract and concrete knowledge: Evidence from a bilingual patient with a semantic refractory access dysphasia. *Neurocase*, *12* (3), 151–163.

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

Crutch, S., & Warrington, E. (2010). The differential dependence of abstract and concrete words upon associative and similarity-based information: Complementary semantic interference and facilitation effects. *Cognitive Neuropsychology*, *27*(1), 46–71.

De Groot, A. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 1001–1018.

Duñabeitia, J., Avilés, J., Afonso, A., & Scheeper, O. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*, 284–292.

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

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gentner, D., & Markman, A. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*, 45–56.

Gopnik, A., & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences*, *8*(8), 372–377.

Griffiths, T., Steyvers, M., & Tenembaum, J. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211–244.

Holmes, M., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Verbal Learning & Verbal Behavior*, *15*, 559–566.

Huang, H., Lee, C., & Federmeier, K. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage*, *49*(1), 1116–1123.

James, C. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception & Performance*, *104*, 130–136.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

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

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*, 11–140.

Lin, D. (1998). An information-theoretic definition of similarity. Proceedings of the Fifteenth International Conference on Machine Learning (ICML '98), Madison, WI.

Markman, A., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language*, *32*, 517–535.

Markman, A, & Stilwell, C. (2001). Role-governed categories. *Journal of Theoretical and Experimental Artificial Intelligence*, *13*, 329–358.

Markman, A., & Wisniewski, E. (1997). Similar and different: The differentiation of basic level categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*(1), 54–70.

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

McRae, K., Khalkhali, S., & Hare, M. (2012). Semantic and associative relations: Examining a tenuous dichotomy. In V. Reyna, S. Chapman, & J. Dougherty (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 39–66). Washington, DC: American Psychological Association.

Medin, D. (1989). Concepts and conceptual structure. *American Psychologist*, *44*, 1469–1481.

Michelbaker, L., Evert, S., & Schütze, H. (2011). Asymmetry in corpus-derived and human word associations. *Corpus Linguistics and Linguistic Theory*, *7*(2), 245.

Moeser, S. (1974). Memory for meaning and wording in concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior*, *13*, 682–697.

Nelson, D., & McEvoy, C. (2012). *The University of South Florida Word Association, Rhyme and Word Fragment Norms*. Available at: http://web.usf.edu/FreeAssociation/Intro.html. Accessed August 1, 2012.

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

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

Paivio, A., Yuille, J., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, *76*(1, Pt. 2), 1–25.

Plaut, D. (1995). Semantic and associative priming in a distributed attractor network. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, (pp. 37–42) Mahwah, NJ: Lawrence Erlbaum Associates..

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

Prinz, J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.

Quillian, M. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). Cambridge, MA: MIT Press.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. *Proceedings of IJCAI-95*, 448–453, Montreal, Canada.

Resnik, P. (1999). Semantic similarity in a taxonomy: An information based measure and its application to problems of ambiguity in natural language. *Journal of Artificial Intelligence Research*, *11*, 95–130.

Rogers, T., & McLelland, J. (2003). *Semantic cognition*. Cambridge, MA: MIT Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), (September 1975), 192–233.

Rubenstein, H., & Goodenough, J. (1965). Contextual correlates of synonymy. *Communications of the ACM*, *8*(10), 627–633.

Ruiz-Casado, M., Alfonseca, E., & Castells, P. (2005). Using context-window overlapping in synonym discovery and ontology extension. *Proceedings of RANLP-2005*, Borovets, Bulgaria.

Schickel-Zuber, V., & Faltings, B. (2007). Using hierarchical clustering for learning the ontologies used in recommendation systems. *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* ACM, San Jose, CA.

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., & Shoben, E. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *9*, 82–102.

Shepard, R. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*, 325–345.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Latent semantic analysis: A road to meaning* (pp. 427–448). Hillsdale, NJ: Erlbaum.

Steyvers, M., & Tennembaum, J. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78.

Taylor, K., Devereux, B., & Tyler, L. (2011). Conceptual structure: Towards an integrated neuro-cognitive account. *Language and Cognitive Processes*, *26*(9), 1368–1401.

Taylor, K., Moss, H., & Tyler, L. (2007). The conceptual structure account: A cognitive model of semantic memory and its neural instantiation. In J. Hart & M. Kraut (Eds.), *Neural basis of semantic memory* (pp. 265–301). London: Cambridge University Press.

Toglia, M., & Battig, W. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.

Turney, P, & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research (JAIR)*, *37*, 141–188.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Tyler, L., & Moss, H. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*(6), 244–252.

Tyler, L., Moss, H., & Jennings, F. (1995). Abstract word deficits in aphasia: Evidence from semantic priming. *Neuropsychology*, *9*(3), 354–363.

Warrington, E. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*(4), 635–657.

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*(5), 719–736.

Wilson-Mendenhall, C., Simmons, W., Martin, A., & Barsalou, L. (2013). Contextual processing of abstract concepts reveals neural representations of non-linguistic semantic content. *Journal of Cognitive Neuroscience*, *25*(6), 920–935.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford, England: Blackwell.

Wu, L., & Barsalou, L. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, *132*, 173–189.