Journal of Experimental Psychology:
Learning, Memory, and Cognition
2017, Vol. 43, No. 9, 1470–1489

0278-7393/17/$12.00 http://dx.doi.org/10.1037/xlm0000391

The Semantic Distance Task: Quantifying Semantic Distance With Semantic Network Path Length

Yoed N. Kenett
University of Pennsylvania

Effi Levi
The Hebrew University at Jerusalem

David Anaki and Miriam Faust
Bar-Ilan University

Semantic distance is a determining factor in cognitive processes, such as semantic priming, operating upon semantic memory. The main computational approach to compute semantic distance is through latent semantic analysis (LSA). However, objections have been raised against this approach, mainly in its failure at predicting semantic priming. We propose a novel approach to computing semantic distance, based on network science methodology. Path length in a semantic network represents the amount of steps needed to traverse from 1 word in the network to the other. We examine whether path length can be used as a measure of semantic distance, by investigating how path length affect performance in a semantic relatedness judgment task and recall from memory. Our results show a differential effect on performance: Up to 4 steps separating between word-pairs, participants exhibit an increase in reaction time (RT) and decrease in the percentage of word-pairs judged as related. From 4 steps onward, participants exhibit a significant decrease in RT and the word-pairs are dominantly judged as unrelated. Furthermore, we show that as path length between word-pairs increases, success in free- and cued-recall decreases. Finally, we demonstrate how our measure outperforms computational methods measuring semantic distance (LSA and positive pointwise mutual information) in predicting participants RT and subjective judgments of semantic strength. Thus, we provide a computational alternative to computing semantic distance. Furthermore, this approach addresses key issues in cognitive theory, namely the breadth of the spreading activation process and the effect of semantic distance on memory retrieval.

Keywords: LSA, memory recall, network science, semantic distance, spreading activation

Semantic priming is a central concept in language and memory research (Jones & Estes, 2012; Lerner, Bentin, & Shriki, 2014; McNamara, 2005; Meyer & Schvaneveldt, 1971; Neely, 1991). It refers to the finding that a target word (e.g., *tiger*) is processed faster and more accurately when it is preceded by a related prime word (e.g., *lion*) than when it is preceded by an unrelated prime word (e.g., *radio*). It is generally agreed that semantic priming is attributable to either associative relations, semantic feature overlap, or a mixture of both, between prime and target (Hutchison, 2003; Hutchison, Balota, Cortese, & Watson, 2008; Lerner & Shriki, 2014; Lucas, 2000; Masson, 1995; Thompson-Schill, Kurtz, & Gabrieli, 1998).

The cognitive mechanism, considered to enable semantic priming, is based on spreading activation models (Anderson, 1983; Collins & Loftus, 1975; McNamara, 2005; Neely, 1991). According to the spreading activation model, concepts are represented in semantic memory as nodes in a semantic network which are linked together, based on a semantic similarity principal. Concepts which are semantically related are located closer to each other and have stronger links connecting them. Processing of a concept leads to the activation of its mental representation. This activation spreads to all other concepts connected to it, quickly dissipating as the distance, namely the number of links, or associative steps, increases (Balota & Lorch, 1986; Den-Heyer & Briand, 1986; McNamara, 1992, 2005; McNamara & Altarriba, 1988). Thus, priming between a directly related prime and target (e.g., *moon* - *sun*) is greater than priming between an indirectly related prime and target (e.g., *crater* - *sun*). This finding has been interpreted as decreasing activation over intervening nodes in the semantic network (de Groot, 1983; McNamara, 1992). Computational modeling approaches have been applied to provide a plausible neural model that can account for semantic priming and spreading acti-

This article was published Online First February 27, 2017.

Yoed N. Kenett, Department of Psychology, University of Pennsylvania; Effi Levi, Institute of Computer Science, The Hebrew University at Jerusalem; David Anaki and Miriam Faust, The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, and Department of Psychology, Bar-Ilan University.

We thank Niv Reggev for his advice on the research design and his fruitful comments on the manuscript. We thank Cherut Zilberfarb, Ayal Klein, and Amit Bashari for their help in collecting the data. The majority of this work was conducted as part of the PhD research of Yoed N. Kenett. This work was partially supported by the Binational Science Fund (BSF) grant (number 2013106) to M. Faust, the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (grant 51/11) to M. Faust, and by the Israel Science Foundation (grant 841/10) to D. Anaki.

Correspondence concerning this article should be addressed to Yoed N. Kenett, Department of Psychology, University of Pennsylvania, Philadelphia, PA 19104. E-mail: yoedk@sas.upenn.edu

vation (Brunel & Lavigne, 2009; Huber & O'Reilly, 2003; Lavigne, Dumercy, & Darmon, 2011; Lerner, Bentin, & Shriki, 2012; Lerner & Shriki, 2014; Masson, 1995; Plaut, 1995). In these models, semantic memory is modeled as a recurrent neural network in which concepts are stored as units of features, or memory patterns, which form attractors in the network. Relations between concepts are implemented as correlations of memory patterns, reflecting the overlap of semantic features. Lerner and Shriki (2014) proposed such an attractor based model to account for semantic priming, based on a latching dynamics—activation in the network "jumps" from one attractor to another due to several parameters regulating the system (such as attention, strength between concepts and activation depression mechanisms).

However, these computational models focus on modeling the mechanisms of semantic priming effects, and not on how the structure of semantic memory affects such processes. For example, an important unanswered question that remains is the breadth of the spreading activation process. As the spreading activation quickly dissipates over semantic distance, how many steps between a prime and a target word can this process traverse? We attempted to address this issue by means of network science. Network science provides computational tools to examine complex systems, such as semantic memory, representing them as a graph where nodes (e.g., concepts) are linked to each other according to some organization measure (e.g., semantic similarity). Such analysis allows examining path lengths between concepts in the network, namely the number of steps traversed from one node to another. In the current study, we examined whether path length between words, derived from such a network approach, can be used as a measure of semantic distance. This is achieved by examining the effect of path length on behavioral performance in a semantic relatedness judgment and free- and cued-recall tasks. Further, we examined how word-pairs, with varying path length between them are subjectively rated for associative strength. Finally, we show how this measure outperforms the current conventional computational methods assessing semantic distance, latent semantic analysis and point wise mutual information, in predicting participants' performance.

Semantic Distance and Mediated Priming

Semantic distance is the "shortest path [direct or indirect] between two nodes" (Collins & Loftus, 1975, p. 412, note 3). As such, semantic distance can be defined as the number of steps that intervene between the prime and the target in memory. In a network model, for example, *mane* and *lion* might be directly separated by a distance of one step, whereas *mane* and *tiger* might be connected only through the mediating word *lion* and thus separated by a distance of two steps (McNamara, 1992). Therefore, one approach to investigate the role of spreading activation in semantic priming is via mediated priming (Jones & Estes, 2012). Mediated priming refers to priming for target words that are only indirectly related in semantic memory (Chwilla & Kolk, 2002; Jones, 2010, 2012; Jones & Estes, 2012; McNamara, 1992; McNamara & Altarriba, 1988). In the case of *one* intervening concept, the facilitation is referred to as two-step priming (e.g., when the prime is *lion*, the target is *stripes* and the nonpresented mediator is *tiger*). In the case of *two* intervening concepts, the facilitation is referred to as three-step priming (e.g., when the prime is *mane*, the target is *stripes* and the nonpresented mediators are *lion*, *tiger*). The majority of research in mediated priming focuses on two-step priming, and is conducted via the lexical-decision task (LDT; Meyer & Schvaneveldt, 1971). In the LDT, participants are presented with a target word which is either preceded by a related or unrelated prime word. The participants are required to decide whether the target word is a real word or not. Manipulations on the prime, target, relations between prime and target and other factors allow delicate examination of semantic priming, in general, and mediated priming in particular (Jones & Estes, 2012). Only a few attempts have been made to examine three-step priming (Chwilla & Kolk, 2002; McNamara, 1992). This is mainly attributable to difficulties in constructing valid stimulus to study this effect, where the prime and target do not share any common associate, which would lead to two-step, instead of three-step, priming (McNamara & Altarriba, 1988). The scarcity of studies on multistep priming is unfortunate since empirical examination of the effect of semantic distance on cognitive processes can shed further light on these processes and the organization of semantic memory (i.e., Lorch, 1982).

Computational Measures of Semantic Distance

To examine the effect of semantic distance on cognitive processes such as semantic priming, models representing semantic memory are needed. In the past two decades, the development of computational models to represent semantic memory has advanced rapidly (Jones, Willits, & Dennis, 2015; McRae & Jones, 2013). One main family of computational models that have been developed to model semantic representations are distributional models, which share in common the distributional hypothesis (Harris, 1970). This hypothesis states that words that appear in similar linguistic contexts are likely to have related meanings (Jones et al., 2015; McRae & Jones, 2013). In recent years a large number of corpus-based methods have been developed (Mandera, Keuleers, & Brysbaert, 2017). These methods differ in terms of how they define a word's context (e.g., the paragraph, the document), the extent to which they use grammatical information (e.g., word order), and how the meaning is represented (e.g., latent spaces, mixture models). One leading distributional model that has been used to extract semantic distance is Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998). LSA quantifies the semantic similarity between words in a given semantic space by determining the probability of a given word co-occurring in a specific context (e.g., a paragraph of a document). LSA has been empirically applied to examine semantic similarity, semantic priming, memory and creativity (Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014; Chwilla & Kolk, 2002; Coane & Balota, 2011; Green, 2016; Griffiths, Steyvers, & Firl, 2007; Howard & Kahana, 2002; Jones & Golonka, 2012; Pakhomov et al., 2010; Prabhakaran, Green, & Gray, 2014; Steyvers, Shiffrin, & Nelson, 2004). Chwilla and Kolk, for example, examined three-step priming with LSA analysis (Chwilla & Kolk, 2002). The authors show the fruitfulness of using LSA to examine more subtle differences in semantic relations between words. The results of this study suggest that LSA can be used as a method to assess semantic distance between words (Chwilla & Kolk, 2002).

However, objections have been raised at the validity of this approach as a measure of semantic distance and in predicting semantic priming (Hutchison et al., 2008; Mandera, Keuleers, & Brysbaert, 2015; Recchia & Jones, 2009; Simmons & Estes, 2006). Hutchison et al. (2008) examined whether LSA can predict semantic priming in a lexical decision and a naming task, in both short and long SOAs. The results of this study revealed that LSA did not predict priming effects in neither task or SOA. Research has indicated that performance of such models strongly depends on the choice and scope of the text corpus used, which can become the determining factor in how well such models capture human performance (Recchia & Jones, 2009). Further, it is yet to be determined the validity of estimating semantic distance based on analysis of textual corpora, as opposed to using free associations (De Deyne, Kenett, Anaki, Faust, & Navarro, 2016; De Deyne, Verheyen, & Storms, 2016).

An alternative computational method to measure semantic distance is based on Pointwise Mutual Information (PMI; Church & Hanks, 1990; Paperno, Marelli, Tentori, & Baroni, 2014; Recchia & Jones, 2009). PMI measures the ratio between the probability of observing word x and word y together (their joint probability) and the probability of observing word *x* and word *y* independently (Church & Hanks, 1990). It has been shown to yield high performance on forced-choice tests of semantic similarity (Budiu, Royer, & Pirolli, 2007; Bullinaria & Levy, 2007; Terra & Clarke, 2003; Turney, 2001). However, to the best of our knowledge currently only one study examined how PMI predicts semantic relatedness judgments (Recchia & Jones, 2009). This study demonstrated how a PMI measure trained on Wikipedia corpora outperformed several publically available measures of semantic relatedness (Recchia & Jones, 2009). However, similar to LSA, PMI is also highly dependent on the type and size of the training corpora. Furthermore, the validity of using such a measure based on textual corpora is still debated (De Deyne, Kenett, et al., 2016; De Deyne, Verheyen, et al., 2016).

Measuring Semantic Distance as Path Length in a Semantic Network

A new and different computational approach to assess semantic distance is to analyze the structure of semantic memory with network science tools. Network science is based on mathematical graph theory, providing quantitative methods to investigate complex systems as networks (Baronchelli, Ferrer-i-Cancho, Pastor-Satorras, Chater, & Christiansen, 2013; Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006; Borge-Holthoefer & Arenas, 2010b). A network comprises a set of nodes, which represent the basic unit of the system (e.g., semantic memory) and links, or edges, that signify the relations between them (e.g., semantic similarity). Thus, network science applied at the cognitive level can directly and quantitatively examine classic cognitive theory on language and memory being structured as networks (Baronchelli et al., 2013).

A growing body of research uses network science tools at the cognitive level to investigate the structure of language and memory (Baronchelli et al., 2013; Borge-Holthoefer & Arenas, 2010b; De Deyne, Kenett, et al., 2016; Karuza, Thompson-Schill, & Bassett, 2016). For example, network science in cognitive science has enabled the direct examination of the theory that high creative individuals have a more flexible semantic memory structure (Kenett, Anaki, & Faust, 2014; Kenett, Beaty, Silvia, Anaki, & Faust, 2016), identified mechanisms of language development through preferential attachment (Hills, Maouene, Maouene, Sheya, & Smith, 2009; Steyvers & Tenenbaum, 2005), have shown how specific semantic memory network parameters influence memory retrieval (Vitevitch, Chan, & Goldstein, 2014; Vitevitch, Chan, & Roodenrys, 2012; Vitevitch, Goldstein, & Johnson, 2016), and provided new insight on the structure of semantic network of second language in bilinguals (Borodkin, Kenett, Faust, & Mashal, 2016). More recently network science at the cognitive level has been applied to examine atypical thought processes exhibited by clinical populations. Such research is shedding new quantitative light on these populations, such as relating rigidity of thought expressed in persons with autism to rigid semantic network structure (Kenett, Gold, & Faust, 2016), quantifying differences in semantic network related to atypical language development (Beckage, Smith, & Hills, 2011; Kenett et al., 2013), and reshaping the diagnostic definitions of psychopathology (Borsboom & Cramer, 2013; Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011).

Kenett et al. (2011) have recently introduced a novel approach to the study of semantic networks. Their approach uses correlation and network methodologies to define semantic similarity between concepts in the semantic network. The core idea of this method is the definition of connections between concepts in the semantic network as the overlap of associative responses generated to these concepts (see Cilibrasi & Vitanyi, 2007 for a similar approach). This notion is in accordance with Collins and Loftus (1975) definition of semantic similarity, and is thus a cognitive inspired approach to examine semantic memory structure (Kenett et al., 2011). Currently, growing evidence points to the strong coupling between associative and semantic relations (McRae, Khalkahli, & Hare, 2012). As such, it is plausible to use analysis based on association data to investigate semantic distance.

One main measure of networks is path length. Path length represents the amount of steps (nodes being traversed) needed to be taken between any pair of nodes in the network. As many paths can exist between two nodes in the network, what is usually measured is the shortest path between any pair of nodes. Thus, path length may be related to semantic distance and predict semantic priming effects. This notion, however, has not yet been examined. Furthermore, network science examines how the structure of a network influences the dynamic processes operating upon it (Watts & Strogatz, 1998). In this regard, certain processes might operate more efficiently on a network with a certain structure (Faust & Kenett, 2014; Kenett et al., 2014; Schilling, 2005; Vitevitch et al., 2014). Thus, path length between concepts, extracted from network analysis, may be related to other cognitive processes operating over semantic memory. One such example is research on false memories, as elicited in the Deese-Roediger-McDermott paradigm (Gallo, 2006). In this regard, research has shown how spreading activation and associative structure lead to false memories (Hutchison & Balota, 2005; Meade, Watson, Balota, & Roediger III, 2007). Another such example is memory search, examined with the free-recall task.

Path Length, Semantic Distance, and Free Recall

In the free-recall task, participants study a sequence of individually presented items. At a later testing stage, they are asked to recall as many items they can remember in any order (Kahana, Howard, & Polyn, 2008). Analyzing the order in which participants recall list items provides insight into the search processes operating upon memory. Memory search in free recall is considered a multiple constrained process. According to this view, the probability of recalling an item and the order in which these items are recalled, are influenced by various factors (Polyn, Norman, & Kahana, 2009). One factor investigated is the effect of a semantic component on free recall performance. These studies usually use free association norms to study the effect of semantic memory structure on the dynamics of free recall. Howard and Kahana (2002), for example, have shown how semantic distance, measured with LSA, is related to free recall performance. They found that closely related words to a recalled word have a high probability to be recalled as well. However, the use of path length to predict free recall performance has not yet been investigated. Previous research using a similar computational approach based on proximity responses—known as Pathfinder (Schvaneveldt, Dearholt, & Durso, 1988)—has been used to examine performance in serial- and free-recall (Cooke, Durso, & Schvaneveldt, 1986). This research found that this method, estimating semantic memory structure, better facilitates list learning, as exhibited in better serial recall performance, and was related to participants responses in a free-recall task (Cooke et al., 1986). However, to the best of our knowledge, a more general network account that directly relates memory recall and semantic priming to path length is lacking.

Current Study

In the current study we examine how path length can be used as a measure of semantic distance. To examine such hypotheses, we relied upon a large network analysis of the Hebrew mental lexicon (Kenett et al., 2011). This study analyzed the organization of associative responses to 800 cue words in Hebrew, resulting in a semantic network representing the organization of these words in the Hebrew mental lexicon. In this network, nodes represent the 800 cue words and links represent semantic relations between cue words, based on the overlap of associative responses to these cue words. From this network analysis, a distance matrix is constructed which represents the shortest amount of steps connecting any pair of cue words. As such, semantic distance is operationalized as the shortest path length between a pair of words, as represented in the distance matrix. To study the effect of path length on semantic priming, we used a semantic relatedness task. In this task, participants are presented with a word-pair and are required to judge whether the two words are related to each-other or not (Faust & Lavidor, 2003). This task was chosen to minimize any possible confounds, such as those found with regard to using the LDT in mediated priming research (Balota & Lorch, 1986; McNamara & Altarriba, 1988). Moreover, natural language is focused on meaning-level integration processes, such as those required in the semantic relatedness judgment task (Balota & Paul, 1996; Faust & Lavidor, 2003). Furthermore, recent research on memory retrieval has shown the relation between semantic relatedness and the congruity effect (Bein et al., 2015). The congruity effect refers to better memory performance for items that are presented within a compatible, rather than incompatible, semantic context (Craik & Tulving, 1975). In a series of studies, Bein et al. (2015) show how the congruity effect and semantic relatedness share a common mechanism, which they claim is related to the structure of the mental lexicon (see also Epstein, Phillips, & Johnson, 1975; Mathews, Maples, & Elkins, 1981). In the semantic relatedness task used here, we manipulated path length between prime and target words, by choosing word-pairs with varying path lengths based on the network analysis of Kenett et al. (2011). In a series of experiments, we examined how long (Experiment 1) and short (Experiment 2) path lengths are related to behavioral performance in semantic relatedness judgments and retrieval from memory tasks (see also De Deyne, Navarro, Perfors, & Storms, 2016 for a related study). We predicted that short path lengths (i.e., close or strong semantic relations) would be judged as semantically related whereas long path lengths (i.e., far or weak semantic relations) would be considered as unrelated. According to spreading activation models (Anderson, 1983; Collins & Loftus, 1975), we predicted slower RT in semantic judgments as path length (distance) grows. In addition, in line with the relation between spreading activation and memory retrieval, we examined whether path length can predict memory retrieval, either by free- (Experiments 1 and 2) or cued-recall (Experiment 3). In accordance with the spreading activation model (Collins & Loftus, 1975), we predicted a reduction in success of recalling of cue words as path length grows. Finally, we examined the performance of path length, PMI, and LSA measures in predicting participants' RT in the semantic relatedness task and in subjective judgments of the semantic strength of the word pairs. In accordance with recent comparisons of textual versus behavioral based networks (De Deyne, Verheyen, et al., 2016), we predicted that our path length measure of semantic distance will outperform LSA and PMI in predicting participants performance (Experiment 4).

Experiment 1

In Experiment 1 we examined whether path length can be used as a measure of semantic distance. To the best of our knowledge, this is the first of such an attempt. As such, the relation between path length and semantic distance is yet to be determined. Thus, in Experiment 1 we examined relatively long path lengths in a semantic relatedness judgment and free-recall tasks (see Table 1). Because of individual differences, in this task there are no correct and incorrect responses: While one participant may judge a specific word-pair as related, another might judge it as unrelated. To examine the validity of this task, we a priori assigned extreme conditions as related (1-step condition) and unrelated (20-step condition). We also examined two intermediate path lengths of 5- and 10- steps between word-pairs. These two intermediate levels were used to examine the breadth of the spread of activation, and whether participants will judge these word-pairs as related or as unrelated. We predicted that directly related word-pairs (1-step) will have the shortest RT in judging their relatedness and will have the highest free-recall performance.

Method

Participants. Thirty-eight participants were initially recruited to Experiment 1. Nine participants were removed due to either low accuracy rates (<50%) or extremely slow RTs (i.e., average RT greater than 2.5 standard deviations than the group mean) in the a priori 1-step condition. Analysis was performed on the remaining

Table 1
An Example of Stimuli Used in Both Experiment 1 (1-, 5-, 10-, and 20-Step Conditions) and Experiments 2 and 3 (1-, 2-, 3-, 4-, 6-, and 15-Step Conditions) and the Path of Words Connecting Prime and Target

Distance	Path
1-step	*bus*-*car*
2-step	*letter*-homesick-*family*
3-step	*elevator*-ladder-Persian lilac-*bench*
4-step	*storm*-bay-sunset-pale blue-*kite*
5-step	*sand*-tent-pleasure-excited-happy-*joke*
6-step	*cheater*-greedy-fortune-coin-pants-linen-*carpet*
10-step	*clean*-faucet-puddle-coat-sailboat-bay-view-backpack-pleasure-happy-*surprise*
15-step	*swan*-daffodil-stem-zucchini-noodle-pot-fed-aid-charity-luxury-jewel-earlobe-listened-tune-flute-*cymbals*
20-step	*clown*-laughter-pleasure-tent-dune-bay-sailboat-coat-puddle-water sprinkler-flowerbed-zucchini-noodle-pot-fed-aid-charity-luxury-tunnel-cricket-*mosquito*

Note. Words were translated to English, and prime and target words are marked in italics.

29 participants (13 males, 16 females; mean age 24.3 ($SD = 3.8$). All participants had normal or corrected to normal vision. Participants either took part in the experiment for partial fulfillment of academic credit or were paid an equivalent of 8 USD for their participation. All participants were dominantly right-handed, with a mean score of 90 ($SD = 9.9$) on the Edinburgh Handedness Inventory (Oldfield, 1971). This experiment was approved by the Bar-Ilan University institutional review board.

Stimuli and tasks.

The semantic distance task. the semantic distance task (SDT) is a semantic relatedness judgment task. In this task, subjects are required to decide whether two words are related to each other or not. The stimuli were taken from a network analysis of the Hebrew mental lexicon, which examined the organization of 800 cue words in the Hebrew mental lexicon (Kenett et al., 2011). This was achieved by computing the association correlations, namely, the overlap of associative responses, between each pair of cue words. The result of this analysis is an 800 × 800 connectivity (adjacency) matrix, which denotes the association correlation between every pair of nodes in the network. This connectivity matrix is symmetrical and sparse, as each node is only directly connected to a small amount of other (neighboring) nodes (Kenett et al., 2011). Because the authors were interested in the global, structural properties of the Hebrew lexicon semantic network, they binarized the association correlations so that all weights equal "1" and analyzed the network as an unweighted, undirected network (Kenett et al., 2011). As such, a link between a pair of cue words represents a symmetrical relation between them. From this binarized connectivity matrix, paths can be calculated, based on the amount of discrete steps needed to be taken from cue word i to cue word j. Because the network is unweighted and undirected, many paths can lead from one cue word to another. However, paths between a pair of nodes in the network is examined based on the shortest path connecting them (Boccaletti et al., 2006). Based on the connectivity matrix, a distance matrix can be constructed which denotes the shortest amount of steps separating cue word i from cue word j in the network. Word-pairs were then constructed from this distance matrix, which spans from a distance of 1 to 22 steps separating nodes in the network. Binarizing the links between words in the network from their association correlations to a uniform weight (of "1") may lead to loss of important information conveyed in the correlation weights. To control for this possibility, we examined the average weighted distance between word pairs for all word-pair conditions constructed. This revealed a similar linear trend—as the unweighted distance grew, so did the average weighted distance. Thus, we chose to remain with the more simple unweighted, binarized distance matrix.

The task consisted of four conditions, each containing 40 word-pairs – 1-step (word pairs directly connected), 5-step (five steps separating the word pair), 10-step (10 steps separating the word pair) and 20-step (20 steps separating the word pair). The 1-step condition was a priori considered as related word-pairs and the 20-step condition was considered a priori as unrelated word-pairs. Words were chosen so that a word appeared only once in the sample, either as a prime or a target (see Table 1 for three examples in each condition). Words were matched for length, frequency, and concreteness.

To examine the validity of the stimuli, we examined the correlation between path length and subjective ratings of the associative strength of the word-pairs. Eleven independent judges judged the associative strength of the word pairs on a 7-point Likert scale (1 – *unrelated* to 7 – *strongly related*). The participants were matched to the sample of participants in the experiment, but did not take part in it. Next, we conducted a Pearson correlation analysis between these two variables. This analysis revealed a significant negative correlation between associative distance and subjective judgments of associative strength, $r(160) = -.59$, $p < .01$.

Face sex recognition distraction task. A distraction task was used to separate between the SDT and the free-recall tasks. A shortened version of the paradigm used in Kenett, Anaki, and Faust (2015) was used. Participants saw color pictures of natural faces with neutral expressions, randomly appearing either on the left or right side of the screen, and had to recognize the sex of the face by pressing a button. The stimuli for this task comprised of 30 faces, equally divided into male and female.

Free recall. A free recall paradigm was used to examine the relation between distance and free memory retrieval. After completion of the distraction task, participants were required to try and recall as many of the words they could remember, which were presented to them during the SDT. Participants were instructed to recall any word they could think of, regardless of whether it was the prime or the target of a specific word pair. Furthermore, the participants were encouraged to try and recall as many word pairs they could. Participants were not aware during the SDT or distractor task that they will undergo the free-recall task.

Procedure. Participants sat 50 cm from a CRT screen. Both the SDT and distractor tasks were conducted using the E-prime software (Schneider, Eschman, & Zuccolotto, 2002). The stimuli were presented against a black screen. First, the participant completed the SDT, after being instructed and provided examples of the task. In the SDT, each trial began with a fixation cross appearing in the center of the screen for 80 ms. Next, the prime word appeared for 120 ms. Following the presentation of the prime word, a second fixation cross appeared in the center of the screen for 80 ms. Finally, the target word appeared for 120 ms. The participant decided whether the pair of words were related to each other or not by pressing a button. Participants were instructed to use their right hand to make their decision, using the index and middle fingers to indicate related and unrelated decisions. Once the participant pressed the button, the next trial was immediately initiated. Next, participants completed the distractor task. For in the face sex recognition distraction task, each trial began with a fixation cross appearing at the center of the screen for 200 ms. Next, the stimulus appeared for 120 ms to either the left or right of the center of the screen, with the inside edge of the stimulus presented 1.5° from the central fixation point. After the stimulus disappeared, the participant was required to judge the sex of the face, by pressing one button for male and another for female. Once the participant pressed the button, the next trial was immediately initiated. Finally, participants completed the free-recall task. For the free-recall task, participants were instructed to recall as many of the words presented in the SDT task. Participants were encouraged to try as hard as they could, by promising monetary rewards for the participant with the first, second and third highest score in the sample. Participants had 10 minutes for this task.

Results

Trials in which response time (RT) was lower than 250 ms were removed. In addition, for each participant, trials which were above or below 2.5 *SD* for each condition were also deleted from final data analysis. To verify the validity of the a priori conditions, we conducted an item-analysis on the accuracy of the sample to make the a priori relatedness judgment. Word-pairs for which group accuracy was lower than 50% were removed from final analysis (six word-pairs from the 1-step condition and 1 word-pair from the 20-step condition). To determine whether the non a priori conditions should be classified as related or unrelated word-pairs, we examined the tendency of participants to judge the 5-step and 10-step word-pairs as related or unrelated. This examination revealed that both 5-step (91%) and 10-step (92%) word-pairs are strongly considered as unrelated word pairs. Thus, these word-pairs were analyzed as unrelated word pairs.

A Distance (1-step, 5-step, 10-step, and 20-step) repeated measures ANOVA was conducted to examine the effect of distance on participants (p) and item (i) mean SDT RT. RT was analyzed for only successful trials based on condition classification (a priori and post-priori), as described above. This analysis revealed a significant main effect of Distance, $F_p(3, 84) = 3.981$, $p < .011$, $\eta^2 = .124$; $F_i(3, 160) = 6.613$, $p < .001$, $\eta^2 = .11$ (Table 2 and Figure 1a). Post hoc analyses (corrected for multiple comparisons) revealed that this difference stems from a significant rise in RT when distance grows from 1- to 5-step ($p_p < .04$ and $p_i < .001$) and a significant reduction in RT for items when distance grows from 10- to 20-step ($p_p < .6$ and $p_i < .01$). No significant differences were found when distance grows from 5- to 10-step ($p_p < .6$ and $p_i < .34$).

A distance (1-step, 5-step, 10-step, 20-step) repeated measures ANOVA was conducted to examine the effect of distance on mean free recall performance. This analysis was performed on participants only due to the low number of words recalled. This analysis revealed a significant main effect of Distance, $F(3, 94) = 49.921$, $p < .01$, $\eta^2 = .641$ (Table 2 and Figure 1b). Post hoc analyses (corrected for multiple comparisons) revealed that this difference stems from a significant difference between the free recall performances of distance of 1-step compared with all other conditions (all p's $< .01$). Similar independent analyses for recall of only the prime words, target words, or word pairs revealed similar results.

Table 2
RT, % of Word-Pairs Judged as Related/Unrelated, and Amount of Words Successfully Recalled in Free Recall as a Function of Distance in Experiment 1 (SD in Parentheses)

Distance	SDT RT	% un/ related	Free recall
1-step	699 (141)	.11/.89 (.07)	14.14 (7.37)
5-step	792 (284)	.91/.09 (.07)	5.38 (3.43)
10-step	802 (262)	.92/.08 (.09)	4.38 (2.37)
20-step	758 (244)	.92/.08 (.09)	4.41 (3.36)

Note. SDT RT = average RT in ms in SDT task; free recall = average amount of words recollected; 1-step = word pairs directly linked to each other; 5-step = word pairs with five steps between them; 10-step = word pairs with ten steps between them; 20-step = word pairs with twenty steps between them.

Discussion

In Experiment 1 we examined the use of path length between word-pairs as a measure of semantic distance. Relatively long path lengths were used to examine how they relate to performance on semantic judgments and free-recall tasks. This was achieved in a semantic relatedness judgment task where participants had to decide whether pairs of words are related to each other or not. The word-pairs were created based on quantitative path lengths (amount of steps being traversed) taken from a large network analysis of the Hebrew lexicon (Kenett et al., 2011). A distance matrix was constructed, representing the shortest amount of steps connecting any pair of the 800 cue words analyzed. Semantic distance was operationalized from the path length computed in this distance matrix. We used word-pairs that were either directly related (1-step), separated by five steps (5-step), separated by 10 steps (10-step), and separated by 20 steps (20-step). To examine the validity of our stimuli, we a priori assigned the 1-step word-pair condition as related and the 20-step word-pair condition as unrelated. Two intermediate path length conditions (5-step and 10-step) were also used, and were assigned as unrelated word-pairs post hoc, based on participants' performance.

The results show a differential relationship between path length and RT. A significant slowing down of RT in SDT was found when moving from processing 1-step to 5-step word-pairs. This prolonged RT remained relatively constant at 10-step and then significantly shortened when processing the 20-step word-pairs. Furthermore, we show how path length is related to performance on free recall word retrieval. Participants were mainly able to recall 1-step words compared with 5-, 10-, and 20-step word-pairs. Finally, we found that path length strongly correlated with subjective judgments of associative strength of the word-pairs. More generally, these results indicate that the breadth of the spreading activation process when attempting to relate between two words is bounded by at least five steps. The results of Experiment 1 provide initial support for using path length as a measure of semantic distance. However, the results indicate a very gross difference, between 1-step and all other conditions. Can path length also be used as a predictor for shorter semantic distances? This issue was examined in Experiment 2.

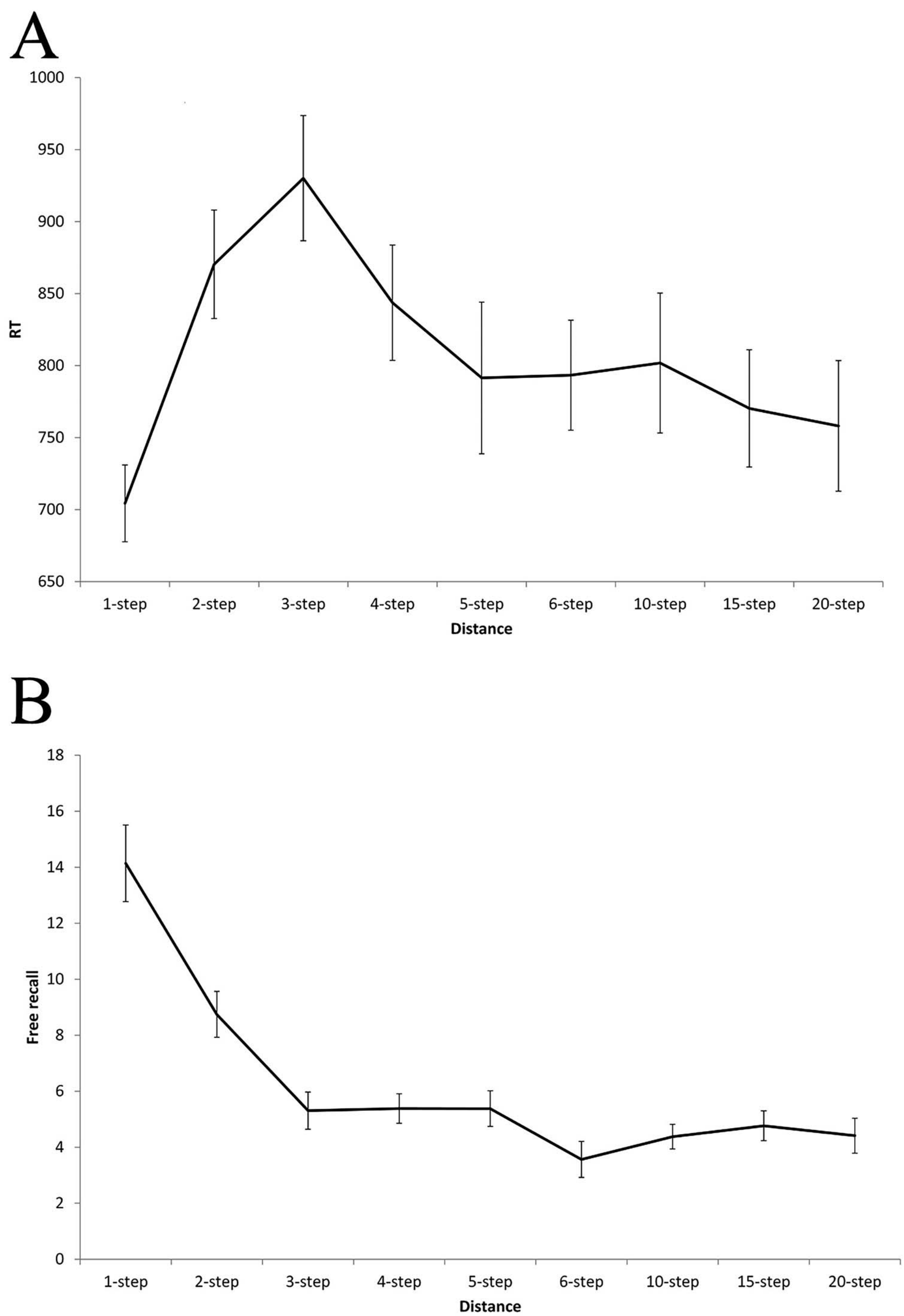

Figure 1. Effect of path length on RT (A) and free recall (B), collapsed from Experiments 1 and 2 (Experiment 1: 1-, 5-, 10-, and 20-step; Experiment 2: 1-, 2-, 3-, 4-, 6-, and 15-step). *x* axis—the different conditions varying in path length (distance). *y* axis—dependent variables, including error bars (RT and mean amount of words recalled in free recall).

Experiment 2

The results of Experiment 1 indicated that the breadth of the spread of activation is at least five steps. Therefore, in Experiment 2 we examined the use of shorter path length (1-, 2-, 3-, and 4-step distances between word-pairs) to determine a more accurate extent of activation spread (see Table 1). Furthermore, to replicate the findings of Experiment 1, we added two additional conditions of 6- and 15-step word-pairs. Based on the findings of Experiment 1, the 1-step condition was a priori classified as related word-pairs and the 6- and 15-step conditions were a priori classified as unrelated.

The 2-, 3-, and 4-step conditions were to be assigned as related or unrelated post hoc. This was achieved by examining whether participants judged the word-pairs in these conditions as related or not.

In accordance with the spreading activation model (Balota & Lorch, 1986; Collins & Loftus, 1975; Den-Heyer & Briand, 1986), we predicted an increase in RT and a decrease in percentage of word-pairs judged as related in the SDT, as path length grows until 4 steps. Furthermore, in accordance with our results in Experiment 1, we predicted a drop in RT for the 6- and 15-step word-pairs. Finally, similar to Experiment 1, we predicted that path length will be related to free recall performance—as path length grows, successful retrieval of cue words diminishes.

Method

Participants. Forty-four participants were initially recruited to Experiment 2. Five participants were removed because of either low accuracy rates (<50%) or extremely slow RTs (average RT greater than 2.5 standard deviations than the group mean) in the a priori 1-step condition. Analysis was performed on the remaining 39 participants (7 males, 32 females; mean age 22.1 [$SD = 3.2$]). All participants had normal or corrected to normal vision. Participants either took part in the experiment for partial fulfillment of academic credit or were paid an equivalent of 8 USD for their participation. All participants were dominantly right-handed, with a mean score of 90 ($SD = 8.7$) on the Edinburgh Handedness Inventory (Oldfield, 1971). This experiment was approved by the Bar-Ilan University institutional review board.

Stimuli and tasks. The tasks used were similar to Experiment 1. In this experiment, the SDT consisted of six conditions, each containing 40 word-pairs – 1-, 2-, 3-, 4-, 6-, and 15-step. The word-pairs of the 1-step condition were the stimuli used in Experiment 1 and word-pairs removed from analysis in Experiment 1 (6 word-pairs) were replaced with new stimuli. All other conditions were constructed from Kenett et al. (2011). Words were chosen so that a word appeared only once in the sample, either as a prime or a target (see Table 1). Words were matched for length, frequency, and concreteness. Similar to Experiment 1, a distractor task and a free recall paradigm were used. The 1-step condition was a priori considered as related word-pairs and the 6- and 15-step conditions were considered a priori as unrelated word pairs. Finally, to examine the validity of our stimuli, we examined the correlation between path length and subjective ratings of the associative strength of the word-pairs. Ten independent judges judged the associative strength of the word pairs in relation to their path length. The participants were matched to the sample of participants in the experiment, but did not take part in it. Next, we conducted a Pearson correlation analysis between these two variables. This analysis revealed a significant negative correlation between associative distance and subjective judgment of associative strength, $r(240) = -.61$, $p < .01$.

Procedure. The procedure of Experiment 2 was similar to that of Experiment 1. Participants sat 50 cm from a CRT screen. Both the SDT and distractor tasks were conducted using the E-prime software (Schneider et al., 2002). After completion of the distractor task, participants completed the free-recall task. Participants had 10 minutes for the free-recall task.

Results

Trials in which RT was lower than 250 ms were removed. In addition, for each participant, trials which were above or below 2.5 *SD* for each condition were also deleted from final data analysis. To determine whether the non a priori conditions should be classified as related or unrelated word-pairs, we examined the tendency of participants to judge the 2-, 3-, and 4-step word-pairs as related or unrelated. This examination revealed that both 2-step (69%) and 3-step (64%) word-pairs are more considered as related word-pairs, whereas the 4-step condition is strongly considered as unrelated word-pairs (86%). Thus the 2- and 3-step conditions were analyzed as related word-pairs and the 4-step condition was analyzed as unrelated word-pairs.

A Distance (1-step, 2-step, 3-step, 4-step, 6-step, and 15-step) repeated measures ANOVA was conducted to examine the effect of distance on participants (*p*) and item (*i*) mean SDT RT. RT was analyzed for only successful trials based on condition classification (a priori and post-priori), as described above. This analysis revealed a significant main effect of Distance, $F_p(5, 190) = 20.261$, $p < .01$, $\eta^2 = .348$; $F_i(5, 230) = 22.198$, $p < .001$, $\eta^2 = .325$ (Table 3 and Figure 1a). Post hoc analyses (corrected for multiple comparisons) revealed that this difference stemmed from a significant rise in RT when distance grows from 1- to 2-step step ($p_p < .01$ and $p_i < .001$), and from 2- to 3-step, step ($p_p < .03$ and $p_i < .004$). Furthermore, a significant reduction in RT was found only across participants when distance grows from 3- to 4-step step ($p_p < .01$ and $p_i < .4$), and both across participants and items when the distance grows from 4- to 6-step step ($p_p < .01$ and $p_i < .001$). No significant differences were found when distance grew from 6- to 15-step step ($p_p < .2$ and $p_i < .8$).

A Distance (1-step, 2-step, 3-step, 4-step, 6-step, and 15-step) repeated measures ANOVA was conducted to examine the effect of distance on mean free recall performance. This analysis revealed a significant main effect of Distance, $F(5, 190) = 18.462$, $p < .01$, $\eta^2 = .327$ (Table 3 and Figure 1b). Post hoc analyses (corrected for multiple comparisons) revealed that this difference stems from a linear function relating decline in free recall performance related to growing distance: no significant reduction in successful recall of cue words was found between free recall performance as distance grew from 1-step to 2-step ($p < .2$). In

Table 3
RT in the SDT, % of Word-Pairs Judged as Related/Unrelated, and Amount of Words Successfully Recalled in Free Recall as a Function of Distance in Experiment 2 (SD in Parentheses)

Distance	SDT RT	% un/ related	Free recall
1-step	704 (169)	.09/.91 (.08)	9.79 (5.38)
2-step	870 (239)	.31/.69 (.14)	8.74 (5.19)
3-step	930 (275)	.36/.64 (.14)	5.31 (4.21)
4-step	844 (253)	.86/.14 (.10)	5.38 (3.34)
6-step	793 (241)	.95/.05 (.06)	3.56 (4.06)
15-step	770 (258)	.97/.03 (.05)	4.77 (3.37)

Note. SDT RT = average RT (in ms) in the semantic distance task; free recall = average amount of words recollected; 1-step = word pairs directly linked to each other; 2-step = word pairs with two steps between them; 3-step = word pairs with three steps between them; 4-step = word pairs with four steps between them; 6-step = word pairs with six steps between them; 15-step = word pairs with fifteen steps between them.

contrast, a significant reduction in successful recall of cue words is found when distance grew from 2-step to 3-step ($p < .01$). Moreover, no significant reduction in successful recall of cue words was found when distance grew from 3-step to 4-step ($p < .88$), while a significant reduction in successful recall of cue words was found when distance grew from 4-step to 6-step ($p < .02$). Finally, a significant increase in successful recall of cue words was found when distance grew from 6-step to 15-step ($p < .02$).

Discussion

In Experiment 2 we examined the effect of short path length on performance in a semantic relatedness judgment task and free recall from memory. We used word-pairs that were either directly related (1-step), or had short path length of 2-, 3-, and 4-steps. As a control, and based on the findings of Experiment 1, we also added two longer conditions, 6- and 15-steps. The 1-step condition was taken from Experiment 1 and a priori assigned as related word-pairs. The 6- and 15- conditions, based on the findings of Experiment 1, were a priori assigned as unrelated word-pairs. The 2-, 3-, and 4-step conditions were assigned post hoc, based on participant's performance. In this regard, the 2-step (69% related) and 3-step (64% related) conditions were determined as related word-pairs, and the 4-step condition was defined as unrelated word-pairs (86% unrelated).

We found that up to 3-steps word-pairs, as path length grew, RT increased and the percentage of word-pairs judged as related decreased. From 4-steps onward, RT decreased and the word-pairs are dominantly judged as unrelated. Furthermore, we found a significant negative correlation between the path lengths used in Experiment 2 and subjective ratings of semantic strength of these word-pairs. Finally, as path length increased, participants' success in free-recalling the words generally decreased (except when path length increased from 6- to 15-steps). These findings indicate that the breadth of the spreading activation process may be bounded by three steps.

Experiment 3

The results of Experiments 1 and 2 revealed that as path length increases, success in free recall of the words decreased. However, in general the free recall performance of the participants was low. As we analyzed the successful recall of the word-pairs, regardless of whether it was the prime or target word, the highest performance for free recall for the 1-step condition was about 20% (14 words out of 80 of 1-step words). This low performance may be related to the paradigm of the SDT, namely a semantic relatedness judgment task of word pairs with a large number of words (160 in Experiment 1 and 240 in Experiment 2). Free recall paradigms are usually conducted as a single word short list learning phase followed by a test phase (Kahana et al., 2008; Polyn et al., 2009). Thus, in Experiment 3 we aimed to extend the investigation of the effect of path length on memory retrieval with a different memory paradigm, via cued recall. Prior research proposes that successful cued recall is related to the number and strength of preexisting associative links in semantic memory (Nelson, Schreiber, & McEvoy, 1992). In a series of studies, Nelson, Bennett, Gee, Schreiber, and McKinney (1993) demonstrated how cued recall is affected by set size (amount of associative responses to the cue word) and interconnectivity (the extent of interconnections between the associative responses of a cue word). Thus, in the presence of an external prime word, a target word with a small set size and high interconnectivity will be more likely to be successfully recalled (Nelson et al., 1993). However, both set size and interconnectivity in these studies are based on frequency of associative responses (Nelson, McEvoy, & Schreiber, 2004). From a network perspective, interconnectivity may be related to the structure of the network, which shortens path length between a prime and target word. Thus, the higher this interconnectivity, the shorter the path length between concepts that are not directly related in the mental lexicon.

Here we use the same paradigm used in Experiment 1 and 2, but replaced the free-recall with a cued-recall task. In accordance with our hypothesis, we predicted that as path length increases, success in memory retrieval will decrease. Thus, we predicted a similar pattern in a cued recall task, which will replicate our previous free recall findings.

Method

Participants. Thirty-four participants were initially recruited to Experiment 3. Three participants who did not follow the instructions were removed. One participant was removed because of technical issues. One participant was removed because of low accuracy rate in the a priori 1-step condition. Analysis was performed on the remaining 29 participants (23 males, 6 females; mean age 23.4 [$SD = 2.5$]). All participants had normal or corrected to normal vision. Participants either took part in the experiment for partial fulfillment of academic credit or were paid an equivalent of 8 USD for their participation. All participants were dominantly right-handed, with a mean score of 84 ($SD = 15.3$) on the Edinburgh Handedness Inventory (Oldfield, 1971). This experiment was approved by the Bar-Ilan University institutional review board.

Stimuli and tasks.

SDT and distractor tasks. The SDT and distractor tasks used were similar to Experiment 2. In this experiment, the SDT consisted of six conditions, each containing 40 word-pairs – 1-, 2-, 3-, 4-, 6- and 15-step. Based on the findings of Experiment 2, the 1-, 2-, and 3-step condition were a priori classified as related word-pairs and the 4-, 6-, and 15-step conditions were a priori classified as unrelated word-pairs.

Cued recall. A cued recall paradigm was used to examine the relation between associative distance and cued memory retrieval. The cued recall task was based on the task used by Bein et al. (2015). After completion of the distraction task, participants were presented with the prime word from each of the word-pairs and were required to try and recall the matched target word from that word-pair.

Procedure. The procedure of Experiment 3 followed the procedure used in Experiment 2. Participants sat 50 cm from a CRT screen. Both the SDT and distractor tasks were conducted using the E-prime software (Schneider et al., 2002). The cued recall task was conducted using the Presentation software (Neuro- behavioral Systems, U.S.A., http://www.neurobs.com). In each cued recall trial, the prime word appeared on the screen for up to six seconds. During that time, the participant could indicate that he or she recalled the target word by pressing a key. Once the key indicating a successful recall of the target word, the participant entered his or her response by typing it on the computer and pressed another key to move on to the next trial.

Results

SDT trials in which RT was lower than 250 ms were removed. In addition, for each participant, trials which were above or below 2.5 *SD* for each condition were also deleted from final data analysis. The responses generated by the participants in the cued recall task were manually scanned and classified as successful or unsuccessful retrieved responses.

A Distance (1-, 2-, 3-, 4-, 6-, and 15-step) repeated measures ANOVA was conducted to examine the effect of distance on participants (*p*) and item (*i*) mean SDT RT. RT was analyzed for only successful trials based on condition classification (a priori and post-priori), as described above. This analysis revealed a significant main effect of Distance, $F_p(5, 140) = 13.3$, $p < .001, \eta^2 = .32$; $F_i(5, 230) = 20.47$, $p < .001$, $\eta^2 = .304$ (see Table 4). These results replicate the results of Experiment 2, where RT increases until 3-steps and then decreases from 4-step onward. Post hoc analyses (corrected for multiple comparisons) revealed a significant increase in RT from 1- to 2-step, ($p_p < .001$ and $p_i < .001$), and from 2- to 3-step, ($p_p < .01$ and $p_i < .09$). Furthermore, a significant reduction in RT was found when distance grew from 3- to 4-step, ($p_p < .001$ and $p_i < 05$). No significant differences were found when distance grew from 4-step to 6-step and from 6- to 15-step ($p_p < .6$ and $p_i < .57$).

A Distance (1-, 2-, 3-, 4-, 6-, and 15-step) repeated measures ANOVA was conducted to examine the effect of distance on mean successful cued recall of target words. This analysis revealed a significant main effect of Distance, $F(5, 140) = 64.7$, $p < .001$, $\eta^2 = .70$ (see Table 4). Post hoc analyses (corrected for multiple comparisons) revealed that this difference stems from a decline in cued recall performance as path length grows: A significant reduction in successful recall of cue words were found between cued recall performance as distance grew from 1-step to 2-step ($p < .001$) and from 2-step to 3-step ($p < .001$). No significant reduction in successful recall of cue words was found when distance grew from 3-step to 4-step ($p < .61$) while a significant reduction was found when distance grew from 4-step to 6-step ($p < .001$). Finally, no significant reduction in successful recall of cue words was found when distance grew from 6-step to 15-step ($p < .84$).

Table 4
RT in the SDT, % of Word-Pairs Judged as Related/Unrelated, and Percentage of Words Successfully Recalled in Cued Recall as a Function of the Distance in Experiment 3 (SD in Parentheses)

Distance	SDT RT	% un/ related	% CR
1-step	620 (117)	.15/.85 (.11)	.30 (.17)
2-step	739 (179)	.42/.58 (.17)	.25 (.16)
3-step	844 (313)	.27/.73 (.13)	.11 (.10)
4-step	681 (204)	.84/.16 (.09)	.12 (.09)
6-step	654 (179)	.96/.04 (.04)	.04 (.06)
15-step	639 (179)	.96/.04 (.04)	.05 (.06)

Note. SDT RT = average RT (in ms) in the semantic distance task; % un/related = percentage of word-pairs judged as un/related in the SDT; % CR = percentage of successful recognition in the cued recall task. 1-step = word pairs directly linked to each other; 2-step = word pairs with two steps between them; 3-step = word pairs with three steps between them; 4-step = word pairs with four steps between them; 6-step = word pairs with six steps between them; 15-step = word pairs with fifteen steps between them.

Discussion

In Experiment 3 we replicated and generalized our findings on the effect of path length on memory retrieval. This was achieved by conducting a cued-recall, rather than a free-recall, memory retrieval task. Cued-recall is better suited to examine memory recall in word-pairs and has been related to theoretical semantic memory network properties such as set size and interconnectivity (Nelson et al., 1993; Nelson et al., 1992). Similar to Experiment 2, we found a significant effect of path on RT in the SDT: Up to 3 steps separating between word-pairs, participants exhibit increased RT and a decrease in the percentage of word-pairs judged as related. From 4 steps onward, participants exhibit a decrease in RT and judge the majority of word-pairs as unrelated.

The results of the cued-recall task replicate the free-recall results of Experiment 2: As path length increases, the percentage of target words successfully recalled from memory decreases. Experiment 3 also replicates a surprising finding from Experiment 2 regarding the 4-step condition: While RT significantly decreases and the percentage of word-pairs judged as unrelated increases, there is no significant difference in either free- or cued-recall compared with the 3-step condition. This may be related to the spreading activation mechanism and requires further research. Thus, the results of Experiment 3 replicates and extends the findings of Experiment 1 and 2 on the relation between path length and successful retrieval from memory. Our findings from both free- (Experiments 1 and 2) and cued- (Experiment 3) recall from memory illustrate the general effect of path length on memory retrieval.

The results presented so far in Experiments 1–3 demonstrate the effect of long and short path lengths on semantic relatedness judgments and retrieval from memory. Thus, path length can be used as a measure of semantic distance. However, to argue that this method offers an alternative to the conventional LSA approach, a comparison of the ability of both LSA and path length to predict participants' performance in the SDT is required. This was examined in Experiment 4.

Experiment 4

To truly demonstrate how path length can be used as a measure of semantic distance, a comparison with LSA is required. Such a comparison is especially important due to the debate on the use of LSA in studying cognition. On one hand, LSA has been empirically applied to examine semantic priming, memory retrieval and creativity (Beaty et al., 2014; Green, 2016; Howard & Kahana, 2002; Jones & Golonka, 2012). On the other hand, objections have been made that LSA does not truly capture human performance (Hutchison et al., 2008; Recchia & Jones, 2009; Simmons & Estes, 2006), and recent research has found that semantic networks derived from free associations is more valid than semantic networks derived from textual-corpora, LSA approach (De Deyne, Verheyen, et al., 2016).

However, conducting LSA in Hebrew is a computational challenge. The rich morphology inherent in the Hebrew language poses a problem in the case of word-count based models such as LSA. For one, the agglomerative nature of the language means that the same lexeme (a basic semantic unit) can be represented by a large number of strings. For example, the Hebrew language has seven *inseparable prepositions*, which are one-letter prefixes that can be added to words for various purposes. To complicate things further, most of these may also be combined, creating dozens of possible complex prefixes. In addition, many forms of suffixes exist in Hebrew, as well as other complex morphological phenomena. Consequently, in contrast to English, where stemming is generally deemed unnecessary, some sort of leximization is required in Hebrew for the model to be efficient. In Cohen, Ben-Simon, and Levi (2014)—the only work on LSA in Hebrew we are aware of—lexeme representations (produced by a morphological disambiguator) were used instead of strings, stripping the words of prefixes, suffixes, temporal properties and so forth. Here, we followed the same procedure applied by Cohen and colleagues.

We also compared our path length results to another word similarity measure popular in the natural language processing community—positive pointwise mutual information (PPMI; Bullinaria & Levy, 2007; Niwa & Nitta, 1994). This is a variant of pointwise mutual information (PMI), an information-theory based similarity measure (Church & Hanks, 1990; Paperno et al., 2014). PMI has been shown to outperform LSA in predicting human responses in semantic similarity judgments, similar to the SDT (Mandera et al., 2015, 2017; Recchia & Jones, 2009). Its main advantage is being based solely on statistics and does not require any complex algorithms or parameter choices to compute. In PPMI, negative values are transformed to zero, preserving only positive values (Mandera et al., 2017). This variation better accounts for computational measures of word similarity, which are larger or equal to zero (either have some similarity or not).

Here we use a similar approach to the one applied by Cohen et al. (2014) to compute LSA values for the 346 word-pairs used in Experiment 1 and Experiment 2 (after item exclusion). Furthermore, we computed the PPMI scores for these word-pairs. This allows us to compare the performance of our path-length measure of semantic distance with LSA and PPMI values in predicting participants' performance in the SDT.

Method

Data. To compare between LSA, PPMI, and path length measures of semantic distance, we examine how well they correlate with our RT data for all word pairs used in Experiment 1 and 2 (after item exclusion). Thus, we examine how well LSA, PPMI, and path length measures correlate with the SDT RT for 346 word-pairs.

Computing Hebrew LSA scores. LSA comprises several stages. First, a *co-occurrence* matrix is computed from a given corpus for a predefined set of contexts (usually the set of documents or coherent paragraphs comprising the corpus). Then, *weighting* methods are applied to the matrix. Next, a *low rank approximation* is computed from the matrix (for a predefined rank). Finally, the similarity between words (or contexts) is computed from the result.

Co-occurrence matrix. Given a corpus, we divide it to a set of semantic contexts. Common choices are a set of documents which comprise the corpus (e.g., newspaper articles) or simply coherent paragraphs. We then count how many times each word in the lexicon (denoted 'term') appears in each one of the contexts. The result is the *co-occurrence* matrix, where each context is represented by a column and each term is represented by a row.

Weighting. To reflect the fact that rarer terms are more informative, a weighting procedure is applied to each value in the co-occurrence matrix, consisting of *local weighting* and *global weighting*. A popular choice which has been shown to perform well (Landauer et al., 1998) is log-entropy weighting, that is, taking the log of $1 + M_{ij}$ (*local weight*) and dividing by the term's entropy over the contexts (*global weight*):

$$M_{ij} = \frac{\log(1 + M_{ij})}{-\sum_j M_{ij}\log(M_{ij})}$$

Low rank approximation. Using reduced-rank Singular Value Decomposition (SVD), a low rank approximation of the weighted co-occurrence matrix is computed for a predetermined rank *k*. The rationale behind this stage stems from various reasons, which include reducing corpus noise, unification of similar semantic dimensions (transferring overly sparse representation to a denser, compact one), and computation efficiency. There is no known rule for choosing *k*; the choice is generally done empirically, where values in the neighborhood of 300–400 are commonly used. Sometimes an inherent dimension in the task at hand is a natural choice (e.g., a known number of clusters in word clustering).

Term similarity. Once we have a vector representation for each term (the rows of the low rank approximation), we need only to define a method for measuring the distance between these vectors. The most commonly used method in this context is the cosine-similarity (Landauer et al., 1998)—the cosine of the angle between the two vectors, *u* and *v*, computed as:

$$sim(u, v) = \cos(q_{u,v}) = \frac{u^* v}{[\![u]\!] * [\![v]\!]}$$

Using this similarity measure, we can measure the similarity between each pair of terms in our lexicon (as well as between any pair of contexts in our corpus).

Computing Hebrew PPMI scores. For two random variables *x* and *y*, their PPMI score is defined as:

$$PPMI(x, y) = \max\left(\log\frac{P(x, y)}{P(x)P(y)}, 0\right)$$

where in our case, *x* and *y* are two words (or baseforms), $P(x,y)$ is the empirical probability of finding *x* and *y* in the same context, and $P(x)$ and $P(y)$ is the empirical probability of encountering *x* and *y* in a given context.

Corpus. We used a corpus of Hebrew Wikipedia articles supplied to us by the Hebrew Language Project (Cohen & Ben-Simon, 2011) at the National Institute for Testing and Evaluation (NITE; https://hlp.nite.org.il/). This corpus was collected on December 2012 and contains all the existing articles in that time. It consists of a total of 61,102,234 tokens (words) in 1,220,031 paragraphs comprising 138,327 articles. We also used the project's automatic Hebrew morphological disambiguator (Cohen & Ben-Simon, 2011) to extract a baseform (lemma) representation for each word in the texts to construct our semantic space over baseforms instead of strings.

Results

A co-occurrence matrix was computed for the corpus, using Wikipedia articles as contexts and baseforms of all the words in the corpus as the lexicon. Local & global weighting (log-entropy) was then applied to the resulting matrix, and a low-rank decomposition was applied to it for various ranks ($k = 1000, 900 \ldots 100$) using the SVDS function in MATLAB. This was used to compute LSA similarity scores for every pair of words used in our experiments. Based on previous studies, we transformed the LSA scores to measures of semantic distance by subtracting the LSA score from the value of 1 (Beaty et al., 2014; Prabhakaran et al., 2014). Finally, we calculated the PPMI for every pair of words used in our experiments, based on the Hebrew Wikipedia corpus, with Wikipedia articles as contexts and baseforms as terms. We then compared how well both the path length measure, the various LSA measures (full matrix without decomposition and all ranks), and PPMI measures are related to participants' performance in the SDT.

First, we conducted a correlation analysis which included SDT RT, path length, PPMI, and all LSA variables. Based on the behavioral findings, this analysis was conducted separately for short (up to four steps) and long (from four steps) distances. This analysis revealed that path length was the variable with the strongest correlation with RT, for both short, $r(118) = .567$, $p < .001$, and long, $r(228) = -.298$, $p < .01$, distances. PPMI was also significantly correlated, albeit weaker, with RT, for both short, $r(118) = -.349$, $p < .01$, and long, $r(228) = .171$, $p < .01$, distances. In contrast, none of the LSA measures were significantly correlated with RT, for both short and long distances (see Table 5).

Next, we examined how well these measures correlate with the subjective strength judgments for each of the word pairs. We conducted a correlation analysis which included subjective judgments of semantic strength (SJ), path length, PPMI, and all LSA variables (see Table 5). This analysis found a strong significant correlation between SJ and path length $r(346) = -.543$, $p < .001$ (as described above separately for Experiments 1 and 2). This analysis also found a medium significant correlation between SJ and PPMI $r(346) = .392$, $p < .01$. Finally, this analysis also found weak significant correlations between SJ and LSA measures based on rank of 1000–500 (all $ps < .05$).

Lastly, we examined how well these variables predict participants' performance (RT in the SDT), through a linear regression analysis (Table 6). To avoid issues of collinearity, we entered the path length variable and all LSA variables in a stepwise regression. This resulted in a two-step solution for the short distances, where path length is included as a significant predictor for the first step, and PPMI as a significant predictor for the second step. For long distances, the regression analysis resulted in a one-step solution where only path length was included as a significant predictor of RT (Table 6).

Table 5
Correlation Values Between SDT Average RT (Short/long Distance), Subjective Judgment of Word-Pair Strength, Path Length, the Different LSA Measures, and PPMI

Variable	SDT RT		Subjective judgment
	Short distances	Long distances	
Path length	.567***	−.298**	−.543**
PPMI	−.349**	.171**	.392**
LSA measures			
Full	.001	.033	−.002
1000	.060	.091	.124*
900	.055	.089	.124*
800	.057	.090	.120*
700	.049	.110	.110*
600	.050	.111	.108*
500	.052	.119	.108*
400	.030	.100	.099
300	.008	.071	.058
200	−.057	.093	.055
100	−.093	.028	.006

Note. SDT RT = average RT (in ms) in the semantic distance task; Path Length = path length measure; PPMI = positive pointwise mutual information; full = Full LSA matrix without rank decomposition; 1000–100 = different rank decomposition of the LSA matrix.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

In Experiment 4 we examined how path length, PPMI, and LSA measures predict participants' performance in the SDT. This was achieved by computing LSA and PPMI scores based on a corpus of Wikipedia in Hebrew documents (Cohen & Ben-Simon, 2011; Cohen et al., 2014). We show that path length was the strongest predictor of participants RT in the SDT, PPMI was a medium predictor, and LSA did not predict at all participants RT in the SDT. Examining the correlation between the three different computational measures and subjective judgments of semantic strength of the word pairs (SJ), we found a strong correlation between path length and SJ, a medium correlation between PPMI and SJ, and a weak correlation between LSA and SJ. Finally, a regression analysis confirmed that path length was the main predictor of RT in the SDT, while PPMI played a part only in short distances, and LSA did not play a part at all. This demonstrates the strength of path length over LSA and PPMI as a measure of semantic distance (see also De Deyne, Verheyen, et al., 2016).

Despite computing LSA with different ranks, none of the LSA measures predicted Participants RT in the SDT. This could be a result of the relation between the word-pairs being more complex than pure semantic relations. Further, it is possible that the corpus used, Wikipedia in Hebrew, is not a suitable corpus to use for computing LSA. Importantly, this is the first LSA research in Hebrew so we do not have any reference to compare to. Indeed, in our approach, we have followed as closely as possible the approach used to compute LSA in English (Landauer & Dumais, 1997; Landauer et al., 1998). Further research in computational linguistics in Hebrew is needed, such as analyzing alternative corpuses to compute LSA scores from, and comparing cross-linguistic LSA models. However, this only strengthens the problematic bias of the corpus used on computing LSA values (Recchia & Jones, 2009). Finally, we find weak significant positive correlations between some of the ranks used for computing LSA scores and subjective judgments of the semantic strength of the word-pairs. Our results demonstrate that LSA measures do capture semantic information, but only in a weak way. Thus, our findings strengthen previous concerns raised against the use of LSA in

Table 6
Regression Analysis of the Different Semantic Distance Variables on Participants' Performance in the SDT (RT)

Step	Short distances ΔR^2	Short distances β	Long distances ΔR^2	Long distances β
Step 1	.32***		.09***	
Path length		.57***	—	.30***
Step 2	.36***			
Path length		.49***		
PPMI		.21***		

Note. PPMI = positive pointwise mutual information. Step 1: Stepwise; Step 2: Enter.
*** $p < .001$.

studying issues such as semantic priming or semantic relatedness tasks (De Deyne, Verheyen, et al., 2016; Hutchison et al., 2008).

In computing our LSA scores, we followed Cohen et al. (2014)—the only work on LSA in Hebrew we are aware of—as closely as possible. However, it is important to note that the task attempted by Cohen et al. (2014) was modeling corpus documents similarity, which is very different from our task (modeling word similarity), possibly affecting the parameter choices as well. To achieve the optimal results from the LSA procedure, it is necessary to try a very large number of parameter choice combinations. Because LSA requires performing complex operations on very large matrices, parameter choice testing is very costly in terms of computation time. In this work, we tried different values for the rank k (assuming it would have the largest effect on the results). We believe that anything more than that would constitute a full research on Hebrew LSA, which is far beyond the scope of this work.

As an alternative to LSA measures of semantic distance, we also computed positive pointwise mutual information (PPMI) a variant of pointwise mutual information (PMI; Bullinaria & Levy, 2007; Church & Hanks, 1990; Niwa & Nitta, 1994). PMI has been shown to outperform LSA methods in corpora based semantic similarity ratings (Budiu et al., 2007; Bullinaria & Levy, 2007; Terra & Clarke, 2003; Turney, 2001), and have recently been used to examine its performance in predicting behavioral semantic similarity judgments (Mandera et al., 2015, 2017; Recchia & Jones, 2009). Our findings provide further support to PMI outperforming LSA in predicting behavioral performance. However, the path length measure greatly outperformed the PPMI measure.

General Discussion

In the present study we examined whether path length, calculated with network science tools, can be used as a measure of semantic distance. Path length in a semantic network represents the amount of steps needed to traverse from one word in the network to the other. Thus, path length may serve as a measure of semantic distance (Collins & Loftus, 1975). In this study we constructed word pairs which varied in the path length between them, based on a large scale network analysis of the Hebrew lexicon (Kenett et al., 2011). From this analysis, a distance matrix was constructed, which represented the shortest amount of steps connecting any pair of words. This distance matrix was used to operationalize semantic distance, from which the word-pairs were constructed. These word-pairs were used to examine how path length is related to behavioral performance in a semantic relatedness task and to free- and cued-recall from memory. This was examined for both long and short path lengths (Experiments 1–3).

We found a differential effect of path length on behavioral performance. As path length increases from 1 to 3 steps, participants exhibit an increase of RT and a decrease in the percentage of word-pairs judged as related. From a path length of 4 steps, participants exhibit a rapid decrease of RT and dominantly judge the word-pairs as unrelated. We also show a linear effect of path length on free recall from memory. The larger the path length between the word-pairs, the harder it was for participants to recall these words. This effect was further generalized and replicated in a cued-recall task. Furthermore, examining the validity of the semantic distance task, we found a strong significant correlation between path length and subjective judgments of semantic strength of the word-pairs. In this regard, as path length grows, word-pairs are judged to have weaker semantic strength between them. Finally, in Experiment 4, we show how path length as a measure of semantic distance outperforms LSA and PPMI measures in predicting participants' RT in the SDT.

Path Length and Memory Retrieval

Examining growing orders of distance between word pairs can also be used to examine memory retrieval processes, such as those taking place in free- and cued- recall from memory (Kahana et al., 2008; Morton & Polyn, 2016). Research conducted on free recall from memory usually analyzes the order in which participants recall list items from memory (Kahana et al., 2008; Polyn et al., 2009). This analysis provides insights into the search processes operating in this task. Currently, this task is considered as a multiply constrained process, influenced by various factors (Polyn et al., 2009). One main factor extensively investigated is the effect of a semantic factor on free recall performance. These studies usually use free association norms to study the effect of semantic memory structure on the dynamics of free recall (Howard & Kahana, 2002). However, these studies make basic assumptions regarding the role of semantic structure in free recall (Morton & Polyn, 2016; Rao & Howard, 2008; Sirotin, Kimball, & Kahana, 2005). One of these assumptions focus narrowly only on the semantic relations between the specific items studied in the learning phase. Another assumption is concerned with the way the semantic relation between the studied items is determined. Previous studies have compared LSA and Word Association Space vectors (Steyvers et al., 2004) as measures of semantic strength in such computational models of memory retrieval (Morton & Polyn, 2016; Sirotin et al., 2005). These studies have consistently shown how semantic strength measures, based on word association, better accounts for participant's performance in free recall.

Our study contributes to these studies by directly examining the effect of the structure of semantic memory on free recall. We found a negative linear relation between path length and free recall performance. The larger the distance between word-pairs, the harder it was for participants to recall these words. This was regardless of whether the words recalled were the prime or target from the word-pairs. Participants were most successful in recalling word-pairs with distances of 1- and 2-step, and to a lesser extent, of 3-step word-pairs. Notably, although free recall performance decreased from 4-steps onward, participant's RT remained generally constant in judging these word-pairs as unrelated. This indicates dissociation between the effect of path length on semantic relatedness judgment and free recall. Free recall is sensitive to path length even in long distances where semantic relatedness is already indeterminable. Thus, path length can contribute to the research of the dynamical processes taking place during free recall from memory. Finally, our results are more generally related to the effect of semantic memory structure on memory retrieval. Thus, our measure of semantic distance based on path length may improve computational models of memory retrieval (Morton & Polyn, 2016). Furthermore, our finding on the breadth of semantic distance can inform and develop these models.

Recently, Bein et al. (2015) examined the effect of preexisting memory structure on memory retrieval via semantic relatedness and the congruity effect. The congruity effect is the enhanced memory performance for items that are presented within a compatible, rather than incompatible, semantic context (Craik & Tulving, 1975). The authors provide evidence that semantic relatedness and the congruity effect share a common mechanism, which they claim is related to the structure of semantic memory (see also Epstein et al., 1975; Mathews et al., 1981). The results of our research provide support for this view, by relating path length and semantic relatedness judgments.

Finally, we replicated our findings from free-recall to cued-recall memory retrieval. Previous studies have demonstrated how free-recall and cued-recall are attributed to different aspects of memory retrieval (Guzel & Higham, 2013; Higham & Tam, 2005; Tulving & Pearlstone, 1966). Participants better retrieve from memory with cued-recall, performance which improves with list size. The authors argue that this difference demonstrates the availability of the memory trace and how the cue facilitates its accessibility (Tulving & Pearlstone, 1966). Thus, cued-recall is considered to result in stronger memory retrieval than free-recall (Guzel & Higham, 2013). The fact that both free- and cued-recall tasks were similarly affected by path length strengthens the validity of this approach and its generality in regard to memory retrieval from semantic memory.

Path Length as a Measure of Semantic Distance

To the best of our knowledge, this is the first study to examine semantic distance based on path length between words in a semantic network. Currently, the main computational method to derive semantic distance is latent semantic analysis (LSA; Landauer & Dumais, 1997; Landauer et al., 1998). LSA quantifies the semantic similarity between words in a given semantic space and can thus be used to compute semantic distance between words which can be experimentally manipulated to examine semantic priming, creativity and memory retrieval (Beaty et al., 2014; Chwilla & Kolk, 2002; Griffiths et al., 2007; Howard & Kahana, 2002; Hutchison et al., 2008; Pakhomov et al., 2010; Prabhakaran et al., 2014; Steyvers et al., 2004).

However, objections have been raised at using LSA to examine the structure of the mental lexicon (De Deyne, Kenett, et al., 2016; Hutchison et al., 2008; Recchia & Jones, 2009; Simmons & Estes, 2006). For example, Hutchison et al. (2008) have shown that LSA measures fail at predicting semantic priming effects in both a LDT and a naming task at the item level. This, regardless of whether the SOA between prime and target was short or long. The authors point out that research showing the success of LSA in predicting semantic priming have done so only for overall priming effects. Finally, De Deyne, Kenett, et al. (2016) compared a semantic network derived from textual corpora versus a semantic network derived from free association data. These authors show how the textual corpora based network did not predict semantic relatedness judgments as well as the association-based network (De Deyne, Verheyen, et al., 2016). The failure of the LSA measures to successfully predict participants' RT in the SDT, together with the weak correlations with the subjective judgments of semantic strength, strengthens these objections.

As an alternative measure to LSA, we also computed semantic distance scores using positive pointwise mutual information (PPMI), a variant of pointwise mutual information (PMI; Bullinaria & Levy, 2007; Church & Hanks, 1990; Niwa & Nitta, 1994). PMI has been shown to outperform LSA methods in corpora-based semantic similarity ratings (Budiu et al., 2007; Bullinaria & Levy, 2007; Terra & Clarke, 2003; Turney, 2001), and have recently been applied to predict behavioral semantic similarity judgments (Mandera et al., 2015, 2017; Recchia & Jones, 2009). Contrary to LSA, our measure of PPMI had a medium correlation with participants RT performance in the SDT and the SJ of the word-pairs. However, the path length measure strongly outperformed it. Finally, both path length and PPMI contributed as significant predictors of RT in short distances, indicating that they may tap into different components of semantic relations between the word-pairs used in our studies.

The effect of path length on RT in our study demonstrates the feasibility of using path length as a measure of semantic distance. Furthermore, we found a significant correlation between path length and subjective judgment of semantic strength. This correlation strengthens the validity of using path length to measure semantic distance. This correlation, coupled with outperforming the PPMI measure, the SDT RT results, and the inability of the LSA measures to predict participants' performance in the SDT and their weak relations to subjective strength judgments, demonstrate that using path length to examine semantic distance is a valid method, sensitive to intermediate levels of semantic distance, and successful at capturing differences in behavioral performance.

Path Length and Mediated Priming

Even though the notion of semantic distance is fundamental in various cognitive fields, empirical operationalization of it remains an open issue. This may be attributable to alternating approaches to define the principles determining the relations between nodes in semantic memory, may they be semantic, featural, or associative (Hutchison, 2003; Jones & Estes, 2012; Lucas, 2000; McNamara, 2005; McNamara & Altarriba, 1988). Collins and Loftus (1975) defined semantic distance as the "shortest path [direct or indirect] between two nodes" (Collins & Loftus, 1975, p. 412, note 3). Thus, semantic distance can be defined as the number of steps that intervene between the prime and the target in memory. Therefore, the application of network science to model semantic memory and the use of path length as a measure of semantic distance appears a viable approach.

Currently, the main approach to behaviorally examine semantic distance and semantic priming is with mediated priming (Chwilla & Kolk, 2002; Jones, 2012; Jones & Estes, 2012; McNamara, 1992). Mediated priming is usually examined with one mediating concept between prime and target (not presented to the participant), also referred to as 2-step priming, and is based on the spreading activation theory (McNamara, 1992). Some studies have also used 3-step priming, in which two concepts mediate between prime and target (not presented to the participant, Chwilla & Kolk, 2002; McNamara, 1992). However, 3-step priming is much harder to investigate and to the best of our knowledge no studies have examined higher orders of mediated priming. The SDT provides a way to examine *n*-step priming, based on path length (see Siew, 2016 for a similar approach in phonological networks). Our results provide empirical evidence for the existence of 3-step priming and initial results for priming effects of higher-order distances.

An alternative account of mediated priming, and more generally of semantic priming, has been argued on the basis of the compound-cue account (McKoon & Ratcliff, 1992; Ratcliff & McKoon, 1981, 1994). According to this account, items processed by the cognitive system are joined together in short-term memory to form a compound-cue, with some degree of familiarity that is based on the associative links between the items in the compound cue (McKoon & Ratcliff, 1992). Semantic priming according to this theory is based on the high familiarity of the prime word to the compound, which is contingent on the task. According to the compound-cue theory, mediated priming is a result of weak direct associative links between prime and target, and not via mediating concepts (McKoon & Ratcliff, 1992). The main criticism of the compound-cue theory on mediated priming findings—which are based on the spreading activation theory (McNamara, 2005)—is that these findings are established via free association norms as a measure of semantic distance. In this regard, Ratcliff and McKoon (1994) argue that free association probabilities do not accurately predict priming effects. Thus, while spreading activation (based on spread of activation over a semantic network) and compound-cue (based on familiarity of compound-cue in short term memory) accounts offer alternative mechanisms for semantic priming, in both models semantic strength between concepts plays a critical role. However, they mostly diverge on the use of free association probabilities as a measure of semantic strength, or distance (McNamara, 1992, 2005; McNamara & Altarriba, 1988; Ratcliff & McKoon, 1981, 1994). The approach we take here, of measuring semantic distance via path length over semantic network, provides a quantitative measure of semantic distance that may be used to better address the different mechanisms offered by these two models. Applying a quantitative method to represent semantic memory may help differentiate between competing mechanisms operating on semantic memory structure.

Spreading Activation Boundary and Individual Differences

The results presented here indicate that the breadth of the semantic priming spread of activation is 3 steps. From 4 steps between word-pairs, participants dominantly judge them as unrelated (more than 80% for 4-step word pairs and higher afterward). What might be the cause for this three-step boundary in the ability to consider two words as related to each other? A partial answer may arise from the study of navigation in complex networks (Cohen & Havlin, 2003; Kleinberg, 2000). This research has shown that the optimal distance scales to the logarithm of the number of nodes (words) in the network (Kleinberg, 2000), and also in some real-world networks to the logarithm of this logarithm (Cohen & Havlin, 2003). The path lengths used in this study were taken from a large scale network analysis of the Hebrew mental lexicon (Kenett et al., 2011). This study analyzed a network of 800 Hebrew words. According to this navigation principle, based on the size of the analyzed network, the optimal distance navigated should be 2.9 (logarithm of 800), lower than the empirical boundary of three steps.

However, such insights and findings from network science on network navigation can greatly contribute to shed further light on cognitive dynamics of processes operating on semantic memory structure (i.e., relatedness judgment). Further research is needed in larger semantic networks to replicate and validate this boundary of three steps and to examine the cognitive constraints that establish it. Such constraints may be related to research examining constraints in memory capacity, which indicates a capacity limit of approximately three items (Cowan, 2001, 2010). According to spreading activation models (Anderson, 1983; Anderson & Pirolli, 1984; Collins & Loftus, 1975), the spread of activation dissipates quickly over semantic distance (Balota & Lorch, 1986; Den-Heyer & Briand, 1986; Ratcliff & McKoon, 1981). Anderson (Anderson, 1983; Anderson & Pirolli, 1984) has proposed three properties of the spreading activation mechanism: the total activation spreading from a node is less than its own activation; activation decreases exponentially with distance; and activation has an additive effect. Currently, few studies examine algorithmic realizations of search processes over semantic networks (Abbott, Austerweil, & Griffiths, 2015; Borge-Holthoefer & Arenas, 2010a; Capitán et al., 2012; Hills et al., 2015; Kenett & Austerweil, 2016). However, none of these models take into account the rapid decay of activation, theorized by Collins and Loftus (1975). Our results indicate a rapid increase in RT as path length increases until three steps and then a rapid decrease in RT as distance increases. This trend is associated with participant's performance in judging whether word pairs are related to each other or not. These two properties (differential RT and a 3-step boundary) need to be incorporated in any models examining semantic navigation.

Finally, our results indicate relatively high variance in the performance of judging 2- and 3-step word-pairs. This is apparent both in the average RT and percentage of judging these word-pairs as un/related. These findings may indicate that further individual differences are related to how participant's judge word-pair relatedness. Such factors, for example, may be related to a participants' general fluid and retrieval abilities, which have also been related to creative ability (Beaty et al., 2014; Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; Kenett, Beaty, et al., 2016; Nusbaum & Silvia, 2011; Silvia, Beaty, & Nusbaum, 2013). Recently, Faust and Kenett (2014) proposed a novel theory which relates lexicon structure to typical and atypical semantic processing (Faust & Kenett, 2014). According to this theory, different types of lexicon structure may be related to individual differences. In this regard, Kenett, Anaki, and Faust (2014) have recently shown a difference in the organization of semantic memory structure between low and high creative individuals (Kenett et al., 2014). The semantic memory structure of high creative individuals was more flexible than that of low creative individuals. This flexibility was interpreted by the authors as facilitating creative processing, such as connecting between weakly related words in the lexicon (Rossmann & Fink, 2010; Schilling, 2005). These findings provide further support for how different populations exhibit a different semantic memory structure which affects processes operating upon this network.

Limitations

A few possible limitations exist in this study. First, while 1-step word-pairs are a priori considered as directly related, about 10% of these word-pairs were judged as unrelated. This was unexpected, as it was assumed that all of these word-pairs would be judged as related. This might be related to individual differences in how participants judge what defines relatedness. Further research is needed to replicate and more closely examine the nature of this variance. Second, the behavioral results might be affected by the difference in judging the word-pairs as related (yes) versus unrelated (no). However, unlike other semantic tasks, such as the LDT, where there are correct and incorrect responses, there are no clear a priori responses in this task. In this regard, when describing the instructions of the task, it was emphasized to the participant that there was no right or wrong answer and that they were free to choose how they saw fit. No direct description of what was regarded as relatedness was provided, to minimize biasing. Further research is needed replicating our results with alternative methods of judging semantic similarity, such as continuous similarity ratings (Benedek et al., 2017), or choosing a word that is the least similar to two other out of a triplet of words (Connolly, Gleitman, & Thompson-Schill, 2007). Finally, the use of LSA in this work is inherently problematic for a few reasons. The procedure has many parameters with a wide range of possible values, from the choice of corpus and contexts, to the choice of weighting function combinations, the rank *k* to which the matrix is reduced and the choice of similarity measure, each of which may have a drastic effect on the outcome. In addition, distinct properties of the Hebrew language, such as its highly morphological nature, requires additional modifications to the procedure—computing the baseform for each string in the corpus—possibly further affecting the choice of parameter values.

Conclusions

In conclusion, in this work we examined whether path length, as derived from computational analysis of semantic networks, can be used as a measure of semantic distance. Our results prove the feasibility of using path length as a measure of semantic distance, by showing how it affects behavioral performance in a semantic distance task and free- and cued-recall from memory. We validated our approach by showing that it strongly correlates with subjective judgments of associative strength. Finally, we show how our method outperforms LSA and PPMI measures in predicting the behavioral RT data in the SDT and subjective judgments of semantic strength. As such, our approach provides a novel alternative computational method to current methods that derive semantic distance, such as LSA and PPMI. Our results have a more general significance by shedding new light on the breadth of spreading activation and on dynamical processes operating over semantic memory. As the application of network science tools in cognitive systems of language and memory develops, understanding of fundamental cognitive phenomena will grow. Such understanding will facilitate new cognitive theory, and also new cognitive tools to empirically examine such theory.

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. http://dx.doi.org/10.1037/a0038693

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior, 22,* 261–295. http://dx.doi.org/10.1016/S0022-5371(83)90201-3

Anderson, J. R., & Pirolli, P. L. (1984). Spread of activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 791–798. http://dx.doi.org/10.1037/0278-7393.10.4.791

Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 336–345. http://dx.doi.org/10.1037/0278-7393.12.3.336

Balota, D. A., & Paul, S. T. (1996). Summation of activation: Evidence from multiple primes that converge and diverge within semantic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 827–845. http://dx.doi.org/10.1037/0278-7393.22.4.827

Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in cognitive science. *Trends in Cognitive Sciences, 17,* 348–360. http://dx.doi.org/10.1016/j.tics.2013.04.010

Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42,* 1186–1197. http://dx.doi.org/10.3758/s13421-014-0428-8

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE, 6,* e19348. http://dx.doi.org/10.1371/journal.pone.0019348

Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the effect of semantic congruency on episodic memory: The role of integration and relatedness. *PLoS ONE, 10,* e0115624. http://dx.doi.org/10.1371/journal.pone.0115624

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence, 46,* 73–83. http://dx.doi.org/10.1016/j.intell.2014.05.007

Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2017). How semantic memory structure and intelligence contribute to creative thought: A network science approach. Thinking & Reasoning, 1–26. http://dx.doi.org/10.1080/13546783.2016.1278034

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. U. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424,* 175–308. http://dx.doi.org/10.1016/j.physrep.2005.10.009

Borge-Holthoefer, J., & Arenas, A. (2010a). Categorizing words through semantic memory navigation. *The European Physical Journal B, 74,* 265–270. http://dx.doi.org/10.1140/epjb/e2010-00058-9

Borge-Holthoefer, J., & Arenas, A. (2010b). Semantic networks: Structure and dynamics. *Entropy, 12,* 1264–1302. http://dx.doi.org/10.3390/e12051264

Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition, 156,* 60–70. http://dx.doi.org/10.1016/j.cognition.2016.07.014

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of*

Clinical Psychology, 9, 91–121. http://dx.doi.org/10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE, 6,* e27407. http://dx.doi.org/10.1371/journal.pone.0027407

Brunel, N., & Lavigne, F. (2009). Semantic priming in a cortical network model. *Journal of Cognitive Neuroscience, 21,* 2300–2319. http://dx.doi.org/10.1162/jocn.2008.21156

Budiu, R., Royer, C., & Pirolli, P. (2007). Modeling information scent: A comparison of LSA, PMI and GLSA similarity measures on common tests and corpora. *Proceedings of the International Conference on Large Scale Semantic Access to Content (Text, Image, Video, and Sound) (RIAO'07),* 314–332.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods, 39,* 510–526. http://dx.doi.org/10.3758/BF03193020

Capitán, J. A., Borge-Holthoefer, J., Gómez, S., Martinez-Romo, J., Araujo, L., Cuesta, J. A., & Arenas, A. (2012). Local-based semantic navigation on a networked representation of information. *PLoS ONE, 7,* e43694. http://dx.doi.org/10.1371/journal.pone.0043694

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics, 16,* 22–29.

Chwilla, D. J., & Kolk, H. H. J. (2002). Three-step priming in lexical decision. *Memory & Cognition, 30,* 217–225. http://dx.doi.org/10.3758/BF03195282

Cilibrasi, R. L., & Vitanyi, P. M. B. (2007). The Google similarity distance. *IEEE Transactions on Knowledge and Data Engineering, 19,* 370–383. http://dx.doi.org/10.1109/TKDE.2007.48

Coane, J. H., & Balota, D. A. (2011). Face (and nose) priming for book: The malleability of semantic memory. *Experimental Psychology, 58,* 62–70. http://dx.doi.org/10.1027/1618-3169/a000068

Cohen, R., & Havlin, S. (2003). Scale-free networks are ultrasmall. *Physical Review Letters, 90,* 058701. http://dx.doi.org/10.1103/PhysRevLett.90.058701

Cohen, Y., & Ben-Simon, A. (2011). The Hebrew language project: Automated essay scoring & readability analysis, *IAEA Annual Conference*. Vienna, Austria.

Cohen, Y., Ben-Simon, A., & Levi, E. (2014). *Indices of semantic similarity for automated essay scoring (AES)*. Annual meeting of the National Council on Measurement in Education (NCME). Philadelphia, PA.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407–428. http://dx.doi.org/10.1037/0033-295X.82.6.407

Connolly, A. C., Gleitman, L. R., & Thompson-Schill, S. L. (2007). Effect of congenital blindness on the semantic representation of some everyday concepts. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 8241–8246. http://dx.doi.org/10.1073/pnas.0702812104

Cooke, N. M., Durso, F. T., & Schvaneveldt, R. W. (1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 538–549. http://dx.doi.org/10.1037/0278-7393.12.4.538

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. http://dx.doi.org/10.1017/S0140525X01003922

Cowan, N. (2010). The magical mystery four: How Is working memory capacity limited, and why? *Current Directions in Psychological Science, 19,* 51–57. http://dx.doi.org/10.1177/0963721409359277

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268–294. http://dx.doi.org/10.1037/0096-3445.104.3.268

De Deyne, S., Kenett, Y. N., Anaki, D., Faust, M., & Navarro, D. J. (2016). Large-scale network representations of semantics in the mental lexicon. In M. N. Jones (Ed.), *Big data in cognitive science: From methods to insights* (pp. 174–202). New York, NY: Psychology Press/Taylor & Francis.

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. http://dx.doi.org/10.1037/xge0000192

De Deyne, S., Verheyen, S., & Storms, G. (2016). Structure and organization of the mental lexicon: A network approach derived from syntactic dependency relations and word associations. In A. Mehler, P. Blanchard, B. Job, & S. Banish (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 47–79). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-662-47238-5_3

de Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *Journal of Verbal Learning & Verbal Behavior, 22,* 417–436. http://dx.doi.org/10.1016/S0022-5371(83)90273-6

Den-Heyer, K., & Briand, K. (1986). Priming single digit numbers: Automatic spreading activation dissipates as a function of semantic distance. *The American Journal of Psychology, 99,* 315–340. http://dx.doi.org/10.2307/1422488

Epstein, M. L., Phillips, W. D., & Johnson, S. J. (1975). Recall of related and unrelated word pairs as a function of processing level. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 149–152. http://dx.doi.org/10.1037/0278-7393.1.2.149

Faust, M., & Kenett, Y. N. (2014). Rigidity, chaos and integration: Hemispheric interaction and individual differences in metaphor comprehension. *Frontiers in Human Neuroscience, 8,* 511, http://dx.doi.org/10.3389/fnhum.2014.00511

Faust, M., & Lavidor, M. (2003). Semantically convergent and semantically divergent priming in the cerebral hemispheres: Lexical decision and semantic judgment. *Cognitive Brain Research, 17,* 585–597. http://dx.doi.org/10.1016/S0926-6410(03)00172-1

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York, NY: Psychology Press.

Green, A. E. (2016). Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. *Current Directions in Psychological Science, 25,* 28–35. http://dx.doi.org/10.1177/0963721415618485

Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind: Predicting fluency with PageRank. *Psychological Science, 18,* 1069–1076. http://dx.doi.org/10.1111/j.1467-9280.2007.02027.x

Guzel, M. A., & Higham, P. A. (2013). Dissociating early- and late-selection processes in recall: The mixed blessing of categorized study lists. *Memory & Cognition, 41,* 683–697. http://dx.doi.org/10.3758/s13421-012-0292-3

Harris, Z. S. (1970). Distributional structure, *Papers in structural and transformational linguistics* (pp. 775–794). Dordrecht, Holland: D. Reidel Publishing Company. http://dx.doi.org/10.1007/978-94-017-6059-1_36

Higham, P. A., & Tam, H. (2005). Generation failure: Estimating metacognition in cued recall. *Journal of Memory and Language, 52,* 595–617. http://dx.doi.org/10.1016/j.jml.2005.01.015

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science, 20,* 729–739. http://dx.doi.org/10.1111/j.1467-9280.2009.02365.x

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & the Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences, 19,* 46–54. http://dx.doi.org/10.1016/j.tics.2014.10.004

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46,* 85–98. http://dx.doi.org/10.1006/jmla.2001.2798

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: Temporal segregation through synaptic depression. *Cognitive Science, 27,* 403–430. http://dx.doi.org/10.1207/s15516709cog2703_4

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10,* 785–813. http://dx.doi.org/10.3758/BF03196544

Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language, 52,* 1–28. http://dx.doi.org/10.1016/j.jml.2004.08.003

Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *The Quarterly Journal of Experimental Psychology, 61,* 1036–1066. http://dx.doi.org/10.1080/17470210701438111

Jones, L. L. (2010). Pure mediated priming: A retrospective semantic matching model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 135–146. http://dx.doi.org/10.1037/a0017517

Jones, L. L. (2012). Prospective and retrospective processing in associative mediated priming. *Journal of Memory and Language, 66,* 52–67. http://dx.doi.org/10.1016/j.jml.2011.08.005

Jones, L. L., & Estes, Z. (2012). Lexical priming: Associative, semantic, and thematic influences on word recognition. In J. S. Adelman (Ed.), *Visual word recognition, Vol. 2: Meaning and context: Individuals and development* (pp. 44–72). Hove, UK: Psychology Press.

Jones, L. L., & Golonka, S. (2012). Different influences on lexical priming for integrative, thematic, and taxonomic relations. *Frontiers in Human Neuroscience, 6,* 205. http://dx.doi.org/10.3389/fnhum.2012.00205

Jones, M. N., Willits, J., & Dennis, S. (2015). Models of semantic memory. In J. Busemeyer & J. Townsend (Eds.), *Oxford handbook of mathematical and computational psychology* (pp. 232–254). Oxford, UK: Oxford University Press.

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In H. L. Roediger, III, (Ed.), *Cognitive psychology of memory* (pp. 467–490). Oxford, UK: Elsevier. http://dx.doi.org/10.1016/B978-012370509-9.00185-6

Karuza, E. A., Thompson-Schill, S. L., & Bassett, D. S. (2016). Local patterns to global architectures: Influences of network topology on human learning. *Trends in Cognitive Sciences, 20,* 629–640. http://dx.doi.org/10.1016/j.tics.2016.06.003

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. http://dx.doi.org/10.3389/fnhum.2014.00407

Kenett, Y. N., Anaki, D., & Faust, M. (2015). Processing of unconventional stimuli requires the recruitment of the non-specialized hemisphere. *Frontiers in Human Neuroscience, 9,* 32. http://dx.doi.org/10.3389/fnhum.2015.00032

Kenett, Y. N., & Austerweil, J. L. (2016). Examining search processes in low and high creative individuals with random walks. In A. Papafragou & D. Grodner & D. Mirman & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 313–318). Austin, TX: Cognitive Science Society.

Kenett, Y. N., Beaty, R. E., Silvia, P. J., Anaki, D., & Faust, M. (2016). Structure and flexibility: Investigating the relation between the structure of the mental lexicon, fluid intelligence, and creative achievement. *Psychology of Aesthetics, Creativity, and the Arts, 10,* 377–388. http://dx.doi.org/10.1037/aca0000056

Kenett, Y. N., Gold, R., & Faust, M. (2016). The hyper-modular associative mind: A computational analysis of associative responses of persons with Asperger Syndrome. *Language and Speech, 59,* 297–317. http://dx.doi.org/10.1177/0023830915589397

Kenett, Y. N., Kenett, D. Y., Ben-Jacob, E., & Faust, M. (2011). Global and local features of semantic networks: Evidence from the Hebrew mental lexicon. *PLoS ONE, 6,* e23912. http://dx.doi.org/10.1371/journal.pone.0023912

Kenett, Y. N., Wechsler-Kashi, D., Kenett, D. Y., Schwartz, R. G., Ben-Jacob, E., & Faust, M. (2013). Semantic organization in children with cochlear implants: Computational analysis of verbal fluency. *Frontiers in Psychology, 4,* 543. http://dx.doi.org/10.3389/fpsyg.2013.00543

Kleinberg, J. M. (2000). Navigation in a small world. *Nature, 406,* 845. http://dx.doi.org/10.1038/35022643

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. http://dx.doi.org/10.1037/0033-295X.104.2.211

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284. http://dx.doi.org/10.1080/01638539809545028

Lavigne, F., Dumercy, L., & Darmon, N. (2011). Determinants of multiple semantic priming: A meta-analysis and spike frequency adaptive model of a cortical network. *Journal of Cognitive Neuroscience, 23,* 1447–1474. http://dx.doi.org/10.1162/jocn.2010.21504

Lerner, I., Bentin, S., & Shriki, O. (2012). Spreading activation in an attractor network with latching dynamics: Automatic semantic priming revisited. *Cognitive Science, 36,* 1339–1382. http://dx.doi.org/10.1111/cogs.12007

Lerner, I., Bentin, S., & Shriki, O. (2014). Integrating the automatic and the controlled: Strategies in semantic priming in an attractor network with latching dynamics. *Cognitive Science, 38,* 1562–1603. http://dx.doi.org/10.1111/cogs.12133

Lerner, I., & Shriki, O. (2014). Internally- and externally-driven network transitions as a basis for automatic and strategic processes in semantic priming: Theory and experimental validation. *Frontiers in Psychology, 5,* 314. http://dx.doi.org/10.3389/fpsyg.2014.00314

Lorch, R. F. (1982). Priming and search processes in semantic memory: A test of three models of spreading activation. *Journal of Verbal Learning & Verbal Behavior, 21,* 468–492. http://dx.doi.org/10.1016/S0022-5371(82)90736-8

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618–630. http://dx.doi.org/10.3758/BF03212999

Mandera, P., Keuleers, E., & Brysbaert, M. (2015). How useful are corpus-based methods for extrapolating psycholinguistic variables? *The Quarterly Journal of Experimental Psychology, 68,* 1623–1642. http://dx.doi.org/10.1080/17470218.2014.988735

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language, 92,* 57–78. http://dx.doi.org/10.1016/j.jml.2016.04.001

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 3–23. http://dx.doi.org/10.1037/0278-7393.21.1.3

Mathews, R. C., Maples, R. C., & Elkins, R. (1981). Semantic judgments as encoding operations in recall: The encoding of task-relevant and task-Irrelevant semantic attributes of words. *Journal of General Psychology, 105,* 311–320. http://dx.doi.org/10.1080/00221309.1981.9921084

McKoon, G., & Ratcliff, R. (1992). Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1155–1172. http://dx.doi.org/10.1037/0278-7393.18.6.1155

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1173–1190. http://dx.doi.org/10.1037/0278-7393.18.6.1173

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York, NY: Psychology Press. http://dx.doi.org/10.4324/9780203338001

McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language, 27,* 545–559. http://dx.doi.org/10.1016/0749-596X(88)90025-3

McRae, K., & Jones, M. N. (2013). Semantic memory. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (pp. 206–219). Oxford, UK: Oxford University Press.

McRae, K., Khalkahli, S., & Hare, M. (2012). Semantic and associative relations: Examining a tenuous dichotomy. In V. F. Reyna, S. B. Chapman, M. R. Dougherty, & J. Confrey (Eds.), *The adolescent brain: Learning, reasoning, and decision making* (pp. 39–66). Washington, DC: APA. http://dx.doi.org/10.1037/13493-002

Meade, M. L., Watson, J. M., Balota, D. A., & Roediger, H. L., III. (2007). The role of spreading actviation and retrieval mode in producing false recognition in the DRM paradigm. *Journal of Memory and Language, 56,* 305–320. http://dx.doi.org/10.1016/j.jml.2006.07.007

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234. http://dx.doi.org/10.1037/h0031564

Morton, N. W., & Polyn, S. M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language, 86,* 119–140. http://dx.doi.org/10.1016/j.jml.2015.10.002

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. A., & McKinney, V. M. (1993). Implicit memory: Effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 747–764. http://dx.doi.org/10.1037/0278-7393.19.4.747

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36,* 402–407. http://dx.doi.org/10.3758/BF03195588

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review, 99,* 322–348. http://dx.doi.org/10.1037/0033-295X.99.2.322

Niwa, Y., & Nitta, Y. (1994). *Co-occurrence vectors from corpora vs. distance vectors from dictionaries*. *Proceedings of the 17th International Conference on computational Linguistics, COLING'94, 1*, 304–309.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence, 39,* 36–45. http://dx.doi.org/10.1016/j.intell.2010.11.002

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113. http://dx.doi.org/10.1016/0028-3932(71)90067-4

Pakhomov, S. V. S., McInnes, B., Adam, T., Liu, Y., Pedersen, T., & Melton, G. B. (2010). *Semantic similarity and relatedness between clinical terms: An experimental study*. *Proceedings of the Annual Symposium of the American Medical Informatics Association, 2010*, 572–576.

Paperno, D., Marelli, M., Tentori, K., & Baroni, M. (2014). Corpus-based estimates of word association predict biases in judgment of word co-occurrence likelihood. *Cognitive Psychology, 74,* 66–83. http://dx.doi.org/10.1016/j.cogpsych.2014.07.001

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

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review, 116,* 129–156. http://dx.doi.org/10.1037/a0014420

Prabhakaran, R., Green, A. E., & Gray, J. R. (2014). Thin slices of creativity: Using single-word utterances to assess creative cognition. *Behavior Research Methods, 46,* 641–659. http://dx.doi.org/10.3758/s13428-013-0401-7

Rao, V. A., & Howard, M. W. (2008). Retrieved context and the discovery of semantic structure. *Advances in Neural Information Processing Systems, 20,* 1193–1200.

Ratcliff, R., & McKoon, G. (1981). Does activation really spread? *Psychological Review, 88,* 454–462. http://dx.doi.org/10.1037/0033-295X.88.5.454

Ratcliff, R., & McKoon, G. (1994). Retrieving information from memory: Spreading-activation theories versus compound-cue theories. *Psychological Review, 101,* 177–184. http://dx.doi.org/10.1037/0033-295X.101.1.177

Recchia, G., & Jones, M. N. (2009). More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods, 41,* 647–656. http://dx.doi.org/10.3758/BRM.41.3.647

Rossmann, E., & Fink, A. (2010). Do creative people use shorter association pathways? *Personality and Individual Differences, 49,* 891–895. http://dx.doi.org/10.1016/j.paid.2010.07.025

Schilling, M. A. (2005). A "small-world" network model of cognitive insight. *Creativity Research Journal, 17*(2–3), 131–154. http://dx.doi.org/10.1080/10400419.2005.9651475

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

Schvaneveldt, R. W., Dearholt, D. W., & Durso, F. T. (1988). Graph theoretic foundations of pathfinder networks. *Computers & Mathematics with Applications, 15,* 337–345. http://dx.doi.org/10.1016/0898-1221(88)90221-0

Siew, C. S. Q. (2016). The influence of 2-hop network density on spoken word recognition. *Psychonomic Bulletin & Review,* 1–7. http://dx.doi.org/10.3758/s13423-016-1103-9

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence, 41,* 328–340. http://dx.doi.org/10.1016/j.intell.2013.05.004

Simmons, S., & Estes, Z. (2006). *Using latent semantic analysis to estimate similarity*. *Proceedings of the 28th Cognitive Science Society* (pp. 2169–2173). Austin, TX: Cognitive Science Society.

Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review, 12,* 787–805. http://dx.doi.org/10.3758/BF03196773

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications: Festchrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer* (pp. 237–249). Washington, DC: American Psychological Association.

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. http://dx.doi.org/10.1207/s15516709cog2901_3

Terra, E., & Clarke, C. L. A. (2003). *Frequency estimates for statistical word similarity measures*. Paper presented at the Proceedings of the 2003 Conference of the North American chap. of the Association for Computational Linguistics on Human Language Technology-Vol. 1.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language, 38,* 440–458. http://dx.doi.org/10.1006/jmla.1997.2559

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning & Verbal Behavior, 5,* 381–391. http://dx.doi.org/10.1016/S0022-5371(66)80048-8

Turney, P. (2001). Mining the Web for Synonyms: PMI-IR Versus LSA on TOEFL. In *Proceedings of the Twelfth European Conference on Machine Learning* (pp. 491–502). Berlin, Germany: Springer-Verlag.

Vitevitch, M. S., Chan, K. Y., & Goldstein, R. (2014). Insights into failed lexical retrieval from network science. *Cognitive Psychology, 68,* 1–32. http://dx.doi.org/10.1016/j.cogpsych.2013.10.002

Vitevitch, M. S., Chan, K. Y., & Roodenrys, S. (2012). Complex network structure influences processing in long-term and short-term memory. *Journal of Memory and Language, 67,* 30–44. http://dx.doi.org/10.1016/j.jml.2012.02.008

Vitevitch, M. S., Goldstein, R., & Johnson, E. (2016). Path-length and the misperception of speech: Insights from network science and psycholinguistics. In A. Mehler, A. Lücking, S. Banisch, P. Blanchard, & B. Job (Eds.), *Towards a theoretical framework for analyzing complex linguistic networks* (pp. 29–45). Berlin, Germany: Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-662-47238-5_2

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393,* 440–442. http://dx.doi.org/10.1038/30918

Received May 20, 2015
Revision received December 13, 2016
Accepted December 28, 2016 ■