J Psycholinguist Res (2017) 46:715–729
DOI 10.1007/s10936-016-9459-6

The Paradox of Abstraction: Precision Versus Concreteness

Rumen Iliev[1,2] · **Robert Axelrod**[1]

Published online: 22 November 2016

Abstract We introduce a novel measure of abstractness based on the amount of information of a concept computed from its position in a semantic taxonomy. We refer to this measure as *precision*. We propose two alternative ways to measure precision, one based on the path length from a concept to the root of the taxonomic tree, and another one based on the number of direct and indirect descendants. Since more information implies greater processing load, we hypothesize that nouns higher in precision will have a processing disadvantage in a lexical decision task. We contrast precision to *concreteness*, a common measure of abstractness based on the proportion of sensory-based information associated with a concept. Since concreteness facilitates cognitive processing, we predict that while both concreteness and precision are measures of abstractness, they will have opposite effects on performance. In two studies we found empirical support for our hypothesis. Precision and concreteness had opposite effects on latency and accuracy in a lexical decision task, and these opposite effects were observable while controlling for word length, word frequency, affective content and semantic diversity. Our results support the view that concepts organization includes amodal semantic structures which are independent of sensory information. They also suggest that we should distinguish between sensory-based and amount-of-information-based abstractness.

Keywords Abstraction · Concreteness · WordNet · Precision · Information · Entropy

Introduction

The distinction between concreteness and abstractness is central for many areas of cognitive and social psychology (Rosch et al. 1976; Trope and Liberman 2010). Yet, there is little agreement about how to define abstractness (Barsalou 2003) and how to study it (Burgoon

✉ Rumen Iliev
riliev@stanford.edu

1 Gerald R. Ford School of Public Policy, University Of Michigan, Ann Arbor, MI, USA

2 Present Address: Stanford University, Stanford, CA, USA

et al. 2013). In this paper we argue that treating abstractness as a unitary construct leads to a logical paradox where both more abstract and less abstract concepts have cognitive advantages. To support our claim, we juxtapose the two most common interpretations of abstractness, one that measures abstractness of a concept in terms of its lack of physicality, and the other based on hierarchical organization of information content. According to the first interpretation, more concrete concepts should have various cognitive advantages because they can be directly experienced, and sensory information is easier to process than non-sensory information. Thus, the word "beagle" should be easier to process than the word "animal" because it is more concrete. According to the second interpretation, however, more abstract concepts should be easier to process because they contain less information. Under this view, "animal" should be cognitively less costly since it encodes less information compared to "beagle". If these predictions are empirically supported, we can conclude that the view of abstractness as a unitary theoretical construct leads to logical conundrum, where both more concrete and more abstract concepts are easier to process.

The first approach defines the process of abstraction as distancing from physical existence and sensory experience. Concrete concepts represent entities that can be seen or touched; they are material, exist in time and space, and can physically interact with other concrete objects. Abstract concepts, on the other hand, represent entities that lack these properties: they cannot be seen or touched, and cannot be part of physical causal chains. According to this approach/classification, abstract objects are only mental, while concrete objects are both mental and physical. From this perspective, the definition of abstraction presupposes the existence of mental representation. Arguably, the apple on my table will still exist even if there are no minds in the universe to perceive it, while the same might not be true for democracy, justice, or love. Since most of research using this sense of abstractness has operationalized it as the converse of concreteness, from here on when we discuss this aspect of abstractness as the opposition of material to mental we will refer to it as *concreteness*.

Typically, concreteness is measured as "directness of reference to sense experience" (Paivio et al. 1968). More concrete concepts have a plethora of cognitive advantages, which collectively has been referred to as the "concreteness effect". Compared to less concrete words, more concrete words are more easily recalled (Walker and Hulme 1999), more easily associated (Paivio et al. 1968), and are recognized faster in lexical decision tasks (Kroll and Merves 1986); they are also richer in meaning (Toglia and Battig 1978), appear earlier in vocabulary development (Gilhooly and Logie 1980) and are less likely to be lost due to brain damage (Jefferies et al. 2009). While the concreteness effect has become a classic textbook example, its theoretical explanation (Shallice and Cooper 2013; Connell and Lynott 2012), its boundary conditions (Kousta et al. 2011), and its neural underpinnings (Wang et al. 2010) are still a matter of debate.

The second approach to abstractness that we will consider is related to the scope, or inclusiveness, of a concept, and we will refer to it as *precision*.[1] In this interpretation, abstractness is a dimension along the general-specific axis, where greater abstractness is defined as greater generality and inclusiveness (Rosch 1978). In Fig. 1, for example, the concept of dog is less inclusive than the concept of mammal, but more inclusive than the concept of beagle.

Most of the work on precision focuses on the so called basic-level category effect. Rosch et al. (1976) suggested that in terms of hierarchically organized categories, middle level of precision (e.g. the node "dog" in Fig. 1) is cognitively privileged when compared to both superordinate (the node "mammal") or subordinate levels (the nodes "beagle" and "poodle").

[1] We chose the term precision rather than the terms inclusiveness, generality or specificity because it more closely reflects the notion of amount of information in a hierarchical system.

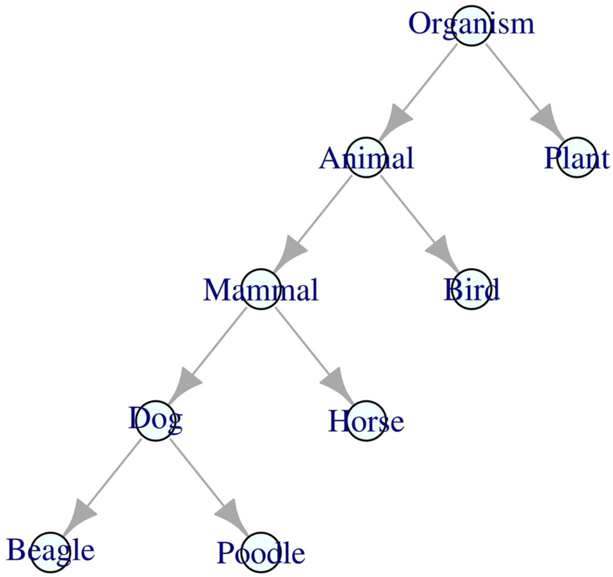

Fig. 1 A hypothetical taxonomy based on an IS-A relationship. Precision increases when moving down the tree, with “beagle” and “poodle” being the most precise concepts and “organism” the least precise

Words representing this basic level tend to be shorter, more frequently used as labels, and learned earlier in life. They also generate more items in feature listing tasks, and are activated faster in picture naming tasks (for a recent review see Hajibayova 2013). While the basic-level advantage is empirically robust, the major challenge for this framework is how to determine which level is basic. Early definitions relied on similarities between objects based on physical features, visual perception, and motor programs for handling these objects, but this approach was challenged by the findings that the most privileged level varies with learning and expertise (Tanaka and Taylor 1991; Johnson and Mervis 1997). Other definitions, those based on maximizing cue validity, between-category differentiation, and informativeness were no more successful, since these measures increase or decrease monotonically as a function of inclusiveness (Medin 1983). In this paper we consider precision in terms of a lack of inclusiveness,[2] and do not address the non-monotonic effects associated with basic levels.

Within the framework of hierarchical organization of concepts, precision is closely related to the amount of information associated with a concept. The more precise a concept is, the more information it provides, since all the information contained in a parent node is inherited by a child node, but the opposite is not true. For example, if we know that something is an animal, we can expect that it is multicellular, that it can probably move, and that it needs food. Yet we will be rather uncertain about its size and appearance. If we also know that it is a dog, we obtain additional information, which will decrease our uncertainty about the object's features, and if we further know that it is a beagle, our inferences will be even more precise.

[2] It is also important to note that precision is not limited to category structure and is relevant to any variable defined in terms of inclusiveness. As reviewed in Burgoon et al. (2013); Vallacher and Wegner (1987) defined abstraction in terms of comprehensiveness, Semin and Fiedler (1991) in terms of enduring qualities, and Watkins et al. (2008) in terms of essential gists and invariance. In this sense, precision can also refer to spatio-temporal inclusiveness, but we do not address this aspect here.

In such a case we can guess its color, weight, and temperament. How exactly this increase of information is represented in the mind is still an open question (see Murphy 2004, chapter 7), yet from a formal perspective an increase in precision is always accompanied by increase in the amount of information.[3]

Since both concreteness and precision have been used as measures of abstractness, it is somewhat surprising that there is relatively little work that directly compares the two constructs. Intuitively, concepts lower on precision should also be less concrete. Thus the two measures should be statistically related. For example, while it is easy to imagine how a dog looks and sounds it is harder to visualize a generic mammal, and even harder to visualize a living thing. Initially, early work supported the intuitive link between concreteness and precision, finding that similarly to more concrete concepts, more precise concepts are easier to visualize, easier to be associated in a pair-association task (Paivio and Olver 1964), faster to be imagined (Paivio 1966) and are recalled better in a free recall task (Paivio 1967). In addition, the two measures were positively correlated. Spreen and Schulz (1966) asked their participants to provide subjective ratings of nouns on two dimensions, abstract-concrete and general-specific, and found that the two measures were correlated at 0.63. Yet, other analyses indicated that the two measures might not stem from the same construct. Paivio and Olver (1964) and Paivio (1966) found that more concrete nouns produce a greater number of associations, while precision had no effect on that variable. Further, Paivio (1968) factor-analyzed 30 word-related measures and found that precision and concreteness loaded on different factors. Overall, the results were found to be puzzling (Paivio 1971, p.83), with concreteness and precision being seemingly closely related, and at the same time being somewhat independent.

In the current paper, we show that concreteness and precision should be treated as independent theoretical constructs rather than as different measures of the same construct (abstractness). We propose that when concreteness and precision are defined in terms of information, there is a clear dissociation between the two. Concreteness is mainly related to the proportion of sensory-based information, while precision is mainly related in the amount of information based on concept's position in a semantic hierarchy. Importantly, this distinction between type versus quantity of information can lead to the surprising prediction that the two measures can have opposite effects on cognitive performance. More concrete concepts should be easier to process because sensory-based information has cognitive advantages. When focusing on the amount of information, however, stimuli with less information have less cognitive load and should be easier to process (see Frank 2013 for a recent discussion). From this perspective, since concepts lower in precision contain less information, they should be cognitively advantageous. In other words, we predict a "precision effect": more abstract words should be easier to process.

To test our hypothesis, we contrast the effect of concreteness and precision on cognitive performance, where cognitive performance is operationalized as latency and accuracy in a lexical decision task. The rest of the paper is organized as follows: In the next section we introduce two different measures of precision. In "Study 1" section we present results from a study where we test if latency and accuracy can be predicted by our precision measures. In a second study, presented in "Study 2" section, we test the generalizability of the results from the previous study by controlling for affective content and semantic diversity, factors which were previously shown to be related both to abstractness and to cognitive processing. Last, we discuss the theoretical and methodological implications of our results.

[3] Each additional level of precision can be seen as adding new sub-partitions to a set, and it is easy to show that the amount of information gained from knowing to which sub-partition an element belongs, increases with the number of sub-partitions (cf. Meilă 2007).

Measuring Precision

While different measures of concreteness and cognitive performance in lexical decision tasks are well established, and large-scale databases are easily available, there are no corresponding databases containing measures of precision. Therefore, we define our own methods for measuring precision. Our notion of precision relies on the existence of a hierarchically structured semantic network where concepts are organized in a directed rooted tree. While concept organization can include many types of relationships which are not hierarchically organized, the IS-A relationship is typically thought to have a tree-like structure, and therefore can be used for operationalizing precision. Take for example the hypothetical tree presented in Fig. 2, where node *a* is the root, and all other nodes are its direct or indirect descendants connected via an IS-A relationship.

We define precision as a function P applied to concepts organized in a directed rooted tree:

$$\text{If c1 is a parent of c2, then } P(c1) < P(c2) \tag{1}$$

Such a simple definition has some appealing properties. We could define the root of the tree as the node which has no parent, and all other nodes are its direct or indirect descendants. Further, we could also define the leaves of the tree (also known as terminal nodes) as the nodes which have no children. Definition (1), then, implies that the root has the lowest precision and that precision is monotonically increasing when traversing a path from the root to a leaf. Note, however, that while this function allows us to compare the precision of nodes which are positioned on the same path linking the root to a leaf (e.g. $a - b - e$ in Fig. 2) it does not allow us to compare the precision of nodes which are not part of the same paths (e.g. comparing nodes b to c or f to c). Accordingly, we further define two more specific measures of precision, *semantic inclusiveness* and *semantic depth*, which will allow us to compare any two nodes in the tree.

Semantic Inclusiveness

In a classical paper Shannon (1948) proposed that information can be quantified as reduction of uncertainty, where the amount of information we gain from an event is a negative function of the likelihood of that event. Rare events bring more information while frequent events bring less information. Resnik (1995) extended this approach to hierarchically organized semantic networks. He proposed an information content measure which is based on the likelihood of an entity c being classified as a member of a particular class. A beagle, for example, can be classified as a beagle, a dog or as an animal, while dog can be classified as a dog or as an animal, but not as a beagle. Since more entities can be classified as a dog than as a beagle, the concept of dog has higher likelihood and therefore lower information content. We use this approach to define precision of a concept based on its inclusiveness. If a concept c has n descendents (both direct and indirect) in a tree that has a total of N nodes, then:

$$\text{P-inclusiveness (c)} = -\log((n+1)/N) \tag{2}$$

In Fig. 2, for example, an entity which can be classified as e, can also be classified as b or a, and there are more entities which can be classified as a than as b. The more inclusive a node is, the less precise it is. Since node e has no descendants it is counted as 1 (itself) and since we have a total of eight nodes, the likelihood of e is $1/8$. If we use 2 as a base of the logarithm, then P-inclusiveness(e) = 3. Similarly, P-inclusiveness (b) = 1.42 and P-inclusiveness (a) = 0.

There is an important difference between Resnik's approach and ours. Specifically, in addition to using the semantic network structure when computing information content Resnik also incorporates word frequencies as a weight for each node. Since word frequencies have observable effect on cognitive performance in lexical decision tasks, we do not use them when estimating P-inclusiveness. Instead, we base our measure exclusively on the structure of the tree. It is also important to note that while our measure of P-inclusiveness stems from information theory, it is very different from recent measures of word information which are based on word co-occurrences in text (Piantadosi et al. 2011; Frank 2013; Mahowald et al. 2013), rather than on the semantic organization of the corresponding concepts.

Semantic Depth

The P-inclusiveness measure follows from our basic definition of precision and captures the intuition that nodes with many descendants should be considered less precise than nodes with fewer descendants, since in the former case there is more remaining uncertainty than in the latter case. One problem with this definition of precision, however, is that it treats all leaves as having maximal precision regardless of their position in the tree. For example, P-inclusiveness for nodes *c* and *e* in Fig. 2 is the same. This could be a problem, since cultural knowledge in one domain might be much richer than in another domain, yet concepts at the lowest levels of abstraction in both domains will be treated as equally precise. One way to overcome this problem is to introduce another measure of precision, which is based on the distance of a node to the root, counting the number of edges to the top. If the function |c1, c2| measures the shortest path between concepts c1 and c2, then:

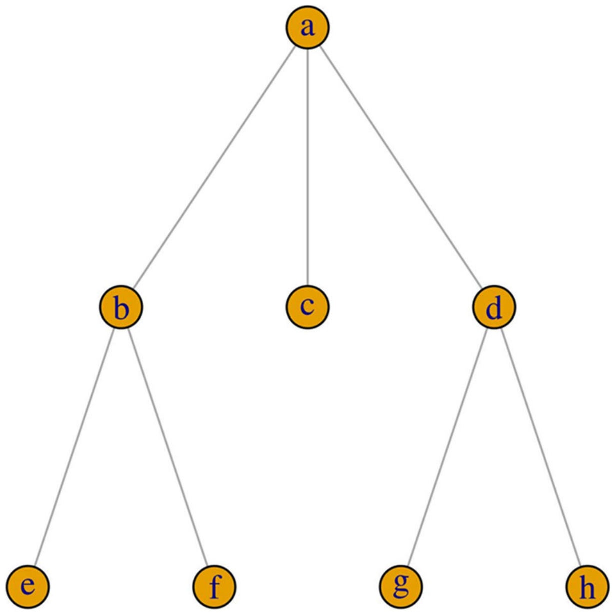

Fig. 2 A hypothetical tree representing an IS-A semantic hierarchy. Node *a* is the root, and nodes *e*, *f*, *g* and *h* are leaves. Precision increases when a node is further away from the root (P-depth) and when it has fewer descendants (P-inclusiveness). The root is the highest level of abstraction according to both measures of precision

$$\text{P-depth}(c) = |c, \text{root}| \tag{3}$$

Both P-depth and P-inclusiveness follow our definition of precision described in (1) and capture the increase in the amount of information with lower levels of abstractness.

The two measures of abstractness will coincide when each non-terminal node has the same number of children and all leaves are equidistant from the root. However, they will differ when there is a sufficient variance in the number of children per node or in the distances between the leaves and the root. Since we do not have a priori reason to expect that one of the two measures will be more closely linked to cognitive performance, we use them both.

In the next part we describe the database which we use to estimate P-depth and P-inclusiveness of English nouns.

WordNet

We computed our measures of precision based on data from WordNet (Miller 1995) version 3.1, implemented in the NLTK package (Bird et al. 2009). WordNet is a semantic network which includes information about IS-A relationships for nouns. The nouns network can be viewed as a taxonomy with 82,125 nodes and a single root labeled "entity.n.01". As an illustration, Fig. 3 presents the network structure of all nodes descending from the concept "animal.n.01". Each node represents a concept, and each concept is associated with one or more words (referred to as lemmas). When there was more than one path linking a node to the root (e.g. a dog is both a mammal and a pet), we followed Resnik (1995) and computed P-inclusiveness for that node based on the path with highest information content. When a word was polysemous (having multiple meanings), such as "bank", we took the average values of the different precision measures for each of the nodes associated with the word.

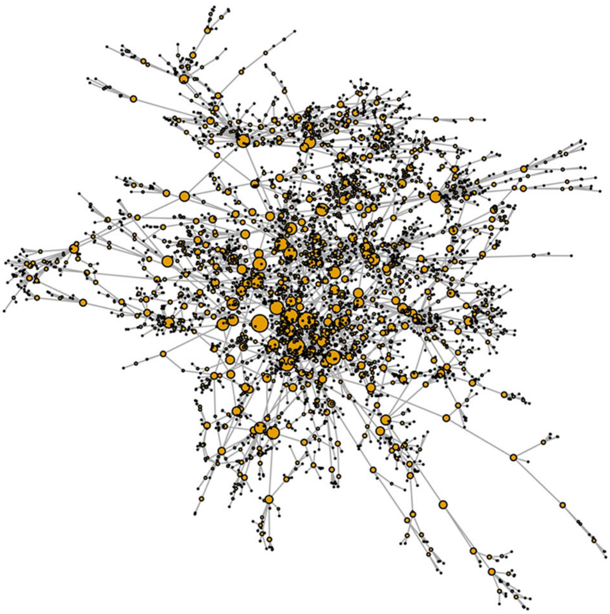

Fig. 3 The part of the WordNet IS-A hierarchy depicting the node "animal.n.01" and its descendants. The size of the nodes is proportional to the number of the descendants. Smaller nodes have higher P-inclusiveness and nodes further from the center have higher P-depth

Before we continue with the results, it is important to note that using WordNet as a source for psychological measurements has both pros and cons. One of its main advantages is the great number of words/concepts included in it. For some of our tests the number of words we were able to use exceeded 10,000. For comparison, when juxtaposing the effects of precision and concreteness (Spreen and Schulz 1966) used 325 words, while Paivio (1968) used only 96 words. Another advantage is that the WordNet taxonomy provides a complex semantic structure which spans 18 levels of inclusiveness. In contrast, in previous studies researchers had either asked for subjective ratings on the specificity/generality dimension, or had directly manipulated up to three levels of inclusiveness. A major disadvantage of using WordNet as an approximation for the semantic organization in the mind is that there is a substantial subjective component—it is based on the judgements of a small number of experts, and is influenced by their theoretical positions and cultural perspectives. However, since we do not see how such a theoretical or a cultural bias will favor our hypothesis (that decrease in precision will enhance cognitive performance), we believe that the advantages of using WordNet to address the theoretical question about the different measures of abstractness significantly outweighs its potential disadvantages.

Study 1

One of the most commonly cited piece of evidence for the "concreteness effect" comes from the fact that in lexical decision tasks less concrete words have longer reaction times (latencies) and lower accuracy rates. In this study we seek initial evidence that precision, when compared to concreteness, will have opposite effects on cognitive performance. More specifically, we expect that less precise words will have shorter latencies and greater accuracy.

Datasets

Concreteness We used norms collected by Brysbaert et al. (2014). The data contain concreteness ratings for 60,000 English words and currently is the richest dataset of its type available. To collect their data, the authors asked human judges to rate the words along the concrete-abstract dimension, describing concrete words as those which one "can experience directly through one of the five senses". Abstract words were defined as those "that cannot be experienced directly but which we know because the meanings can be defined by other words". The words were rated on a 5 point Likert scale, where 1 was labeled as "abstract" and 5 was labeled as "concrete".

Precision We measured P-inclusiveness and P-depth using the procedures described above.

Cognitive performance Balota et al. (2007) present latency and accuracy measures from a lexical decision task for 40,481 words. In a lexical decision task participants have to answer if a particular string of characters is a word or not. Longer latencies and lower accuracy are typically interpreted as lower cognitive performance.

Word frequency Since word frequency is typically a strong predictor of cognitive performance in various lexical task, in our analyses we control for word frequency. Our measure is based on the norms from the SUBTLEX-US database (Brysbaert and New 2009). These norms were derived from movie and TV series subtitles and explain more variance in various psychological variables than much larger datasets. We used log transformation with base 10 of the original frequency data.

Table 1 Summary of the regression models used in Study 1

	Mode 1: Latency	Model 2: Latency	Model 3: Latency	Mode 4: Accuracy	Model 5: Accuracy	Model 6: Accuracy
Concreteness	−0.05*		−0.07*	0.06*		0.07*
	(0.008)		(0.008)	(0.009)		(0.01)
P-inclusiveness		0.08*	0.08*		−0.06*	−0.04*
		(0.008)	(0.008)		(0.008)	(0.01)
P-depth		0.05*	0.07*		−0.04*	−0.05*
		(0.008)	(0.008)		(0.007)	(0.01)
Length	0.05*	0.03*	0.05*	0.20*	0.27*	0.19*
	(0.009)	(0.008)	(0.009)	(0.01)	(0.008)	(0.01)
Frequency	−0.66*	−0.59*	−0.63*	.0.59*	0.65*	0.57*
	(0.009)	(0.008)	(0.009)	(0.01)	(0.008)	(0.011)
df	8436	10,983	8434	8436	10,992	8343
R-squared	0.46	0.41	0.48	0.28	0.41	0.29

The regression coefficients are standardized beta weights, with standard errors of the estimate provided in parenthesis

* $p < .001$

Results

First, we examined the relationship between our three predictors. While concreteness was correlated with P-depth at medium level (r(18,946) = 0.36, p < .001) it was only weakly correlated with P-inclusiveness (r(18,946) = -0.02, p < .05). Further, P-depth and P-inclusiveness were also relatively weakly correlated at r(114,190) = 0.16, p < .001) (all zero-order correlations are presented in the Appendix). Next, we tested if concreteness and precision are related to latency. We run a series of OLS regression models, using latency as a dependent variable and the two different measures of abstractness as predictors. We also controlled for basic word properties by including word frequencies and orthographic lengths in the equation. The results are presented in Table 1. In Model 1 we replicated the classical concreteness effect, finding that concreteness negatively predicts latency. In Model 2 we focused on precision and found that both measures of precision, P-inclusiveness and P-depth, positively predict latency. This was the first evidence that precision is related to cognitive performance. In model 3 we entered simultaneously concreteness and the two precision measures, and, replicating the results from the previous two models, we found that precision and concreteness had opposite effects on latency. While higher concreteness speeds up lexical decisions, higher precision slows them down. Next, we run the same series of analyses using accuracy as the dependent variable (Models 4–6). The results were virtually the same: concreteness consistently predicted better cognitive performance, while both precision measures consistently predicted worse cognitive performance.

The results from these analyses provide strong support for the hypothesis that precision and concreteness have opposite effects on cognitive performance. The analyses of both latency and accuracy converged toward the same conclusion, namely that concreteness helps cognitive performance in lexical decision tasks, while precision hinders it. Further, we also observed that the two measures of precision are not highly correlated, and have independent effects on cognitive performance in the predicted direction.

Study 2

In Study 1 we found support for our hypothesis for diverging effects of precision and concreteness on cognitive performance while controlling for two basic psycholinguistic properties: word frequency and word length. In the current study we further test the robustness of these results by controlling for two additional measures which have been linked to abstraction and cognitive performance: affective content and semantic diversity.

Kousta et al. (2009) found that words with affective associations, regardless if positive or negative, are processed faster. Kousta et al. (2011) reported that when controlling for other psycholinguistic properties (imageability and context availability) the concreteness effect gets reversed and more abstract words become cognitively privileged. The authors explained this reversal with the greater affective content of abstract words, which was the only consistent predictor across their analyses. Further, Hoffman et al. (2013) found that semantic diversity is negatively related to concreteness, and Hoffman and Woollams (2015) found that greater semantic diversity predicts shorter latencies in a lexical decision task. Since there is evidence that both affective content and semantic diversity are related to abstraction and to cognitive performance, it is important to know how these measures are related to precision, and if they can account for the results from Study 1. In the current study we test if the reversed effects of concreteness and precision will hold when controlling for affective content and semantic diversity.

Method

Datasets

In addition to the datasets used in the previous study we also used:

Affective content We used the Word Valence variable from the Warriner et al. (2013) dataset. The words in this variable were rated on a 1–9 scale, where 1 was negative and 9 positive, with 5 being the neutral point. Because we were not interested in the directionality of the affect, we used the absolute value of the difference between each score and the neutral point. Greater numbers indicate greater affective content regardless of the particular polarity.

Semantic diversity We used a measure from Hoffman et al. (2013), which extracts semantic diversity based on the loadings from Latent Semantic Analysis. Words which can be associated with many other words have higher scores than those which can be associated with a smaller set of other words.

Results

We run a series of OLS regression models testing if the opposite effects of concreteness and precision will hold when controlling for affective content and semantic diversity. Summary of the results are presented in Table 2. When latency was used as a measure of cognitive performance, concreteness and precision had the same opposite effects as in Study 1, regardless if we were controlling for affective content or semantic diversity. When accuracy was used as a measure of cognitive performance, adding affective content as a covariate did not change the results. When adding semantic diversity to the model, P-depth was not a reliable predictor, while P-inclusiveness and concreteness continued to be statistically significant predictors.

The results from this study suggest that precision, while being somewhat related to other psycholinguistic measures, is not reducible to them and explains unique variance in cognitive performance. Most importantly, precision and concreteness had opposite effects even when

Table 2 Summary of the regression models used in Study 2

	Mode 1: Latency	Model 2: Latency	Mode 3: Accuracy	Mode 4: Accuracy
Concreteness	−0.08*	−0.16*	0.09*	0.17*
	(0.01)	(0.012)	(0.012)	(0.014)
Precision: inclusiveness	0.08*	0.07*	−0.07*	−0.05*
	(0.01)	(0.012)	(0. 012)	(0.012)
Precision: depth	0.08*	0.06*	−0.05*	−0.02
	(0.01)	(0.012)	(0.012)	(0.013)
Length	0.07*	0.05*	0.18*	0.22*
	(0.01)	(0.012)	(0. 012)	(0.013)
Frequency	−0.58*	−0.50*	0.45*	0.35*
	(0.011)	(0.013)	(0. 013)	(0.015)
Affective content	−0.04*	−0.07*	0.04*	0.07*
	(0.01)	(0.011)	(0.012)	(0.013)
Semantic diversity		−0.17*		0.17*
		(0.013)		(0.015)
Degrees of freedom	6065	5293	6065	5293
R-squared	0.43	0.43	0.22	0.23

The regression coefficients are standardized beta weights, with standard errors of the estimate provided in parenthesis
* $p < .001$

controlling for those other measures. On average, more concrete words and less precise words were processed faster and more accurately. We also found that P-depth was more closely related to semantic diversity than P-inclusiveness, and that when controlling for semantic diversity the only significant precision predictor for accuracy was P-inclusiveness.

Discussion

The central claim in this paper is that abstractness has two separate meanings, which are not only independent from each other, but can have opposite effects on cognitive performance. To distinguish between the two, we used the terms "concreteness" and "precision". While both of these imply lower level of abstraction, the first one is based on the type of the information, while the second one is based on the quantity of information. We operationalized precision as the position of a concept in a hierarchically organized semantic network, and defined two separate measures, one based on distance from the the top of the hierarchy (P-depth), and a second one based on the number of descending nodes (P-inclusiveness). We hypothesized that in sharp contrast to concreteness, higher precision will actually hinder cognitive performance. Study 1 provided strong support for this hypothesis. Concreteness and precision consistently had opposite effects on cognitive performance, regardless if cognitive performance was based on latency or accuracy data. The opposite effects were independent of whether concreteness and precision were entered separately or simultaneously in the model. In Study 2 we further tested the generalizability of these findings by additionally controlling for affective content and semantic diversity. The general pattern was largely the same: precision and concreteness consistently had opposite effects.

The results from our studies have a number of theoretical and methodological implications. On the theoretical level, they clearly demonstrate that we need to distinguish between two measures of abstractness, otherwise we face the paradoxical finding where abstractness both helps and hinders cognitive processing. Since abstractness has been a central concept in many subfields of psychology (Burgoon et al. 2013), we believe that the relevance of our findings goes beyond psycholinguistics. Another theoretical implication concerns the current debate about concept representation (Binder 2016). The last two decades have accumulated evidence in support of an amodal view of concept representation, where concept's storage and retrieval rely on sensory information. While our results are consistent with a sensory component in concept storage and retrieval, the precision effect we observed cannot be explained by a framework which does not include a hierarchical semantic organization.

In addition to its theoretical implications, the current work also has a methodological contribution in that it explicitly quantifies precision for a large number of nouns. Since P-depth and P-inclusiveness captured different aspects of reduction of information, we used both measures when testing the effects of precision on cognitive processing. While the two measures confirmed our hypothesis we also observed important differences. The two measures were only moderately correlated with each other. They were also differentially correlated with other psycholinguistic variables related to abstractness. For example, semantic diversity and concreteness were more strongly related to P-depth than to P-inclusiveness (see appendix). Further, when predicting accuracy, the effect of P-depth seized being significant when we controlled for semantic diversity, while P-inclusiveness remained a significant predictor. While the general pattern of results from Studies 1 and 2 suggests that both measures are important, future research will need to test if they have differential effects on cognitive performance in a broader sets of tasks.

While we found strong empirical support for differentiating between concreteness and precision, a distinction based on type of and quantity of information, our approach has some important limitations. First, we have not addressed possible domain specificity effects, where guiding principles from one domain of knowledge might not be applicable to another domain (see Hirschfeld and Gelman 1994). We have treated the different branches of the semantic hierarchy as qualitatively similar, yet there might be differences between concepts related to folkbiology and concepts related to artefacts, for example. Second, the distinction we have proposed is not applicable to other aspects of abstractness related to global versus local processing (for a review see Kimchi 2014). A triangle, for example, consists of three lines, but it is the combination of these lines into a particular structure rather than the lines themselves that gives the triangle its properties. While such emerging properties are typically framed as being part of hierarchies based on inclusiveness, these hierarchical patterns cannot be seen as reduction of information.

Conclusion

In this paper we contrasted two meanings of abstractness, referring to them as concreteness and precision. While concreteness is typically defined in terms of type of information, we proposed that precision should be defined as the amount of information based on concept's position in a semantic hierarchy. Using large databases, we replicated the classical "concreteness effect", where concrete words were processed faster and more accurately than abstract words. Strikingly, we also found a "precision effect", where more precise words were processed more slowly and less accurately than more abstract words. These results suggest that abstractness should not be treated as an unitary construct since one aspect of abstractness—precision—hinders cognitive performance, while the other—concreteness—facilitates it. The precision effect reported here cannot be explained by current theories linking abstraction to cognitive performance, and it was observed even when controlling for concreteness, affective content, and semantic diversity.

Acknowledgements We are grateful to Doug Medin, Scott Atran, David Balota, Victor Kuperman, Bryor Snefjella, Dermot Lynott, Anastasia Smirnova, and Nadiya Kostyuk for helpful comments and suggestions. This work was supported by Air Force Office of Scientific Research Grant FA9550-10-1-0373.

Author's contribution RI and RA developed the theoretical framework. RI assembled the database and conducted the statistical analyses. RI and RA interpreted the results and prepared the manuscript. Both authors approved the final version of the manuscript for submission.

Appendix

See Table 3.

Table 3 Zero-order correlations of all variables from Studies 1 and 2

	2	3	4	5	6	7	8	9
1 Concreteness	−0.02*	0.36***	0.05***	−0.05***	−0.38***	0.15***	−0.14***	−0.24***
2 P-inclusiveness		0.16***	−0.23***	0.23***	0.07***	−0.23***	−0.04***	−0.1***
3 P-depth			−0.17***	0.16***	0.17***	−0.11***	−0.03**	−0.19***
4 Accuracy				−0.63***	0	0.58***	0.14***	0.32***
5 Latency					0.26***	−0.63***	−0.15***	−0.4***
6 Length						−0.39***	0.05***	−0.13***
7 Frequency							0.15***	0.46***
8 Aff.Content								0.06***
9 Sem.Divers.								

* $p < .05$; ** $p < .01$; *** $p < .001$

References

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods*, *39*, 445–459.

Barsalou, L. W. (2003). Abstraction in perceptual symbol systems. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, *358*, 1177–1187.

Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin & Review*, *23*(4), 1096–1108.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. Sebastopol: O'Reilly.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera 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.

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.

Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There are many ways to see the forest for the trees a tour guide for abstraction. *Perspectives on Psychological Science*, *8*(5), 501–520.

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452–465.

Frank, S. L. (2013). Uncertainty reduction as a measure of cognitive load in sentence comprehension. *Topics in Cognitive Science*, *5*(3), 475–494.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427.

Hajibayova, L. (2013). Basic-level categories: A review. *Journal of Information Science*, *39*(5), 676–687.

Hirschfeld, L. A., & Gelman, S. A. (Eds.). (1994). *Mapping the mind: Domain specificity in cognition and culture*. Cambridge: Cambridge University Press.

Hoffman, P., Ralph, M. A. L., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, *45*(3), 718–730.

Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 385.

Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, *23*(4), 492.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, *126*(3), 248.

Kimchi, R. (2014). The perception of hierarchical structure. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization*. Oxford: Oxford University Press.

Kousta, S. 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.

Kousta, S. T., Vinson, D. P., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition*, *112*(3), 473–481.

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 92.

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, *126*(2), 313–318.

Medin, D. L. (1983). Structural principles in categorization. In T. Tighe & B. Shepp (Eds.), *Development: Interactional analyses* (pp. 203–230). Hillsdale, NJ: Erlbaum.

Meilă, M. (2007). Comparing clusterings—An information based distance. *Journal of Multivariate Analysis*, *98*(5), 873–895.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38*(11), 39–41.

Murphy, G. L. (2004). *The big book of concepts*. Cambridge, MA: MIT Press.

Paivio, A. (1966). Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *20*(4), 378.

Paivio, A. (1967). Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. *Psychological Reports*, *20*(1), 239–245.

Paivio, A. (1968). A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behavior*, *7*(1), 41–49.

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

Paivio, A., & Olver, M. (1964). Denotative-generality, imagery, and meaningfulness in paired-associate learning of nouns. *Psychonomic Science*, *1*(1–12), 183–184.

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

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Resnik, P. (1995). *Disambiguating noun groupings with respect to WordNet senses*. arXiv:cmp-lg/9511006v1.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. London: Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439.

Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European Review of Social Psychology*, *2*(1), 1–30.

Shallice, T., & Cooper, R. P. (2013). Is there a semantic system for abstract words? *Frontiers in Human Neuroscience*, *7*, 1–10.

Shannon, C. E. (1948). A Mathematical theory of communication. *Bell System Technical Journal, 27*, 379–423.

Spreen, O., & Schulz, R. W. (1966). Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning and Verbal Behavior*, *5*(5), 459–468.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*(3), 457–482.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Abingdon-on-Thames: Lawrence Erlbaum.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*, 440–463.

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review, 94*, 3–15.

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*(5), 1256.

Wang, J., Conder, J., Blitzer, D., & Shinkareva, S. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping*, *31*(10), 1459–1468.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207.

Watkins, E., Moberly, N. J., & Moulds, M. L. (2008). Processing mode causally influences emotional reactivity: Distinct effects of abstract versus concrete construal on emotional response. *Emotion*, *8*(3), 364.