



Implementing a concept network model

Sarah H. Solomon¹ · John D. Medaglia² · Sharon L. Thompson-Schill¹

© The Psychonomic Society, Inc. 2019

Abstract

The same concept can mean different things or be instantiated in different forms, depending on context, suggesting a degree of flexibility within the conceptual system. We propose that a feature-based network model can be used to capture and predict this flexibility. We modeled individual concepts (e.g., BANANA, BOTTLE) as graph-theoretical networks, in which properties (e.g., YELLOW, SWEET) were represented as nodes and their associations as edges. In this framework, networks capture within-concept statistics that reflect how properties relate to one another across instances of a concept. We extracted formal measures of these networks that capture different aspects of network structure, and explored whether a concept's network structure relates to its flexibility of use. To do so, we compared network measures to a text-based measure of semantic diversity, as well as to empirical data from a figurative-language task and an alternative-uses task. We found that network-based measures were predictive of the text-based and empirical measures of flexible concept use, highlighting the ability of this approach to formally capture relevant characteristics of conceptual structure. Conceptual flexibility is a fundamental attribute of the cognitive and semantic systems, and in this proof of concept we reveal that variations in concept representation and use can be formally understood in terms of the informational content and topology of concept networks.

Keywords Conceptual knowledge · Conceptual flexibility · Network science

The apple information evoked by “apple pie” is considerably different from that evoked by “apple picking”: The former is soft, warm, and wedge-shaped, whereas the latter is firm, cool, and spherical. If you scour your conceptual space for APPLE information, you will uncover the knowledge that apples can be red, green, yellow, or brown when old; that they can be sweet or tart; that they are crunchy when fresh and soft when baked; that they are naturally round but can be cut into slices; that they are firm, but mushy if blended; that they can be found in bowls, in jars, and on trees. Despite the complexity of this conceptual knowledge, we can generate an appropriate APPLE instance, with the appropriate features, based on the context

we are in at the time. In other words, the APPLE concept can be flexibly adjusted in order to enable a near-infinite number of specific and appropriate APPLE exemplars. This flexibility enables concepts to be represented in varied and fluid ways, a central characteristic of the semantic system.

The concept APPLE can be instantiated as a Granny Smith or as a Macintosh, and either one can easily be brought to mind. The fact that a single conceptual category has many distinct subordinate types that differ from each other is a basic form of conceptual variation that has been embedded within hierarchical semantic models (e.g., Collins & Quillian, 1969). But even a representation of a single instance of APPLE can be flexibly adjusted: activated properties might be RED and ROUND while shopping, whereas they might be SWEET and CRISPY while eating. A concept can also be represented in varied states, each with their own distinct features: the representation of an APPLE is FIRM versus SOFT before and after baking, and SOLID versus LIQUID before and after juicing. Conceptual flexibility is further evidenced in the frequent nonliteral use of concepts: One should stay away from “bad apples” and should not “compare apples with oranges,” and one can use concepts fluidly in novel analogies and metaphors.

Typically, theories assume a “static” view of concepts, in which conceptual information is stable across instances. But this framework makes it hard to model the conceptual shifts

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-019-01217-1>) contains supplementary material, which is available to authorized users.

✉ Sarah H. Solomon
sarahsol@sas.upenn.edu

¹ Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

² Department of Psychology, Drexel University, Philadelphia, PA, USA

over long and short time-scales that occur during context-dependent concept activation and learning (Casasanto & Lupyan, 2015; Yee & Thompson-Schill, 2016). By “context,” we refer to events or situations, whether in the physical environment or in language, that could influence the ways in which conceptual information is activated and represented. The flexibility of meaning is also a challenge in the language domain, and is referred to as *enriched lexical processing*, *type-shifting*, or *coercion* (e.g., McElree, Traxler, Pickering, Seely, & Jackendoff, 2001; Pustejovsky, 1998; Traxler, McElree, Williams, & Pickering, 2005). Though conceptual flexibility is a pervasive phenomenon, it poses a formidable challenge: How is conceptual information organized to enable this flexibility?

We are particularly interested in the structure of individual concepts (e.g., APPLE, SNOW), rather than the structure of semantic space more broadly. This latter pursuit—the modeling of semantic space—has already been approached from various theoretical orientations and methodologies. Some theoretical approaches claim that the meaning of a concept can be decomposed into features and their relationships with each other (e.g., McRae, de Sa, & Seidenberg, 1997; Sloman, Love, & Ahn, 1998; Smith, Shoben, & Rips, 1974; Tversky, 1977; Tyler & Moss, 2001). For example, the “conceptual structure” account (Tyler & Moss, 2001) represents concepts as binary vectors indicating the presence or absence of features, and argues that broad semantic domains (e.g., ANIMALS, TOOLS) differ in their characteristic properties and in their patterns of property correlations (e.g., HAS-WINGS and FLIES tend to co-occur within the ANIMAL domain). Models that represent basic-level concepts in terms of their constituent features are valuable because they can be implemented in computational architectures such as parallel distributed processing models and other connectionist models. For example, the “feature-correlation” account (e.g., McRae, 2004; McRae, Cree, Westmacott, & de Sa, 1999; McRae et al., 1997) pairs empirically derived conceptual feature statistics with a type of connectionist model called an attractor network: Property statistics characterize the structure of semantic space, and the model can leverage these statistics to settle on an appropriate conceptual representation given the current inputs (Cree, McNorgan, & McRae, 2006; Cree, McRae, & McNorgan, 1999).

However, most instantiations of feature-based models represent individual concepts with sets of features that are static and unchanging—a clear limitation if one aims to incorporate flexibility into conceptual structure. Some recent connectionist models have aimed to incorporate context-dependent meaning (Hoffman, McClelland, & Lambon Ralph, 2018), and the flexibility and context dependence of individual features was addressed in prior work (Barsalou, 1982; Sloman et al., 1998). For example, Sloman et al. modeled the pairwise dependencies between features in order to ascertain the mutability or immutability of features. The authors claimed that a

feature is immutable if it central to a concept’s structure: It is harder to imagine a concept missing an immutable feature (e.g., a robin without bones), than a mutable feature (e.g., a jacket without buttons). The authors argued that modeling a concept in terms of formal pairwise relationships makes it possible for concepts to be structured as well as flexible. Whereas the goal of these and other researchers has been to characterize the role of individual conceptual *features* (Cree et al., 2006; Devlin, Gonnerman, Andersen, & Seidenberg, 1998; Sedivy, 2003; Sloman et al., 1998; Tyler & Moss, 2001), our present goal was to examine whether a feature-based conceptual structure can shed light on the flexibility of a concept as a whole.

Another way to model conceptual knowledge is to use a network to capture the relationships *between* concepts in language. The use of networks to model semantic knowledge has a well-established history. The early “semantic network” models (Collins & Loftus, 1975; Collins & Quillian, 1969) represent concepts as nodes in a network; links between these nodes signify associations between concepts in semantic memory. These networks capture the extent to which concepts are related to other concepts and features, and can model the putatively hierarchical nature of conceptual knowledge. Though these models are “network-based,” they are so in a rather informal way. On the other hand, network science, a mathematical descendent of graph theory, has developed a rich set of tools to study networks in a formal, quantitative framework (see Barabási, 2016). For example, current network approaches characterize relationships between concepts in terms of their word-association strengths or corpus-based co-occurrence statistics. Word co-occurrence statistics can be extracted from text corpora and have been used to create probabilistic models of word meanings (Griffiths, Steyvers, & Tenenbaum, 2007), to represent semantic similarity (Landauer & Dumais, 1997), and to characterize the structure of the entire lexicon (e.g., WordNet; Miller & Fellbaum, 2007). In a similar approach, empirically-derived word association data have been used to capture and analyze the structure of semantic space (De Deyne, Navarro, Perfors, & Storms, 2016; Steyvers & Tenenbaum, 2005; Van Rensbergen, Storms, & De Deyne, 2015).

In network science, units and links are referred to as “nodes” and “edges,” respectively, and the pattern of connections between nodes can be precisely described, revealing patterns of network organization. Nodes can represent any number of things (e.g., cities, people, neurons), depending on the system being modeled; edges can likewise represent a range of connection types (e.g., roads, friendship, synapses). Many diverse systems have been described in network science terms, including the World Wide Web (e.g., Cunha et al., 1995), social communities (Wellman, 1926), the nervous system of *Caenorhabditis elegans* (Watts & Strogatz, 1998), and many others (see Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006). Here we will summarize how network science has been applied, and can be further extended, to the study of conceptual knowledge.

Network structure can be characterized at different levels of organization. For example, the large-scale organization of a network (i.e., topology) can be characterized as a “regular,” “random,” or “small-world” structure (Figs. 1A–C; Watts & Strogatz, 1998). In *regular* networks, each node is connected to its k nearest neighbors; in *random* networks, nodes are randomly connected to each other. Regular networks result in long path lengths and high local clustering (modular processing), whereas random networks result in short path lengths and minimal clustering (integrated processing). Between these two extremes is the *small-world* network, which contains high-clustering as well as a few random, long-range connections; this results in the “small-world” phenomenon, in which each node is connected to all other nodes with relatively few degrees of separation. A small-world topology thus maximizes the efficient spread of information, enables both modular and integrated processing, and supports network complexity (Bassett & Bullmore, 2006; Watts & Strogatz, 1998). Much work has revealed that naturally evolving networks have small-world topology (Bassett & Bullmore, 2006), including functional brain networks (Bassett et al., 2011; Salvador, Suckling, Schwarzbauer, & Bullmore, 2005) and language networks (i Cancho & Solé, 2001; Steyvers & Tenenbaum, 2005). These systems exhibit small-world topologies presumably because this structure facilitates local “modular” processing as well as easy communication via a few long-range connections.

The semantic network approaches described above use nodes to represent individual words, and edges to represent their co-occurrence statistics or association strengths. Once modeled in this way, network structure can be quantitatively analyzed and related to other phenomena. As we mentioned above, it has been suggested that human language networks exhibit small-world properties (i Cancho & Solé, 2001; Steyvers & Tenenbaum, 2005). Additionally, semantic networks appear to exhibit an “assortative” structure, meaning that semantic nodes tend to have connections to other semantic nodes with similar characteristics (e.g., valence, arousal, and concreteness; Van Rensbergen et al., 2015). A spreading-activation model applied to these word-association networks makes accurate predictions of weak similarity judgments—for example, between the unrelated concepts of “teacher” and “cup” (De Deyne et al., 2016). Furthermore, Steyvers and Tenenbaum (2005) reported that the degree of a word in a language network (i.e., how many links it has to other word nodes) predicts aspects of language development and processing: A high-degree word, for instance, is likely to be learned at a younger age and engenders faster reaction times on a lexical decision task. These network models are valuable because semantic structure can be analyzed using a rich set of network science tools. However, current network-based implementations do not provide the internal conceptual structure that is necessary—we argue—to model conceptual flexibility. In other words, it is hard to provide a model of conceptual flexibility

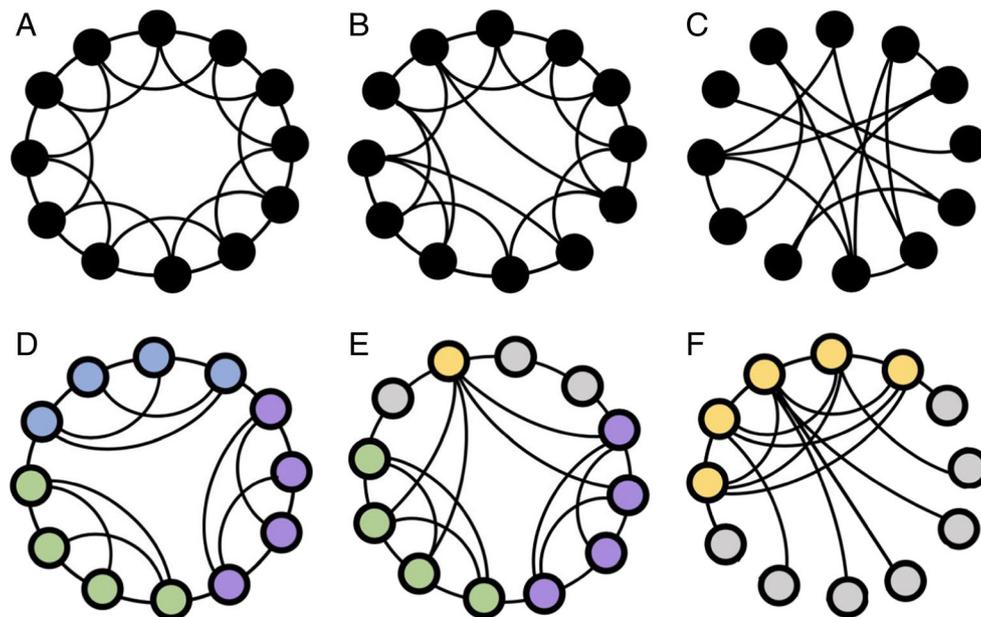


Fig. 1 A visualization of network topologies and measures. Networks are defined in terms of nodes (circles) and edges (lines). Network topologies fall into three main categories: (A) regular, (B) small-world, and (C) random (Watts & Strogatz, 1998). Most naturally evolving networks exhibit small-world topology, including neural networks and language networks. Regular and small-world networks have high clustering. (D) Modularity reflects the extent to which a network can be partitioned into a set of densely connected “modules,” represented here in distinct colors.

(E) Some nodes participate in multiple modules, reflecting a diversity of connections; this is captured in a “diversity coefficient.” A diverse node (yellow) participates in multiple modules (green, purple), whereas other nodes (gray) do not exhibit these diverse connections. (F) A network has strong core–periphery structure if it can be characterized in terms of a single densely connected “core” (yellow) and a sparsely connected “periphery” (gray)

(in the sense described above) when the features that are being flexibly adjusted are not explicitly represented.

We believe that a feature-based conceptual framework paired with network science techniques provides a platform through which conceptual flexibility can be quantified and explored. Here we introduce a new approach in which concepts are represented as their own feature-based networks, and we work through an example as a proof of concept. We created concept-specific networks for 15 concepts (e.g., CHOCOLATE, GRASS, KNIFE), in which the nodes represented conceptual features (e.g., BROWN, GREEN, METAL, SHARP, SWEET) and edges represented how those features co-occurred with each other within each concept. The creation of such networks thus required the calculation of within-concept feature statistics, which may describe how a concept's information may be appropriately adjusted to form valid, yet varied, concept representations. Though here we were interested in analyzing the structure of basic-level concepts, these concept network methods could theoretically be applied at any level of the conceptual hierarchy. Our specific goals were to (1) show that creation of such networks is possible, (2) confirm that these networks contain concept-specific information, and (3) demonstrate that these networks permit the extraction of measures that relate to conceptual flexibility.

We hand-picked a selection of measures to extract and analyze from our concept networks. As we mentioned above, small-world networks (Fig. 1B) are characterized by high network clustering, such that a node's neighbors also tend to be neighbors with each other. Previous work has shown a relationship between the clustering within semantic networks and individual differences in creativity (Kenett, Anaki, & Faust, 2014); because creativity relates to flexible conceptual processing, the *clustering coefficient* was one of our measures of interest. In small-world networks, this clustering paired with random connections results in network modules, which are communities of nodes with dense connections between them. *Modularity* is a formal measure that captures the extent to which a given network can be partitioned in this way (Fig. 1D). A network with a modular structure is able to activate distinct, specialized sets of nodes; because this might translate into a concept's ability to activate distinct sets of features, modularity was another measure of interest. In modular networks, each node can also be characterized in terms of its *diversity* of connections across network modules (Fig. 1E). Some nodes may have links within only one module, whereas others may have links that are highly distributed across different network modules. Because related measures are often used to define network hubs that support flexible network processing (Sporns, 2014; van den Heuvel & Sporns, 2013), we were interested in exploring the relation between network diversity and flexible concept use. Another kind of network topology is *core-periphery* structure (Fig. 1F), in which a network is characterized by one highly connected core and a sparsely

connected periphery. Core-periphery organization, originally observed in social networks (Borgatti & Everett, 2000), has been applied to functional networks in neuroimaging data (Bassett et al., 2013). A core-periphery structure in a concept network would reflect one set of highly associated features (i.e., the core), but also a substantial collection of features that are weakly associated with one another (i.e., the periphery). We included core-periphery structure as a measure of interest because we hypothesized that the “stiff” core and/or “flexible” periphery of a concept network (Bassett et al., 2013) might relate to flexible concept use.

In this proof of concept, we extracted measures of network organization from concept-specific networks (i.e., clustering, modularity, core-periphery, diversity) and explored what these structural characteristics might predict about how a concept is used. Conceptual flexibility manifests when a concept is recruited to represent varied subordinate exemplars, when a concept word is used in a variety of language contexts, and when a concept is differentially activated depending on the task context. We therefore aimed to relate our network measures of interest to a measure of semantic diversity (SemD; Hoffman, Lambon Ralph, & Rogers, 2013), calculated from text-based statistics. We also collected data on two tasks related to conceptual flexibility—a figurative-language task (comprehension of novel similes) and a widely used measure of creative cognition, the alternative-uses task (AUT)—to explore whether network structure relates to how a concept is flexibly used in different language and task contexts. Here we present one variation of the concept network approach, implementing a particular set of methodological decisions on a particular set of concepts, to show the potential of this framework to provide new ways to characterize the structure and flexibility of conceptual knowledge.

Method

Network methods

Introduction to methods

Our goal was to construct feature-based networks that would capture each concept's specific constellation of features and the ways those features relate to each other. There are, however, many ways one could create such networks. Here we walk through one possible instantiation of this method, to reveal the feasibility of this approach and suggest the kinds of analyses that could be used to examine the relationship between concept network topology and conceptual flexibility. We hope that future researchers interested in conceptual knowledge will be able to use, improve, and expand upon these methods. Our data and code are available online: <https://osf.io/vsa2t/>.

We collected data in two rounds, and we refer to these data as *Set 1* and *Set 2*. We collected data for five concepts in Set 1 as a first attempt to construct concept networks. Once we established the success of these methods, we collected data for another ten concepts in Set 2. The concepts in Set 1 were CHOCOLATE, BANANA, BOTTLE, TABLE, and PAPER, and the concepts in Set 2 were KEY, PUMPKIN, GRASS, COOKIE, PICKLE, KNIFE, PILLOW, WOOD, PHONE, and CAR. When statistics are reported separately for the two sets, we report Set 1 followed by Set 2. Once the networks are constructed and we analyze network measures and their relation to other conceptual measures, the sets are no longer treated separately, and each concept is treated as an item ($N = 15$). We use Spearman's rank correlation in all correlational analyses, due to this small sample size. All participants were compensated according to current standard rates, and consent was obtained for all participants in accordance with the University of Pennsylvania IRB.

Network construction

The first step was to define our nodes. Since our nodes represent individual conceptual properties, we compiled a list of properties that could be applied to all of our concepts within each set. Participants ($N = 66$, $N = 60$) were recruited from Amazon Mechanical Turk (AMT) and were asked to list all of the properties that must be true or can be true for each concept. It was emphasized that the properties do not have to be true of all types of the concept. Participants were required to report at least ten properties per concept, but there was no limit on the number of responses they could provide. Once these data were collected, we organized the data as follows. For each concept, we collapsed across different forms of the same property (e.g., “sugar,” “sugary,” “tastes sugary”), and removed responses that were too general (e.g., “taste,” “color”). This was a highly data-driven approach; however, see the Bootstrap Analysis section for an analysis of robustness across properties. For each concept, we only included properties that were given by more than one participant. We then combined properties across all concepts to create our final list of N properties ($N = 129$, $N = 276$) that were represented as nodes in our concept networks.

The same AMT participants who provided conceptual properties also provided subordinate concepts (from now on referred to as “subordinates”) for each of the concepts. For each concept, participants were asked to think about that object and all the different kinds, forms, types, or states in which that object can be found. Participants were required to make at least five responses and could make up to 15 responses. For each concept, we removed subordinates that corresponded to a property for that concept (e.g., “sweet chocolate”), subordinates that were highly similar to other subordinates (e.g., “white chocolate chip cookie,” “chocolate chip cookie”), and responses that were too specific, including some brand names (e.g., “Chiquita banana”). Though this was a data-driven approach, there was some degree

of subjectivity in the final subordinate lists; see the Bootstrap Analysis section for an analysis of robustness across subordinates. In Set 1, we only included responses that were given by more than one participant; due to the increased number of participants and responses in Set 2, we included responses that were given by more than two participants. In both sets, we ended up with a set of K subordinates for each concept (K : $M = 17$, $SD = 3.14$). The included and excluded subordinates for all concepts are presented in Supplementary Table 1.

A separate set of AMT participants ($N = 198$, $N = 108$) were presented with one subordinate of each of the concepts in random order (e.g., “dark chocolate,” “frozen banana”) and asked to select the properties that are true of the specific subordinates (see Fig. 2A). The full list of N properties was displayed in a multiple-choice format. For each subordinate, responses were combined across participants; we thus know, for each subordinate, how many participants reported each of the N properties. In the networks we report here, we used weighted subordinate vectors in which values indicate the percentage of participants who reported each property. However, to reduce noise, a subordinate property was only assigned a > 0 weight if it was reported by more than one participant; if only one participant reported a particular property for a particular subordinate, the weight = 0. For each concept, we excluded properties that were not present in any of its subordinates, resulting in a smaller set of N_c properties that were present in ≥ 1 subordinates (N_c : $M = 126$, $SD = 32.2$). Each concept's data thus included a set of K subordinates, each of which corresponded to an N_c -length vector that indicates each property's weight in that subordinate.

Importantly, these data for each concept can also be considered a set of N_c properties, each corresponding to a vector indicating that property's weight in each of the subordinates. For example, if a concept was described by ten subordinates (K) and 100 meaningful properties (N_c), we have 100 ten-element vectors, each of which represents the contribution of a single property across the subordinates of the concept. The premise behind this concept-network construction was that the ways in which these patterns of property contributions relate to each other, within a single concept, might be an important aspect of conceptual structure. Our networks thus captured the pairwise similarities between properties—that is, between the N_c K -element vectors. To do this, many different distance metrics could be used (e.g., Euclidean, Mahalanobis, or cosine); we used the *pdist()* function in MATLAB, which includes many distance measure alternatives. In the analysis and results we report here, we constructed our networks on the basis of Mahalanobis distance, a measure suited for high-dimensional data and that takes the variance between subordinates and the correlations between subordinates into account. However, there are many other options, the choice of which might depend on other analysis decisions. For example, if the subordinate concepts are binary instead of weighted, a “matching” measure such as the Jaccard distance might be more appropriate.

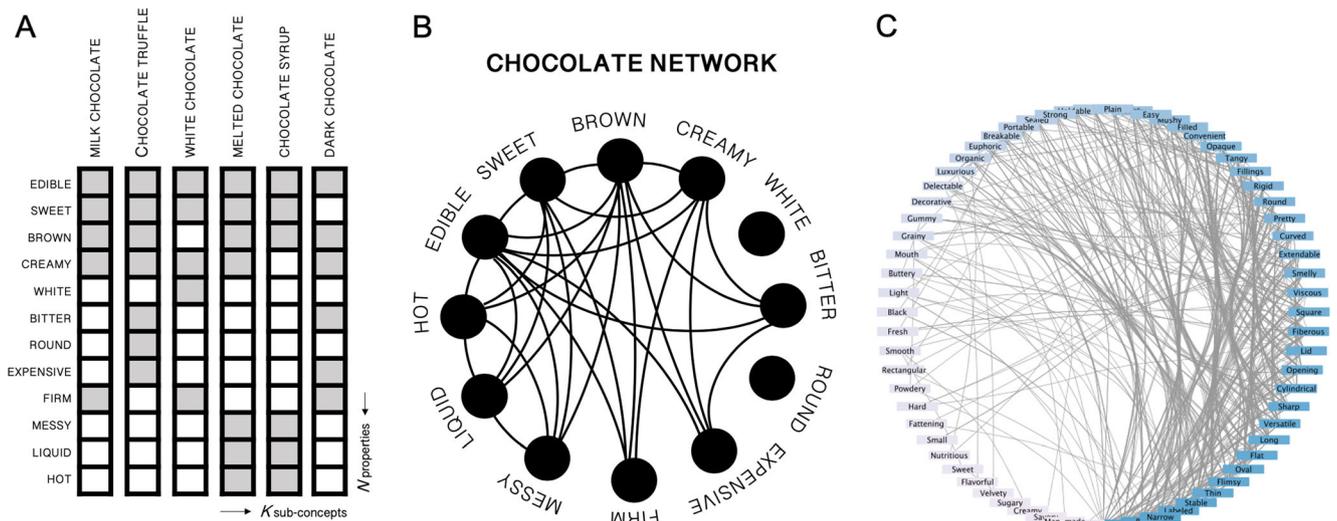


Fig. 2 Visualizing the CHOCOLATE network. (A) The CHOCOLATE concept can be broken down into a range of subordinates, which can each be defined as a property vector (columns). Each property can also be defined as a vector (rows), which can be used to calculate within-concept property relationships. Only a small set of potential subordinates and properties are shown here, for simplicity. (B) A simple schematic of the CHOCOLATE network that reveals a selection of potential property relationships. Certain properties might cluster together in the network—for example,

First, the distance between each of the N_c K -element vectors was calculated. This resulted in a square, symmetrical $N_c \times N_c$ matrix that contains the distance between each pair of properties. These values were scaled between 0 and 1 and converted to a similarity measure by subtracting these values from 1. We thus created a network for a single concept that captures pairwise property–property similarities; this represents the patterns of property relationships across subordinates within a given concept (see Fig. 2). Here we used weighted networks, where edges represent similarity measures between 0 and 1, though it would also be possible to use unweighted networks by binarizing these similarity values according to a given threshold. We repeated this (weighted) network construction process for each of the 15 concepts. These final networks were then analyzed using standard network science methods (see the Network Analysis section).

A simple measure of concept stability The subordinate–property data for each concept enabled us to calculate a simple measure of conceptual stability that did not involve treating concepts as networks. For each concept, we counted the number of properties that were represented across *all* of that concept’s subordinates (i.e., weighted value greater than zero). We then divided this number by N_c in order to calculate the proportion of properties that were universally consistent for that concept. We interpreted this measure as a measure of conceptual stability, because higher values indicate that a large number of properties are not variable across conceptual instances. We refer to this measure as a concept’s *simple stability* and consider it to be an inverse measure of conceptual flexibility.

EDIBLE, SWEET, BROWN, CREAMY and MESSY, LIQUID, HOT. (C) The actual CHOCOLATE network we constructed on the basis of the empirical property statistics. Our constructed CHOCOLATE network was binarized (threshold = 90%) to reduce the number of properties in order to ease visualization. Properties are arranged in order of degree (number of links), from low degree (white) to high degree (blue). The image was generated using cytoscape (Shannon et al., 2003)

Classification analysis

Our main goal was to extract concept-specific measures from our networks, and this goal would only be justified if the network *structures* themselves were concept-specific. Even though different sets of data contributed to the different concept networks, it was not necessarily the case that the resulting networks would differ from each other. It could theoretically be that property relationships are consistent across the entire semantic domain; indeed, this has been the premise of the neural-network models of semantic knowledge created thus far (e.g., Cree et al., 2006; Cree et al., 1999; McClelland & Rogers, 2003). However, our goal was to capture concept-specific property relationships, so our first task was to test whether we succeeded in this goal.

If our concept network models captured concept-specific information, the networks should be able to successfully discriminate between new concept exemplars. Exemplar data were generated from sets of photographs for each concept (see Fig. 3); all subordinates were represented. AMT participants ($N = 60$, $N = 30$) were shown one image per concept and asked to imagine interacting with this object in the real world and to consider what properties it has. The full list of N properties was displayed in multiple-choice format, and participants were asked to select the properties that they believed applied to the object in the image. Individual participants’ responses to each exemplar image were represented as N -length property vectors and were used as test data in the classification analysis. The test data comprised 300 property



Fig. 3 Example images used to generate test data in the classification analysis. The test data used in the classification analysis were generated from participants who made property judgments on images of conceptual

exemplars. A yellow cross indicates the object to be considered. The example images are for grass (top) and cookie (bottom)

vectors (Set 1, 60/concept; Set 2, 30/concept); classification analyses were run separately for Sets 1 and 2.

By performing eigen-decomposition on each adjacency matrix (i.e., concept network) we could assess the extent to which a vector would be expected given an underlying network structure (e.g., Huang et al., 2018; Medaglia et al., 2018). For each adjacency matrix A , V was the set of N_c eigenvectors, ordered by eigenvalue. M was the number of ordered eigenvectors to include in the analysis and designated a subset of V . For each eigenvector v , we found the dot product with signal vector x , which gave us the projection of x on that dimension in the eigenspace of A . That is, it gave us an “alignment” value for that particular signal and that particular eigenvector. We could include all eigenvectors in M by taking the sum of squares of the dot products for each eigenvector. The alignment value for each signal was defined as

$$\tilde{x} = \sum_{i=1}^M (v_i \cdot x)^2 \quad (1)$$

where x is a property vector, M is the number of eigenvectors to include in alignment (sorted by eigenvalue), v_i is one of the M eigenvectors of the adjacency matrix, and \tilde{x} is the scalar alignment value for signal x with adjacency matrix A , given the eigenvectors 1– M . In our case, signal x was a property vector corresponding to a particular exemplar image (e.g., Fig. 3), which we aligned with each of the concept networks. Each exemplar was restricted to the properties included in each concept model before transformation; that is, the exemplar data (x) were reduced to N_c -length vectors. The concept network that resulted in the highest alignment value (\tilde{x}) was taken as the “guess” of the classifier; each exemplar was classified either correctly (1) or incorrectly (0). We averaged these data across all exemplars to calculate the average classifier accuracy.

To calculate a baseline measure of classification accuracy, we created traditional vector models for each concept. These models were similar to those used elsewhere in the literature (McRae, 2004; McRae et al., 1999; McRae et al., 1997; Tyler & Moss, 2001). For each concept, we averaged the K subordinate vectors resulting in an N_c -length vector containing mean property strength values. Each concept’s traditional vector model and network model contained the same conceptual properties. We ran a separate classification analysis using these traditional models and a correlational classifier. Each exemplar property vector was correlated with each of the traditional concept vector models; the concept model that resulted in the highest correlation value was taken as the guess of the classifier. We calculated average measures of classifier performance using the same methods described above.

Network analysis

We extracted network metrics from our concept networks using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010). In the discussion below, the set of nodes in each network is designated as N , and n is the number of nodes. The set of links is L , and l is the number of links. The existence of a link between nodes (i, j) is captured in a_{ij} : $a_{ij} = 1$ if a link is present and $a_{ij} = 0$ if a link is absent. The weight of a link is represented as w_{ij} , and is normalized such that $0 \leq w_{ij} \leq 1$. l^w is the sum of all weights in the network. The network metrics we extracted included clustering coefficients, modularity (Q), core–periphery structure, and diversity coefficients (Fig. 1), for the reasons described above.

The clustering coefficient captures the “cliqishness” of a network—that is, the extent to which a node’s neighbors are also neighbors of each other. The clustering coefficient is calculated for each node individually (C_i), by calculating the percentage of

potential pairwise connections among the neighbors of node i . A “triangle” is formed when node i is linked to j and h while j and h are also linked to each other; the number of existing triangles can be calculated for each node (t_i ; Eq. 2), which is used to calculate the proportion of possible triangles that exist for each node. This proportion is averaged across nodes to result in the clustering coefficient (C ; Eq. 3) for a network (Rubinov & Sporns, 2010; this can also be calculated for weighted networks):

$$t_i = \sum_{h,j \in N} a_{ij} a_{ih} a_{jh} \quad (2)$$

$$C = \frac{1}{n} \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i-1)} \quad (3)$$

Modularity (Q) is a metric that describes a network’s community structure. We can attempt to partition a weighted network into sets of nonoverlapping nodes (i.e., modules) such that within-module connections are maximized and between-module connections are minimized. Some networks exhibit more of a modular structure than others; Q^w is a quantitative measure of modularity for each weighted network (Eq. 4; Rubinov & Sporns, 2010), which is defined as

$$Q^w = \frac{1}{l^w} \sum_{i,j \in N} \left[w_{ij} - \frac{k_i^w k_j^w}{l^w} \right] \delta_{m_i, m_j} \quad (4)$$

where m_i is the module containing node i , and $\delta_{m_i, m_j} = 1$ if $m_i = m_j$, and 0 otherwise. The modularity calculation is stochastic; in our analysis we performed a modularity partition 10,000 times and averaged across these iterations to calculate a mean Q coefficient for each concept. Nodes may have connections to many different modules, or have very few such connections. The diversity coefficient (h_i^\pm) is a measure ascribed to individual nodes that reflects the diversity of connections that each node has to modules in the network. This is a version of the participation coefficient, and is calculated using normalized Shannon entropy; we have previously used entropy to model property flexibility, and so predicted that diversity would be a good candidate for a network-based measure of conceptual flexibility. The diversity coefficient (Eq. 5; Rubinov & Sporns, 2011) for each node is defined as

$$h_i^\pm = -\frac{1}{\log m} \sum_{u \in M} p_i^\pm(u) \log p_i^\pm(u) \quad (5)$$

where $p_i^\pm(u) = \frac{s_i^\pm(u)}{s_i^\pm}$, $s_i^\pm(u)$ is the strength of node i within module u , and m is the number of modules in modularity partition M . We averaged diversity coefficients across nodes in a network to obtain a mean measure of diversity for each concept network. The diversity coefficient is based on Q , which is stochastic; we thus calculated a diversity coefficient for each of the 10,000 modularity partitions, and averaged across these iterations for each concept.

Core–periphery structure is another way to describe the structure of a network. Here, we attempted to partition a network into two nonoverlapping sets of nodes such that connections within one set were maximized (i.e., the “core”) and connections in the other were minimized (i.e., the “periphery”). Core–periphery fit (Q_C) is a quantitative measure of how well each network can be partitioned in this way (Eq. 6), and for weighted networks is defined as

$$Q_C = \frac{1}{v_c} \left[\sum_{i,j \in C_c} (w_{ij} - \gamma_C \bar{w}) - \sum_{i,j \in C_p} (w_{ij} - \gamma_C \bar{w}) \right] \quad (6)$$

where C_c is the set of all nodes in the core, C_p is the set of nodes in the periphery, \bar{w} is the average edge weight, γ_C is a parameter controlling the size of the core, and v_c is a normalization constant (Rubinov, Ypma, Watson, & Bullmore, 2015).

Bootstrap analysis

The properties and subordinates used to create these networks were chosen by study participants, not the experimenters. However, there was a certain degree of subjectivity in how the final lists were constructed, and the properties and subordinates reported by the participants are unlikely to fully represent the total possible sets. We thus ran bootstrap analyses to explore whether relationships between network and non-network measures were dependent on the particular sets of properties and subordinates used in our network construction.

Bootstrapping over subordinates The goal of this analysis was to generate a distribution of correlation values for a specific pair of measures. In each iteration of the analysis, new networks were constructed; for each concept, a random subordinate was removed before network construction. Network measures were extracted from these 15 concept networks and correlated with another measure of choice, and this correlation value was recorded. We performed 1,000 iterations of this analysis, resulting in a distribution of 1,000 correlation values along with a 95% confidence interval.

Bootstrapping over properties The goal of this analysis was to generate a distribution of correlation values for a specific pair of measures. In each iteration of the analysis, new networks were constructed; for each concept, a random 10% of its meaningful properties were removed (N_C) before network construction. Network measures were extracted from these 15 concept networks, and correlated with another measure of choice: this correlation value was recorded. We performed 1,000 iterations of this analysis, resulting in a distribution of 1,000 correlation values along with a 95% confidence interval.

Figurative-language task (similes)

The structure and flexibility of a concept likely has implications for how the concept can be used in creative contexts, such as in figurative-language comprehension (e.g., Sloman et al., 1998). We therefore set out to collect data reflecting the extent to which a given concept is easily interpreted in a figurative context, and to explore the characteristics of conceptual structure that may facilitate this creative process.

Participants

A total of 300 AMT participants contributed data to this study and were compensated according to current standard rates. Consent was obtained for all participants in accordance with the University of Pennsylvania Institutional Review Board (IRB).

Stimuli

Experimental similes of the form “X is like a Y” were constructed using the 15 target concepts in the “vehicle” (Y) position. Fifteen additional concepts were used in the “tenor” (X) position of the similes: TRUTH, TIME, CONVERSATION, SADNESS, CITY, LIFE, DREAM, CAREER, FAMILY, FRIENDSHIP, GOVERNMENT, SCHOOL, HAPPINESS, CELEBRATION, BOREDOM. These tenor concepts were chosen to minimize sensorimotor content, but otherwise were chosen randomly. The 15 target concepts and 15 tenor concepts were fully crossed, resulting in 225 novel similes (e.g., “Truth is like a key,” “Happiness is like chocolate,” “Boredom is like a bottle”). The experimental similes were split into 15 lists consisting of 15 similes each; each target concept and each tenor concept occurred once within each list. An additional ten similes were taken from Blasko and Connine (1993) and were used as control similes; five were “high-apt” similes (e.g., “A book is like a treasure chest”) and five were “moderate-apt” similes (e.g., “Stars are like signposts”). Each list (and therefore each of the experimental similes) was seen by 20 participants. The control similes were seen by all 300 participants.

Task

Each participant read 25 total similes (10 control similes, 15 experimental similes) presented in a randomized order. On each trial, the simile was presented at the top of the screen, with two sliding-scale questions beneath. To assess simile meaningfulness, participants were asked: “How meaningful is this figurative sentence?” ranging from “It does not have any meaning at all” to “It has a very strong meaning.” To assess simile familiarity, participants were asked: “How familiar is this figurative sentence?” ranging from “It is not familiar at all” to “It is very familiar.” These questions were motivated

by prior work on simile comprehension (Blasko & Connine, 1993). For both questions, participants were asked to slide a bar in order to make their desired response. Values on both questions ranged from 0–100, but these values were not displayed to participants.

Analysis

Meaningfulness and familiarity ratings for each simile were separately averaged across participants (experimental: $N = 20$, control: $N = 300$). Inspection of the control similes borrowed from Blasko and Connine (1993) revealed that our measures were sensitive to simile characteristics reported elsewhere: The five high-apt similes were judged as being more meaningful than the five moderate-apt similes [$t(8) = 7.20, p < .0001$]; they were also judged as being more familiar [$t(8) = 5.60, p < .001$].

The meaningfulness and familiarity ratings across all experimental and control similes are shown in Fig. 4. Here it is clear that the meaningfulness and familiarity ratings of our experimental similes are not categorically different from the similes reported in the literature, suggesting that our constructed similes—though novel and pseudo-randomly generated—were meaningful enough for further analysis. The control similes were not included in any subsequent analysis.

Across the 225 experimental similes, the meaningfulness and familiarity ratings were highly correlated [$r(223) = .83, p < .0001$]; since this tight correspondence made it difficult to tease apart the separate measures, we averaged the meaningfulness and familiarity measures within each simile to construct a composite measure we refer to as simile “goodness.”

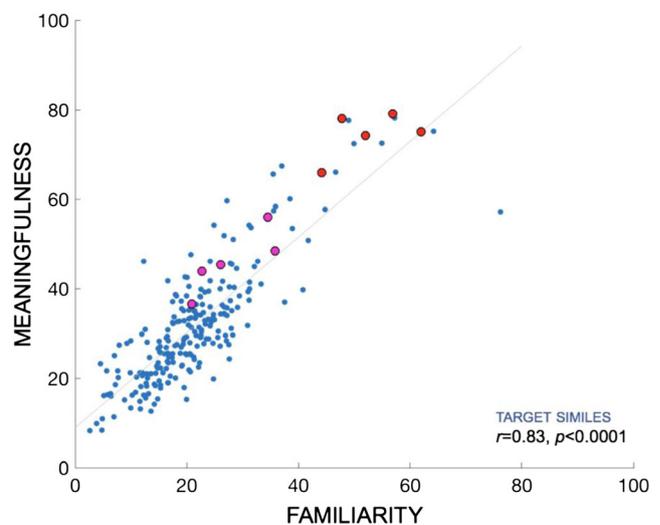


Fig. 4 Simile ratings. The relationship between simile meaningfulness and familiarity ratings for the target similes (blue) and control similes. As expected, the high-apt control similes (red) were rated as more meaningful than the moderate-apt control similes (pink). The range of meaningfulness and familiarity ratings for our target similes subsumes the range of the control similes. Since meaningfulness and familiarity were highly correlated, we averaged these measures to create a single measure of simile goodness

These simile goodness measures were then averaged with respect to each target concept; that is, the goodness ratings for the 15 similes that contained the same target concept (e.g., “chocolate”) were averaged together (e.g., “Truth is like chocolate,” “Happiness is like chocolate”). This resulted in a single simile goodness measure for each of our 15 target concepts.

Alternative-uses task

The AUT is a widely used measure of creative cognition in which participants generate novel uses for common objects. To further explore the relationships between conceptual structure and flexible concept use, we set out to collect data reflecting the extent to which a given concept can be reimagined in creative ways.

Participants

A total of 28 AMT participants generated novel uses in the AUT, and an additional 25 AMT participants provided ratings on these responses. All participants were compensated according to current standard rates. Consent was obtained for all participants in accordance with the University of Pennsylvania IRB.

Alternative-uses task

Participants ($N = 28$) generated alternative uses for the 15 concepts. They were instructed to think of as many novel uses of each object as they could, that their responses should be plausible but significantly different from the common use of the object, and that there were no right or wrong answers. On each trial, the concept label appeared above blank response boxes. Participants had 60 s to answer with as many alternative uses as they could. After 60 s had passed, the next trial immediately began. The presentation order of the 15 concepts was randomized.

After these data were collected, we removed responses that were not task-relevant (e.g., “I can’t think of anything”) and terminal responses that were incomplete (due to the strict 60-s time limit). The number of responses given by each participant for each concept was recorded, and these values were averaged across participants: this resulted in a measure that reflected the mean number of alternative uses generated for each of the 15 concepts ($M = 3.24$, $SD = 0.39$).

From the full set of responses, we selected the first response from each participant for each of the 15 concepts: this resulted in a set of 420 alternative uses (28 per concept). These responses were edited such that they began with a verb (e.g., “Use as a hat,” “Use as a bowling ball,” “Make pie”). This set of responses was rated by an additional set of participants in the next stage of this study.

Alternative-use ratings

An additional set of participants ($N = 25$) provided ratings on the alternative-use data described above. Participants were told that they would be judging other participants’ responses on an AUT, which was used to study creative thinking. On each trial, they were presented with a concept label (e.g., PUMPKIN) and one alternative-use response beneath (e.g., “Use as a hat”). Participants were asked to make three multiple-choice responses on each trial. The first two questions were: “Would this work?” (yes/maybe/no) and “Have you seen someone use this object to do this before?” (yes/no). These questions were not central to our aims of this study, and the corresponding data will not be reported here. The third question, which provided our main measure of creativity, asked participants to rate each response as one of the following: (1) *Very obvious/ordinary use*, (2) *Somewhat obvious use*, (3) *Nonobvious use*, (4) *Somewhat imaginative use*, and (5) *Very imaginative/recontextualized use*. The design of this question was motivated by Hass et al. (2018), in which good reliability was obtained for ratings on a similar AUT. Each participant rated 80 total responses, which were approximately evenly distributed across the 15 concepts; trials were presented in a randomized order. Each alternative-use response was rated by five different participants.

For the main creativity measure, ratings (1–5) for each alternative-use response were averaged across the five participants. The mean ratings of creativity for each alternative use were then averaged within each concept, resulting in a measure that reflects the mean creativity score of alternative-use responses for each of the 15 concepts ($M = 2.91$, $SD = 0.45$).

Results

Classification results

To determine whether our concept networks contained concept-specific information, we ran a classification analysis using eigen-decomposition for both Set 1 and Set 2. We ran multiple analyses using different ranges of eigenvectors, which were sorted by eigenvalue (positive to negative). We started by only using the first eigenvector in each of the concept networks and determined whether this dimension alone could be used to classify the property vector as one of the five concepts in Set 1 or ten concepts in Set 2. One dimension was enough to classify exemplars in both Set 1 (mean accuracy = .31, $SE = .03$, chance = .20) and Set 2 (mean accuracy = .53, $SE = .03$, chance = .10). Classification accuracy in both Set 1 and Set 2 continued to increase as more eigenvectors were included in the analysis (Fig. 5), with performance leveling off around 22–25 eigenvectors. The network-based classification accuracy reached the performance of a more traditional

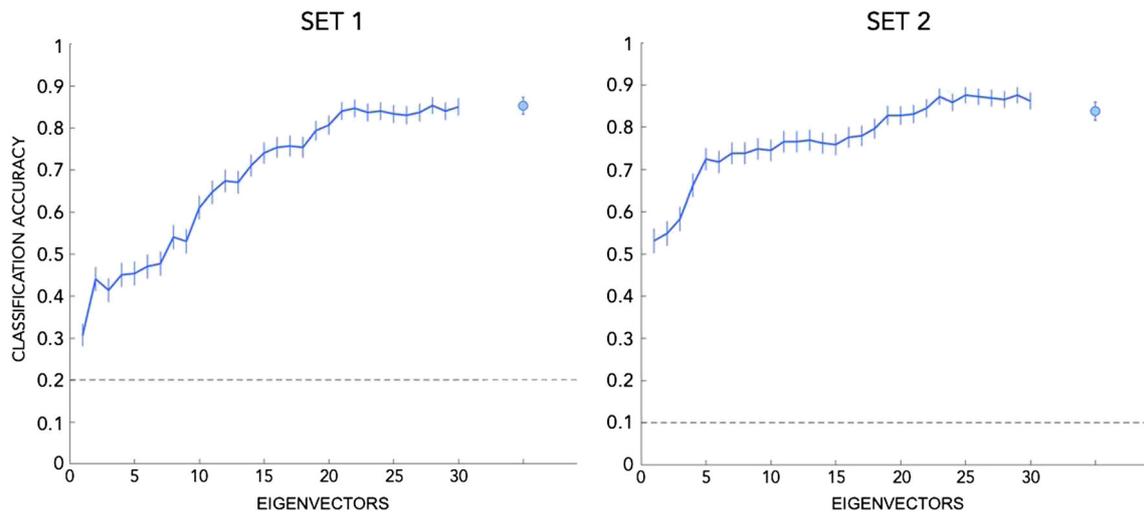


Fig. 5 Classification results. We ran a range of classification analyses using different numbers of eigen-dimensions from our concept networks. The classification was successful using ≥ 1 dimensions in both Sets 1 and

2. Classification performance increased as more dimensions were added, such that the performance of the network models reached the performance of the vector-based models (single data points)

vector-based classifier (rightmost point on each graph), which was successful at classifying the exemplars in Set 1 (mean accuracy = .85, $SD = .06$, chance = .20) and Set 2 (mean accuracy = .84, $SD = .10$, chance = .10). The successful classification of conceptual exemplars using our concept network models suggests that the structures of these networks are concept-specific. We can now extract and analyze traditional network science measures from these concept-specific networks in order to examine the relationships between network topology and flexible concept use.

Network measures of conceptual structure

We extracted network measures from 15 concept networks and explored how they relate to text-based and empirical measures of conceptual flexibility. The correlations between all network measures, along with their means and standard deviations, are shown in Table 1. Hoffman et al. (2013) used word co-occurrence statistics to quantify the context-dependent variations in word meanings found in language. To capture “semantic ambiguity” and “flexibility of word usage” in a computational framework, the authors provide a measure of semantic diversity (SemD) based on latent semantic analysis (LSA; Hoffman et al., 2013). A high-SemD item is a word that occurs in diverse language-based contexts—that is, the verbal context surrounding instances of the word are relatively dissimilar in meaning. Based on the assumption that flexibility of word usage reflects flexibility of meaning, we extracted SemD values for our 15 concepts to determine whether SemD predicts any of our network measures of interest.

We hypothesized that network modularity, network diversity, core–periphery structure, and network clustering might relate to conceptual flexibility, and we tested whether our network

measures correlated with SemD (Hoffman et al., 2013) across our 15 concepts. First we examined modularity and diversity—these measures capture the extent to which a network can be partitioned into distinct clusters of nodes (“modules”) and the extent to which individual nodes participate in these modules, respectively; SemD was not predicted by network modularity ($p > .2$), nor by network diversity ($p > .5$).

Next, we examined core–periphery structure—this measure reflects the extent to which a network can be partitioned into one densely connected core and a sparsely connected periphery; SemD was positively predicted by core–periphery fit ($r = .71$, $p = .003$; Fig. 6A). Core–periphery fit was not significantly related to either K ($p = .16$), Nc ($p > .3$), or simple stability ($p > .3$). Furthermore, core–periphery fit predicted SemD when separately controlling for each of these measures in a general linear model (all $ps < .035$). Our interpretation of the positive relationship between SemD and core–periphery fit is that the presence of a “periphery”—that is, a set of weakly associated features—relates to increased variation of potential word meanings.

The last network measure we explored was the clustering coefficient, which reflects the overall connectivity of a network; SemD was negatively predicted by clustering ($r = -.70$, $p = .004$; Fig. 6B). Clustering was not significantly related to either K ($p = .18$), Nc ($p > .4$), or simple stability ($p > .7$). Furthermore, clustering predicted SemD when separately controlling for each of these measures in a general linear model (all $ps < .04$), except for K ($p = .12$). We interpret this negative relationship between SemD and network clustering as reflecting the fact that networks with high clustering are too intracommunity to facilitate the flexible activation of varied sets of features.

Core–periphery fit and network clustering were significantly negatively related to each other ($r = -.56$, $p = .03$).

Table 1 Statistics and correlations between network measures

	Modularity	Diversity	Core-Periphery	Clustering
Mean (<i>SD</i>)	0.01 (0.002)	0.96 (0.007)	0.17 (0.025)	0.48 (0.042)
Modularity	–	–	–	–
Diversity	.154	–	–	–
Core-periphery	–.089	.446 ⁺	–	–
Clustering	–.461 ⁺	–.186	–.561 [*]	–

The mean and standard deviation across 15 concepts is shown for each network measure of interest in the top row. Beneath are the Spearman correlation values between the four network measures. The only significant relationship is between core-periphery structure and network clustering. ⁺ $p < .10$. ^{*} $p < .05$.

No other measures of interest were significantly correlated, though there was a trend toward a negative relationship between modularity and clustering ($r = -.46$, $p = .08$), and diversity and core-periphery fit were marginally positively correlated ($r = .45$, $p = .095$).

Finally, we explored whether SemD related to our simple stability measure, which reflects the extent to which the properties in a concept are represented across all of that concept's subordinates. We observed a significant negative relationship between SemD and simple stability ($r = -.52$, $p = .046$). Simple stability did not relate significantly to any of the network measures of interest ($ps > .3$).

Bootstrap analysis

We ran bootstrap analyses to test the robustness of the network-SemD relationships when only a subset of subordinates or a subset of properties was used to create the concept networks. These tests were done to confirm that

the network measures we report here are not dependent on the exact subordinates and the exact properties that went into network construction.

The results of the leave-one-out subordinate analysis are shown in Fig. 7A. For each of the four network measures of interest, the bootstrap analysis resulted in a distribution of correlations with SemD (blue histograms), along with a mean and 95% confidence interval (pink bars). The distribution of core-periphery correlations with SemD was significantly greater than zero ($p = 0$), and the distribution of connectivity-SemD correlations was significantly less than zero ($p = 0$). These results confirm that the correlations reported above are robust to variations in the set of subordinates for each concept. Though modularity was not significantly related to SemD, the distribution of SemD correlations was significantly above zero ($p = .01$). On the other hand, the distribution of diversity-SemD correlations was not significantly different from zero ($p = .23$).

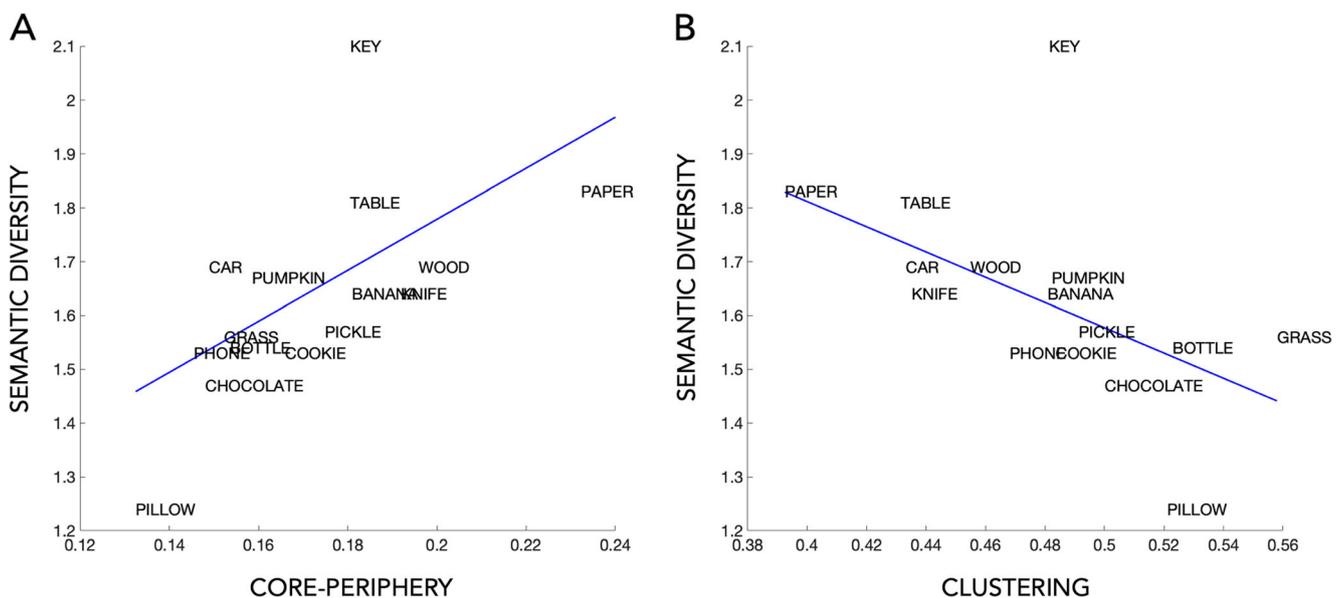


Fig. 6 Network predictors of semantic diversity. The semantic diversity measures, calculated using word co-occurrence statistics (Hoffman et al., 2013), were (A) positively predicted by core-periphery structure ($r = .71$, $p = .003$) and (B) negatively predicted by network clustering ($r = -.70$, $p = .004$)

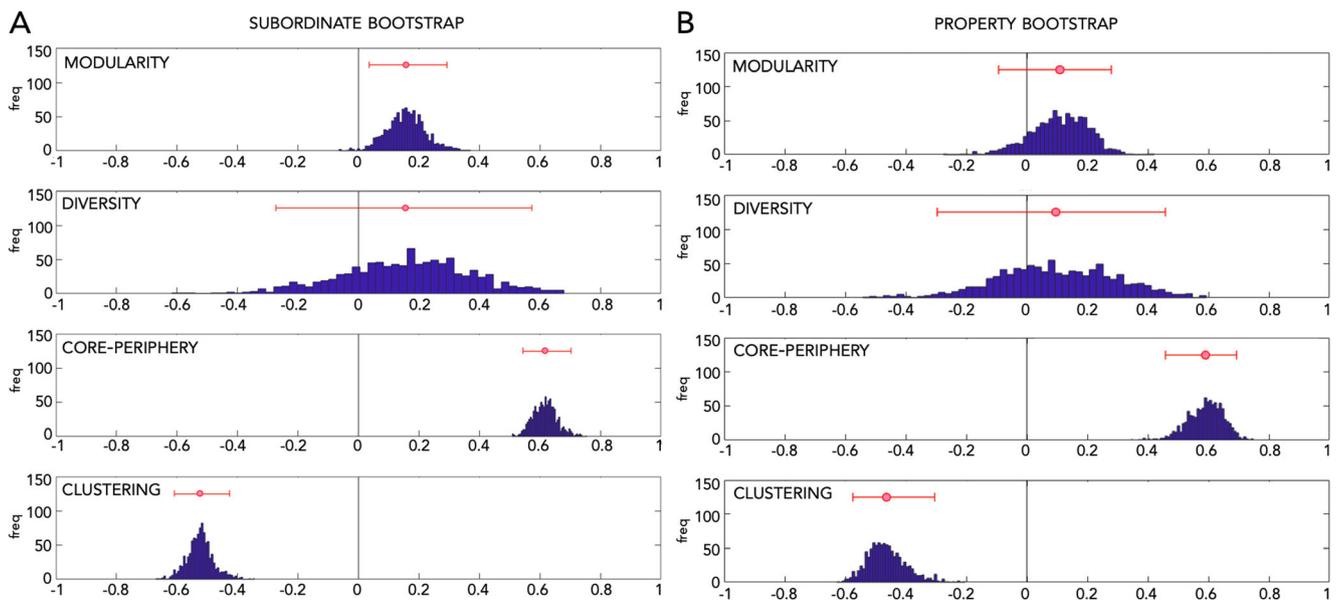


Fig. 7 Results of bootstrap analyses: Distributions of correlations between SemD and network measures when multiple networks were constructed using subsets of (A) subordinates or (B) properties. The positive relationship between core–periphery structure and SemD is robust to

variations in subordinates and properties, and the negative relationship between network clustering and SemD is also robust to variations in subordinates and properties

The results of a leave-10%-out property analysis are shown in Fig. 7B. For each of the four network measures of interest, the bootstrap analysis resulted in a distribution of correlations with SemD (blue histograms), along with a mean and 95% confidence interval (pink bars). The distribution of core–periphery correlations with SemD was significantly greater than zero ($p = 0$), and the distribution of connectivity–SemD correlations was significantly less than zero ($p = 0$). These results confirm that the correlations reported above are robust to variations in the set of properties for each concept. The distributions of SemD correlations were not different from zero for either modularity ($p = .12$) or diversity ($p = .32$).

In both of these bootstrap analyses, the variability in correlations between network diversity and SemD is striking. This instability of network diversity across bootstraps might indicate that network diversity is driven by a small number of nodes with highly varied connections whose presence varies over each iteration. However, the particular local and global structures of a concept network that contribute to network diversity—and the stability of this measure across different concepts and methods of network construction—is an open question for future work.

Similes

Interpreting a simile (e.g., “X is like a Y”) involves context-dependent activation of conceptual meaning: The explicit comparisons contained in a simile imply that X is similar to Y with respect to a certain dimension or a subset of properties of Y, thus requiring the interpreter to select a likely subset of

Y’s properties that is being asserted for X. For example, to interpret the simile “Truth is like a knife,” one must decide which properties of KNIFE can also apply to TRUTH. Since this process involves within-concept property structures and flexible conceptual meaning, we asked which of our measures, if any, predict the simile goodness measure we constructed for each of our 15 target concepts.

SemD did not predict simile goodness ($r = -.14, p > .6$). The two network measures that were significantly related to SemD—core–periphery and clustering—did not predict simile goodness, either (core–periphery: $r = .04, p > .8$; clustering: $r = .25, p > .3$), nor did network modularity ($r = -.11, p > .7$). Interestingly, network diversity positively predicted simile goodness across the 15 target concepts ($r = .54, p = .04$; Fig. 8A). Network diversity still predicted simile goodness when controlling for both K ($p = .04$) and N_c ($p = .03$) separately, as well as simultaneously ($p = .04$). However, the bootstrapped distributions of network-diversity and simile-goodness correlations were not significantly greater than zero (Supplemental Figs. 1 and 2), suggesting that this relationship might not be robust to variations in subordinates and properties used during network construction. The relationships between concept network structure and figurative-language comprehension should be further explored in future work.

Alternative uses

Generating alternative uses for common objects involves thinking about those objects in new, creative ways (Chrysikou & Thompson-Schill, 2011). Theories posit that

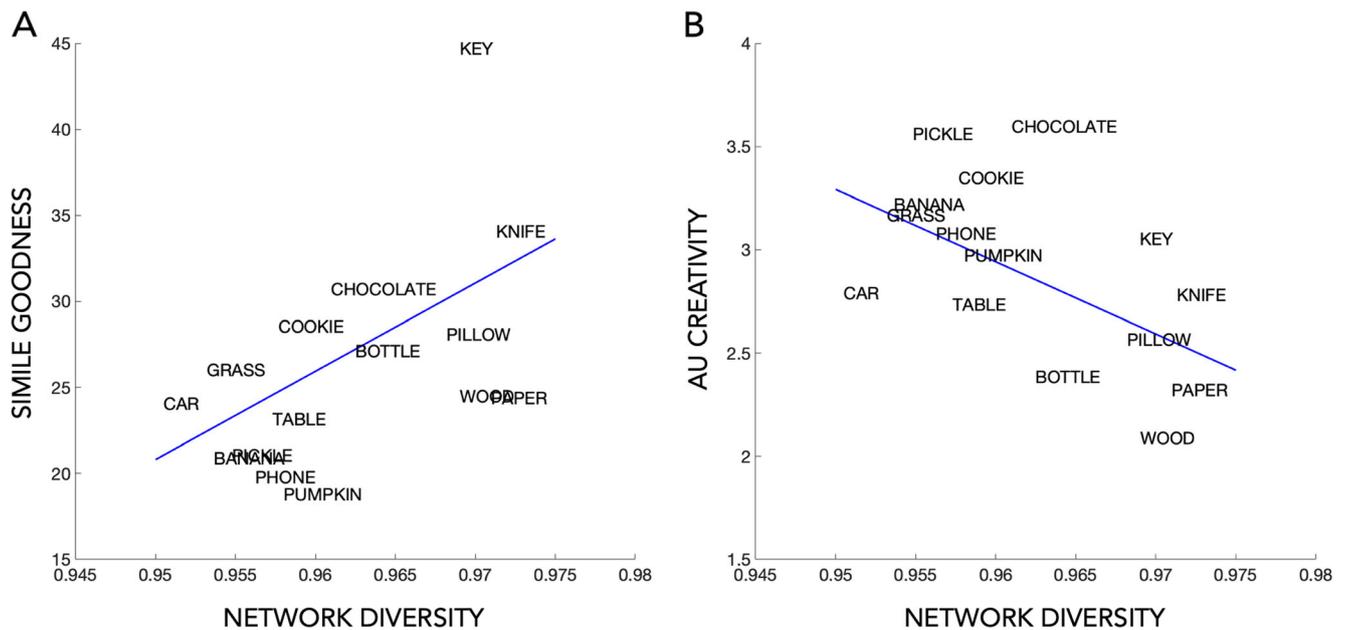


Fig. 8 Network predictors of creative concept use. (A) Network diversity positively predicted the goodness of similes containing the target concept in the vehicle position. (B) Network diversity negatively predicted the creativity of initial responses in an alternative-uses (AU) task

the generation of novel uses requires one to suppress the typical function of the object and to pay attention to its sensorimotor properties. For example, the common use of CHOCOLATE is “to eat”—to realize that it can also be used “as paint,” one must activate the CHOCOLATE properties of MELTABLE and BROWN. Since this process involves consideration of a concept’s properties and property relationships, we asked which of our measures, if any, would predict performance on an AUT.

First, we analyzed the average number of responses given for each of the 15 concepts in the AUT. SemD positively predicted the number of responses ($r = .67, p = .006$), such that if a concept label (e.g., “key”) occurs in a diverse range of textual contexts, participants will generate more potential alternative uses for that concept (e.g., will give more responses for what novel things to do with a KEY). None of the network measures predicted number of alternative-use responses (all $ps > .1$).

Second, we analyzed the average creativity ratings for the first response given for each of the 15 concepts. Mean creativity was strongly negatively predicted by the mean number of responses ($r = -.77, p < .001$), suggesting that concepts that inspire *more* responses also tend to inspire *less creative* (initial) responses. SemD did not predict mean creativity ($r = -.41, p = .13$). The two network measures that were significantly related to SemD—core-periphery and clustering—did not predict creativity, either (core-periphery: $r = -.32, p > .2$; clustering: $r = .40, p = .14$). A negative relationship between network modularity and creativity was observed, but it did not reach statistical significance ($r = -.50, p = .056$). Interestingly, network diversity was a negative predictor of creativity ($r = -.56, p = .03$; Fig. 8B). Network diversity still predicted

creativity when controlling for K ($p = .041$) and Nc ($p = .036$) separately, and it was marginally reliable when both were controlled for simultaneously ($p = .053$). As in the previous analysis, the bootstrapped distributions of network-diversity and creativity correlations were not significantly less than zero (Supplemental Figs. 3 and 4), suggesting that this relationship might not be robust to variations in the subordinates and properties used during network construction. However, the other network measures were reliably related to creativity in the bootstrap analyses, suggesting that a concept’s network structure might relate to people’s ability to use the concept in creative ways.

Discussion

Here our goal was to model basic-level concepts using graph-theoretical networks. A model structured using within-concept feature statistics provides a framework in which varied yet appropriate instantiations of a concept may be flexibly activated. An APPLE network may contain a strong connection between CRUNCHY + FRESH and between SOFT + BAKED, enabling the conceptual system to know what sets of properties should be activated in a particular APPLE instance—for example, in the representations evoked by “apple picking” versus “apple pie.” The property-covariation statistics for a given concept will determine which sets of properties tend to co-occur, and how individual properties relate to those sets and to each other. Here we have demonstrated (1) how to create these concept network models, (2) that these models are

concept-specific, and (3) structural characteristics of these networks can predict other measures of conceptual processing.

The concept network approach we describe here is a general one, and there are many different ways in which feature-based concept networks can be constructed. We have walked through one potential way to do so, in a proof of concept that reveals the feasibility and potential utility of this approach. The specific methodological decisions used in this worked-through example are described above, and we hope that other researchers interested in modeling and capturing conceptual flexibility will use variations of these methods—for example, different concepts, distance metrics, or network measures—to further explore how conceptual structure relates to flexible concept use.

There would be no point in attempting to extract concept-specific measures of conceptual flexibility if the networks themselves did not contain concept-specific feature relationships. Our approach assumes that this is the case, though this was not a theoretical certainty. In fact, many feature-based models of conceptual knowledge rely on feature correlations across the entirety of semantic space or within a large semantic domain, and represent many concepts within this single correlational feature space. For example, it could have been the case that the property BLACK relates to SOFT and ROTTEN across all concepts. However, our analyses suggest that properties relate to each other in different ways across basic-level concepts. For example, BLACK might relate to SOFT and ROTTEN in BANANA, but to FIRM and BITTER in CHOCOLATE. We found that our 15 concept networks could successfully discriminate between new conceptual exemplars, suggesting that within-concept feature statistics differ reliably between basic-level concepts. These results emerged out of a classification analysis based on eigen-decomposition of our concept networks. Eigen-decomposition of graphs has previously been used to assess the correspondences between anatomical brain network structure and patterns of functional activation (Medaglia et al., 2018); here we adapted this method to assess the correspondences between conceptual structure and feature-vectors for individual conceptual exemplars. Empirically demonstrating that networks contain concept-specific feature statistics enabled us to analyze each concept's network structure and relate structural characteristics to aspects of conceptual processing.

The kinds of structure we analyzed here included network clustering, modularity, core–periphery, and diversity. To explore whether these network structures could predict interesting aspects of conceptual processing, we examined three external measures: a text-based measure of semantic diversity (SemD; Hoffman et al., 2013), empirical measures of simile goodness, and empirical measures of creativity on an AUT. We found reliable relationships between network measures and each of these three external data sets, highlighting the potential of this approach to capture aspects of flexible concept use.

Network clustering, quantified in a *clustering coefficient*, captures the extent to which nodes are linked to its nearest neighbors. A network characterized by high clustering is one in which network nodes form “cliques” in which nearby nodes are linked to each other (Watts & Strogatz, 1998). This is intuitive in social networks, in which friends of one person tend to be friends with each other. High network clustering has been observed in text-based semantic networks (Steyvers & Tenenbaum, 2005), and semantic networks exhibit greater clustering in high- versus low-creative individuals (Kenett et al., 2014). Here, we observed that feature-based concept networks with greater clustering exhibit less text-based semantic diversity (SemD). That is, words that do not occur in many text-based contexts correspond with concepts whose features exhibit strong clustering. This result was robust to variations in properties and subordinates used in network construction. This finding suggests that dense local feature associations within a concept network reduce the extent to which word meaning can vary across instances.

Network modularity, quantified in coefficient Q , captures the extent to which a network can be partitioned into densely connected modules (i.e., sets of nodes) with sparse connections between them. Modularity is a defining characteristic of “small-world” networks (Bassett & Bullmore, 2006) and has been observed in semantic networks (Kenett et al., 2014) as well as functional brain networks (Bassett et al., 2011). Here we did not observe direct relationships between network modularity and our other measures, though our additional bootstrap analyses suggest that stronger relationships between modularity and conceptual processing may be found when a larger set of concept networks are analyzed.

Core–periphery structure, quantified in a *core-fit* measure, reflects the extent to which a network can be partitioned into one set of densely connected nodes (core), with sparse connections between all other nodes (periphery). This kind of network structure, originally observed in social networks (Borgatti & Everett, 2000), has also been observed in functional brain networks (Bassett et al., 2013). Here, we found that concept networks with stronger core–periphery structures exhibit greater text-based semantic diversity (SemD); this result was robust to variations in properties and subordinates used in network construction. This finding suggests that the presence of one set of highly associated features (“core”) in addition to a substantial set of weakly associated features (“periphery”) is predictive of conceptual flexibility. In particular, this structure might enable substantial variation in the activation of individual periphery features across instances of concept representation.

Network diversity, quantified in a diversity coefficient, reflects the extent to which nodes in a network participate in few or many network modules. This a version of a “participation” coefficient calculated using Shannon entropy (Rubinov & Sporns, 2011). In functional brain networks, these measures

are typically used to define network “hubs” (Sporns, 2014; Power, Schlaggar, Lessov-Schlaggar, & Petersen 2013), which are particularly important for transitioning between network states (i.e., patterns of activity in a network). Here, we observed that network diversity positively predicted simile goodness judgments and negatively predicted creativity of responses in an AUT. However, these results were not robust to variations in properties and subordinates used in network construction, so these specific relationships should be interpreted with caution until they are replicated in a larger set of concept networks.

We observed relationships among non-network measures that are interesting in their own right. First, we found that the distributional, corpus-based measure of SemD (Hoffman et al., 2013) was negatively correlated with our simple stability measure, which is the proportion of properties that are present in all of a concept’s subordinates. As was discussed in Landauer and Dumais (1997), two important aspects of word meaning are *usage* and *reference*; measures of distributional semantics (e.g., LSA, SemD) are constructed on the basis of usage only, and do not contain nor point to information in the world to which a word refers. Feature-based measures (e.g., McRae et al., 1997; Tyler & Moss, 2001), on the other hand, do incorporate reference into word meaning by pointing to the sets of features contained in each concept. Though distributional semantic approaches have their benefits, it is often difficult to know what their respective measures relate to from a cognitive or psychological standpoint. Our finding that the distributional, corpus-based statistic of semantic diversity was negatively related to a feature-based statistic of conceptual stability provides some insight into how usage- and reference-based measures of conceptual diversity and stability might converge.

The AUT also resulted in findings that warrant further investigation. First, we found that the SemD of a word is a positive predictor for the number of alternative-use responses participants can generate for that item. One interpretation of this finding is that if a word is found in more diverse text-based contexts it is easier to think of alternative uses of the object to which the word refers. We additionally found that the mean number of alternative-use responses for a concept is negatively correlated with the creativity of the initial alternative-use response. Though further analysis of these findings is beyond the scope of the present report, this might be a relevant finding for those interested in the AUT task and creativity more generally.

Taken together, these results reveal the ability of feature-based concept networks to capture meaningful aspects of conceptual structure and use. The analyses reported here were exploratory; we did not have any strong a priori predictions of which network measures would relate to each additional conceptual measure, and in what direction. However, we did predict that network measures—which capture different kinds

of conceptual structure—would predict the ways in which concepts are flexibly used in language and thought. We have thus demonstrated how to create concept-specific networks, and that the structures of these networks can be related to other concept-specific measures. The external measures we report here (SemD, simile goodness, AUT) are intended to serve as examples of measures that could be related to concept network structure. We look forward to future work further exploring the utility of this concept network framework in the study of conceptual knowledge.

Linking back to cognitive theories of conceptual knowledge, this concept network approach has similarities to theories that aim to characterize the flexibility of individual features. Sloman et al. (1998) captured pairwise relations between features in order to model the feature-based structure of individual concepts. These authors were interested in the role of individual features with respect to conceptual coherence, which relates to notions of centrality in the “intuitive theory” view of concepts (e.g., Keil, 1992; Murphy & Medlin, 1985). Sloman et al. (1998) simplified this previous notion of centrality by basing conceptual structure on asymmetrical dependency relationships between features; this structure captures the concept-specific “mutability” of a feature, and offers a framework in which concepts can be structured yet flexible. In the present article, our concept networks are defined by symmetrical feature co-occurrence statistics rather than asymmetrical dependency relationships. However, it would be possible to capture feature dependencies in directed concept networks (i.e., with asymmetrical links). The use of network science tools enables us to analyze not only a concept’s global structure, but also the characteristics of individual feature nodes (e.g., mutability, centrality). These kinds of structures have implications for flexible concept use such as analogies, metaphors, and conceptual combination (Sloman et al., 1998).

Though we believe that a feature-based concept network approach will provide a new set of useful tools with which to study conceptual flexibility, it is not the only way to do so. Other frameworks have the potential to capture the flexibility of the conceptual system, including attractor networks (e.g., Cree et al., 2006; Cree et al., 1999; Rodd, Gaskell, & Marslen-Wilson, 2004) and recent updates of the hub-and-spoke model (Hoffman et al., 2018; Lambon Ralph, Jefferies, Patterson, & Rogers, 2017). The concept network framework proposed here is not in opposition with these other approaches; the development and implementation of all of these methods will greatly benefit our understanding of the semantic system. However, we do believe that a network science approach to conceptual knowledge has unique advantages. The ability of graph-theoretical network science to model a vast range of systems enables us to examine conceptual structure across cognitive, linguistic, and neural levels of analysis. The structure of behavioral, feature-based networks (as discussed here)

can be analyzed and compared with the structure of functional brain networks within specific cortical sites (e.g., anterior temporal lobe; ATL) or across the brain as a whole. There is the additional possibility of analyzing “informational” brain networks, in which networks are constructed on the basis of simultaneous pattern discriminability across cortical sites (informational connectivity; Coutanche, Solomon, & Thompson-Schill, 2013). Network neuroscientists have previously forged links to cognitive processes such as motor-sequence learning (Bassett et al., 2011) and cognitive control (Medaglia et al., 2018), setting a precedent for the application of network science to cognitive neuroscience.

Recent work exploring the intersection of network science with control theory suggests another possible future direction. Network *controllability* refers to the ability to move a network into different network states; this idea has been applied to structural brain networks in order to shed light on how the brain may guide itself into easy- and difficult-to-reach functional states (Gu et al., 2015). There have been additional attempts to link brain network controllability to cognitive control (Medaglia, 2018). The application of control theory to concept networks might provide an additional way to quantify conceptual flexibility by identifying nodes that are well-

positioned to drive the brain into diverse, specific, or integrated states. Perhaps concept networks that are more controllable overall—that is, networks in which it is easier to reach varied network states—correspond to concepts that are more cognitively flexible.

So far, we have discussed mainly event context—a BANANA representation will be slightly different while one is painting as opposed to eating, and it will be different before and after a peeling event has occurred. However, language itself can provide a context: Language is inherently interactive, and the meaning of a word (i.e., the corresponding conceptual content) depends on the words surrounding it (e.g., McElree et al., 2001; Pustejovsky, 1998; Traxler et al., 2005). Researchers interested in conceptual combination aim to understand how the meaning of a combined concept (e.g., “butterfly ballerina”) can be predicted on the basis of the meaning of its individual constituents (Coutanche, Solomon, & Thompson-Schill, 2019). This is not a unique challenge for noun–noun compounds, but also for adjective–noun compounds; even the (putatively) simple concept RED has different effects when combined with the concepts TRUCK, HAIR, and CHEEKS (see Half, Ortony, & Anderson, 1976). Combining a noun concept with an adjective concept might not simply

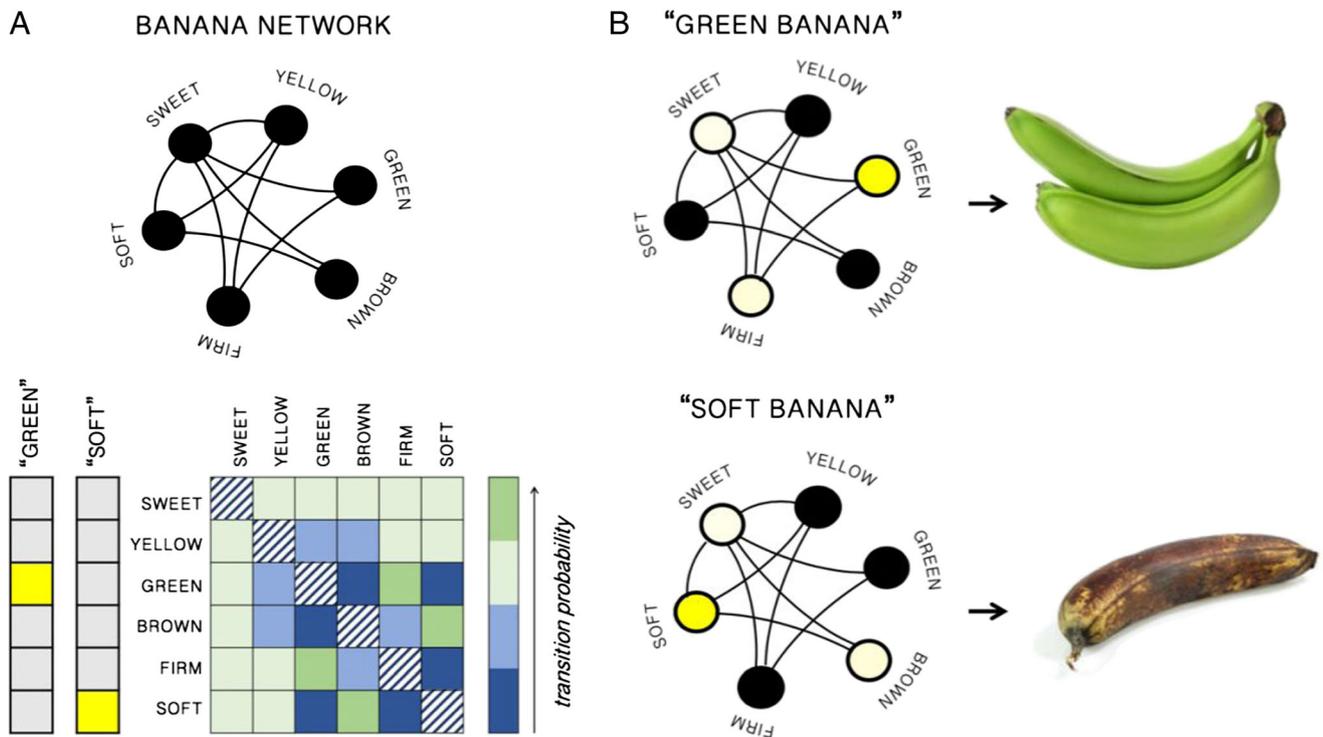


Fig. 9 Predicting conceptual combination. (A) The within-concept feature statistics of BANANA are encoded in its concept network. These data can also be used to define a transition probability matrix that encodes the probabilities that activation will spread from one node to another. We can use these data in a (e.g., random-walk) spreading-activation model to predict network states in different contexts. For example, we can predict the state of the BANANA network during adjective–noun combinations: Here, the adjectives “green” and “soft” are represented as single-node

activations. (B) Activating a single node will cause the spread of activation throughout the network according to the transition probabilities. For example, activating the GREEN node (in “green banana”) will likely cause the activation of FIRM and SWEET (top), whereas activating the SOFT node (in “soft banana”) will likely cause the activation of BROWN and SWEET (bottom). Thus, the structure of the BANANA network enables the activation of a range of states, subsequently generating varied, yet appropriate, representations of bananas

involve the reweighting of a single property node, but a more complex interaction governed by within-concept statistics.

These predictions could be generated using models of signal propagation, such as spreading activation (e.g., Collins & Loftus, 1975; De Deyne et al., 2016) or information diffusion (e.g., Bakshy, Rosenn, Marlow, & Adamic, 2012). In graph-theoretical networks, spreading activation can be formalized by simulating “random walks” over the network (Abbott, Austerweil, & Griffiths, 2015; De Deyne et al., 2016). The edges in the network can be used to define a transition probability matrix (e.g., Fig. 9A), which contains the probabilities of transitioning from one node to another node. A random walk begins at an initial node and then continues to a new node by following one of that node’s edges at random. Eventually the random walk will end (e.g., determined by a maximum path length), and the resulting path and activated nodes can be observed. This approach, which has been used to analyze word association data over large-scale semantic networks (Abbott et al., 2015; De Deyne et al., 2016), can also be used to analyze the concept networks proposed here. Random walks over concept networks could be used to predict the properties activated for a concept in a given context—starting with an initial property node, we could trace the activation of associated properties to predict the cluster of properties that would likely be activated in that instance. For example, a random walk over the BANANA network starting at the node GREEN would likely include GREEN, SWEET, and FIRM, whereas a random walk starting at SOFT would be more likely to include SOFT, SWEET, and BROWN (Fig. 9). These methods could thus be used to predict interpretations of conceptual combinations.

Finally, interesting differences might exist in the flexibility of concept networks across individuals or in the neural networks that support their processing. It has previously been suggested that individual differences in functional neural networks relate to differences in psychopathology (Lynall et al., 2010; van den Heuvel, 2013), and that the structures of semantic networks differ between low- and high-creative persons (Kenett et al., 2014). Though these networks are different in kind from the ones proposed here, the same kinds of comparisons could be explored. The ability to shift a concept network from one state to another—that is, the ability to flexibly modify the activation of certain properties—could relate to an individual’s ability to generate or comprehend novel metaphors or to generate novel uses for common objects. The flexibility of person-specific concept networks and neural networks could be explored in relation to performance on these kinds of tasks. More generally, the structure and flexibility of individuals’ concept networks might differ in meaningful ways, and this could be a fruitful avenue for future research.

We acknowledge that our proposed concept network model framework has some limitations. A large amount of data had to be collected, such that we could calculate within-concept

statistics and run a classification analysis. Here we were only able to construct and analyze 15 concept networks, a small sample with which to work. However, the primary bottlenecks were related to the classification analysis, which required that the concepts within a set be initially defined by the same set of properties, as well as requiring other tasks related to stimulus design and data analysis. Here we have provided evidence that concept networks are concept-specific, and thus classification analyses are not an essential part of the pipeline moving forward. This will greatly reduce the effort and time needed to construct concept networks in future work.

Additionally, a certain degree of experimenter subjectivity is involved in the final selection of properties and subordinate concepts. Both our properties and subordinates were reported by participants and not determined by the experimenters. However, some decisions had to be made, such as whether a response should be considered a subordinate or a property (e.g., “nuts” with respect to CHOCOLATE) and deciding an appropriate level of specificity of subordinates (e.g., excluding some brand names). To mitigate these concerns, we have reported bootstrap analyses in which subsets of subordinates or properties were excluded, in order to assess the robustness of network measures and their relationships. These kinds of analyses can be used in the future to flag any potentially idiosyncratic effects in a given data set. Despite these limitations, we believe that the construction and analysis of concept networks will provide useful insights as to the relationship between concept structure and flexible concept use.

In this proof of concept, we constructed concept network models, confirmed their ability to capture concept-specific information, and extracted network measures that relate to external text-based and behavioral measures. We believe that the application of network science to conceptual knowledge will provide a set of tools that will enable the intrinsic flexibility of the conceptual system to be explored and quantified. We hope that other researchers will be able to use these tools to further our understanding of conceptual flexibility and of the conceptual system more broadly.

References

- Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Random walks on semantic networks can resemble optimal foraging. *Psychological Review*, *122*, 558–569.
- Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 519–528). New York, NY: ACM Press.
- Barabási, A. L. (2016). *Network science*. Cambridge, UK: Cambridge University Press.
- Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*, 82–93. <https://doi.org/10.3758/BF03197629>

- Bassett, D. S., & Bullmore, E. D. (2006). Small-world brain networks. *The Neuroscientist*, 12(6), 512–523.
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108, 7641–7646.
- Bassett, D. S., Wymbs, N. F., Rombach, M. P., Porter, M. A., Mucha, P. J., & Grafton, S. T. (2013). Task-based core–periphery organization of human brain dynamics. *PLoS Computational Biology*, 9, e1003171. <https://doi.org/10.1371/journal.pcbi.1003171>
- Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 295–308. <https://doi.org/10.1037/0278-7393.19.2.295>
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424, 175–308. doi:<https://doi.org/10.1016/j.physrep.2005.10.009>
- Borgatti, S. P., & Everett, M. G. (2000). Models of core/periphery structures. *Social Networks*, 21, 375–395.
- Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 543–566). Cambridge, MA: MIT Press.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240–247.
- Coutanche, M. N., Solomon, S. H., & Thompson-Schill, S. L. (2019). Conceptual combination. In D. Poeppel, G. R. Mangun, & M. S. Gazzaniga (Eds.), *The cognitive neurosciences* (6th ed.). Cambridge, MA: MIT Press.
- Cree, G. S., McNorgan, C., & McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 643–658. <https://doi.org/10.1037/0278-7393.32.4.643>
- Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, 23, 371–414. https://doi.org/10.1207/s15516709cog2303_4
- De Deyne, S., Navarro, D. J., Perfors, A., & Storms, G. (2016). Structure at every scale: A semantic network account of the similarities between unrelated concepts. *Journal of Experimental Psychology: General*, 145, 1228–1254. <https://doi.org/10.1037/xge0000192>
- Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience*, 10, 77–94.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, 114, 211–244. <https://doi.org/10.1037/0033-295X.114.2.211>
- Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q. K., Alfred, B. Y., Kahn, A. E., . . . Bassett, D. S. (2015). Controllability of structural brain networks. *Nature Communications*, 6, 8414.
- Halff, H. M., Ortony, A., & Anderson, R. C. (1976). A context-sensitive representation of word meanings. *Memory & Cognition*, 4, 378–383.
- Hoffman, P., Lambon Ralph, M. A., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity on the basis of variability in the contextual usage of words. *Behavior Research Methods*, 45, 718–730. <https://doi.org/10.3758/s13428-012-0278-x>
- Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review*, 125, 293–328. <https://doi.org/10.1037/rev0000094>
- Huang, W., Bolton, T. A. W., Medaglia, J. D., Bassett, D. S., Ribeiro, A., & Van De Ville, D. (2018). A graph signal processing perspective on functional brain imaging. *Proceedings of the IEEE*, 106, 868–885. <https://doi.org/10.1109/JPROC.2018.2798928>
- i Cancho, R. F., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society B*, 268, 2261–2265.
- Keil, F. C. (1992). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 8, 407. <https://doi.org/10.3389/fnhum.2014.00407>
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18, 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>
- Lynall, M. E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., & Bullmore, E. (2010). Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 30, 9477–9487.
- McClelland, J. L., & Rogers T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4, 310–322. <https://doi.org/10.1038/nm1076>
- McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition*, 78, B17–B25.
- McRae, K. (2004). Semantic memory: Some insights from feature-based connectionist attractor networks. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 45, pp. 41–86). San Diego, CA: Elsevier Academic Press.
- McRae, K., Cree, G. S., Westmacott, R., & de Sa, V. R. (1999). Further evidence for feature correlations in semantic memory. *Canadian Journal of Experimental Psychology*, 53, 360–373. <https://doi.org/10.1037/h0087323>
- McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126, 99–130. <https://doi.org/10.1037/0096-3445.126.2.99>
- Medaglia, J. D. (2018). *Clarifying cognitive control and the controllable connectome*. arXiv preprint. arXiv:1801.08806
- Medaglia, J. D., Huang, W., Karuza, E. A., Kelkar, A., Thompson-Schill, S. L., Ribeiro, A., & Bassett, D. S. (2018). Functional alignment with anatomical networks is associated with cognitive flexibility. *Nature Human Behaviour*, 2, 156–164. <https://doi.org/10.1038/s41562-017-0260-9>
- Miller, G. A., & Fellbaum, C. (2007). WordNet then and now. *Language Resources and Evaluation*, 41, 209–214.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289–316. <https://doi.org/10.1037/0033-295X.92.3.289>
- Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N., & Petersen, S. E. (2013). Evidence for hubs in human functional brain networks. *Neuron*, 79, 798–813.
- Pustejovsky, J. (1998). Generativity and explanation in semantics: A reply to Fodor and Lepore. *Linguistic Inquiry*, 29, 289–311.
- Rodd, J. M., Gaskell, M. G., & Marslen-Wilson, W. D. (2004). Modelling the effects of semantic ambiguity in word recognition. *Cognitive Science*, 28, 89–104. https://doi.org/10.1207/s15516709cog2801_4

- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*, 1059–1069.
- Rubinov, M., & Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. *NeuroImage*, *56*, 2068–2079.
- Rubinov, M., Ypma, R. J., Watson, C., & Bullmore, E. T. (2015). Wiring cost and topological participation of the mouse brain connectome. *Proceedings of the National Academy of Sciences*, *112*, 10032–10037.
- Salvador, R., Suckling, J., Schwarzbauer, C., & Bullmore, E. (2005). Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philosophical Transactions of the Royal Society B*, *360*, 937–946. <https://doi.org/10.1098/rstb.2005.1645>
- Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, *32*, 3–23.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498–2504.
- Sloman, S. A., Love, B. C., & Ahn, W. K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*, 189–228.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*, 214–241. <https://doi.org/10.1037/h0036351>
- Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, *17*, 652–660. <https://doi.org/10.1038/nn.3690>
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41–78. https://doi.org/10.1207/s15516709cog2901_3
- Traxler, M. J., McElree, B., Williams, R. S., & Pickering, M. J. (2005). Context effects in coercion: Evidence from eye movements. *Journal of Memory and Language*, *53*, 1–25.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352. <https://doi.org/10.1037/0033-295X.84.4.327>
- Tyler, L. K., & Moss, H. E. (2001). Toward a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*, 244–252. [https://doi.org/10.1016/S1364-6613\(00\)01651-X](https://doi.org/10.1016/S1364-6613(00)01651-X)
- van den Heuvel, M. P., & Sporns, O. (2013). Network hubs in the human brain. *Trends in Cognitive Sciences*, *17*, 683–696.
- van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C., Cahn, W., . . . Kahn, R. S. (2013). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry*, *70*, 783–792.
- Van Rensbergen, B., Storms, G., & De Deyne, S. (2015). Examining assortativity in the mental lexicon: Evidence from word associations. *Psychonomic Bulletin & Review*, *22*, 1717–1724. <https://doi.org/10.3758/s13423-015-0832-5>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, *393*, 440–442.
- Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic Bulletin & Review*, *23*, 1015–1027.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.