Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Behavioral Sciences

Distributed semantic representations for modeling human judgment

Sudeep Bhatia, Russell Richie and Wanling Zou

People make judgments about thousands of different objects and concepts on a day-to-day basis; however, capturing the knowledge that subserves these judgments has been difficult. Recent advances in computational linguistics are filling this gap, as the statistics of language use yield rich, distributed semantic representations for natural objects and concepts. These representations have been shown to predict semantic and linguistic judgments, such as judgments of meaning and relatedness, and more recently, high-level judgments, including probability judgment and forecasting, stereotyping and various types of social judgment, consumer choice, and perceptions of risk. Distributed semantic representations are now a key component of computational models that represent knowledge, make evaluations and attributions, and give responses, in a human-like manner.

Address
University of Pennsylvania, United States

Corresponding author: Bhatia, Sudeep (bhatiasu@sas.upenn.edu)

Current Opinion in Behavioral Sciences 2019, **29**:31–36

This review comes from a themed issue on **Artificial intelligence**

Edited by **Matt Botvinick** and **Sam Gershman**

https://doi.org/10.1016/j.cobeha.2019.01.020

Representation and judgment

Understanding how people judge, evaluate, and make attributions about objects and concepts is a fundamental goal across the behavioral sciences. Central to this research is the analysis of naturalistic targets of judgment — people, places, and things in the world. Artificial experimental stimuli may be useful to tease apart psychological and neural processes, but many behavioral scientists aim to be able to predict the types of everyday judgments that influence the lives of individuals, and the functioning of societies and economies.

This real world focus has imposed a set of unique challenges. What people know about and associate with natural objects and concepts is not directly accessible to researchers, making it difficult to study the many important types of judgments that people make on a day-to-day basis. How can we uncover *semantic representations* for targets of everyday judgment, such as household objects, food items, consumer products, or political candidates? How can we use these uncovered representations to predict people's judgments, evaluations, and attributions for these targets, and to characterize the complex web of associations that underlie responses and behavior?

Psychologists have developed a number of techniques for uncovering and quantifying representations for natural entities. However, until very recently, these have all involved explicit participant responses obtained through surveys and experiments. Thus, for example, techniques relying on multi-dimensional scaling [e.g. Ref. 1] elicit pairwise similarity ratings from participants, and, through dimensionality reduction, uncover a small number of latent dimensions of semantic representation for the rated items (corresponding perhaps to some intuitive features of the items). Other techniques simply ask participants to list the features that they think are important for each item [e.g. Ref. 2]. The need for explicit participant responses and the difficulty of implementing these methods on a large scale mean that they can only be used to generate relatively impoverished representations for a small number of items.

Fortunately, the past few years have seen the development of a new set of powerful techniques for uncovering and quantifying semantic representations. These techniques, and their underlying theoretical assumptions, are referred to by the term *distributional semantics* (DS), as they propose that semantic representations are reflected in, and can be recovered from, the statistical distribution of words in language. With the recent availability of large scale natural language data, and increased computational power, we are now able to get especially rich and comprehensive distributed semantic representations for millions of objects and concepts (without the need for any explicit participant ratings data). This has opened up new avenues for studying naturalistic human judgment, and made it feasible to build computational models that represent knowledge, make evaluations and attributions, and give responses, in a human-like manner. This paper will provide a summary of these models, with a focus on their applications to human judgment and behavioral science. Other recent reviews provide a more comprehensive overview of the statistical and computational underpinnings of DS models, the differences between different algorithms for building DS models, and the technical steps necessary to apply these algorithms on natural language data [e.g. Refs. 3,4].

Distributional semantics

The core idea underlying distributional semantics (DS) is that word meaning and knowledge are reflected in the distribution of words in language, with closely associated words occurring in similar linguistic contexts. For instance, the words 'big' and 'large' rarely co-occur immediately, but are used interchangeably to refer to similar sizes of objects. These co-occurrences allow distributional semantic models to discern that the two words are related to each other. Recent applications of this idea use word-context co-occurrence relationships to derive high-dimensional vector representations for words, so that objects and concepts with similar word vectors in the underlying semantic space, are predicted to be associated with each other. In Figure 1, we present a hypothetical three-dimensional semantic space and show how this space can be used to measure the association between various objects and concepts.

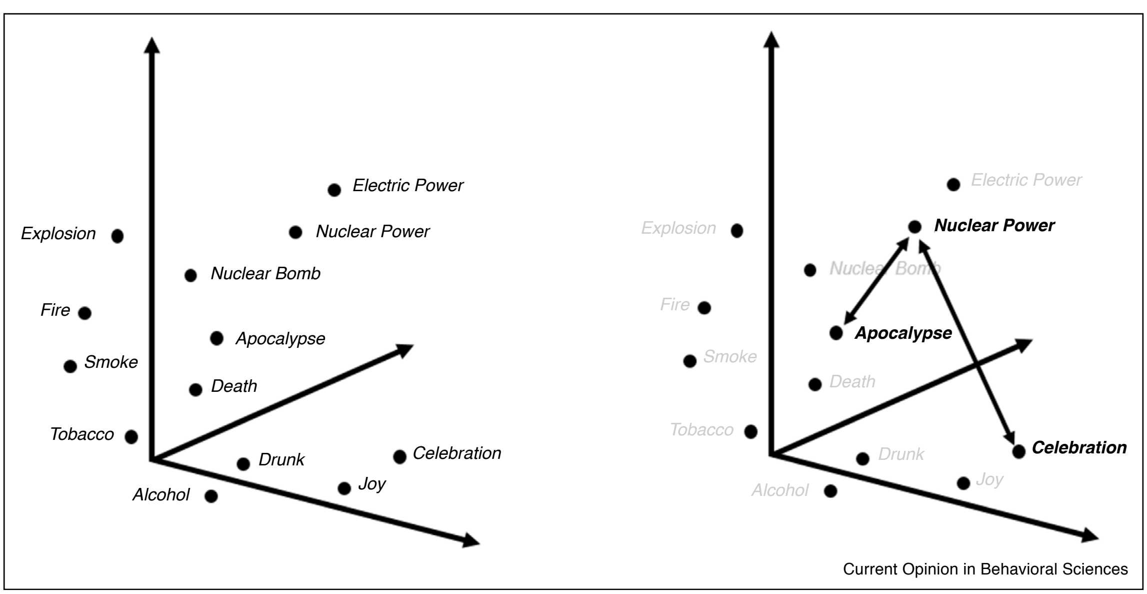
There are, by now, many different algorithms for building semantic spaces [e.g. Refs. 5–9]. Some of these algorithms measure word-context co-occurrence, and using various unsupervised techniques, derive word vectors to preserve measured co-occurrence relationships. Others attempt to predict word-context co-occurrence and thus derive word vectors in a seemingly supervised manner. Understanding the core differences between these two classes of methods (and determining whether or not they are actually different) has been a key focus of recent research in computational linguistics and cognitive science [e.g. Refs. 10,11•].

One aspect of DS that has made it especially valuable to researchers across different fields is the wide availability of off-the-shelf vector representations. These representations have been trained on very large corpora using sophisticated computational methods, and have rich representations for millions of words and phrases. However, they can be downloaded and applied with ease by non-experts to novel domains [12•, see also Refs. 13 and 14 for more recent examples of pretrained DS models with powerful 'transfer learning' applications].

Semantic judgment

DS models have a number of practical applications in natural language processing and artificial intelligence. However, this approach is also based on foundational ideas in linguistics [15] as well as core theories in cognitive science regarding the nature of semantic memory and learning from episodic experience [4,11•]. This makes DS a theoretically grounded tool for predicting human judgment. Indeed, much of the research on DS, both in cognitive science and in computer science, has focused on building models that are capable of mimicking human responses in synonym judgment tasks [e.g. Refs. 5,16], similarity judgment tasks [e.g. Ref. 17], and tasks

Figure 1



Twelve concepts in a hypothetical three dimensional semantic space (left). Such spaces are derived from the distribution of the words in natural language data. Proximity on this space captures the association between concepts (right): Here *Nuclear Power* is more associated with *Apocalypse* than with *Celebration*. This can be used to predict semantic and linguistic judgments for these concepts (e.g. concepts that are judged to be related to *Nuclear Power*); memory effects (e.g. the concepts that are activated when the individual is primed with *Nuclear Power*), as well as high-level judgments (e.g. peoples' perceptions of the risks of *Nuclear Power*).

involving judgments of semantic relatedness and association [18,19]. The rich semantic knowledge possessed by DS representations also makes them useful for modeling other phenomena related to semantic memory, and recent work has used this approach to study priming and lexical access [11•,20–26], as well as free association, semantic search, and list recall [6,11•,12•,27–31]. Even analogical judgments can be predicted by vector arithmetic (e.g. $king - man + woman \sim queen$) [7,8], suggesting that DS representations are capable of representing relational similarity.

Although models of distributional semantics do appear to encode realistic semantic representations, there are settings in which their predictions deviate from the judgments made by human participants, and understanding how to address these deviations is an important topic of continuing inquiry. One such setting pertains to asymmetries in semantic representation. Human similarity judgment displays systematic asymmetries (with object A being considered similar to object B, but object B not being considered similar to object A, but see Ref. [32] for a large-scale failure to replicate asymmetries in similarity judgments), and often violates the triangle inequality (with the judged dissimilarity of object A and object C being greater than the judged dissimilarity of object A and object B plus the judged dissimilarity of object B and object C). Simple geometric measures of distance in semantic spaces, which are often used to predict semantic judgment (e.g. Figure 1), are unable to easily account for these patterns [33–35]. Many DS models also fail to accurately predict human responses in similarity judgment tasks which control for the effect of association [17], when similarity judgments are applied to verbs rather than nouns [36], and when similarity judgments involve antonyms [37] and hypernyms [38]. It seems that most standard DS models are best at capturing association (or, more generally, semantic relatedness), rather than similarity.

High-level judgment

Associations are considered to be at play in a variety of high-level human judgments, making DS models applicable across many areas of psychology and cognitive science, including those that are not explicitly semantic or linguistic. For example, association-based assessments of representativeness (that is, similarity in representation for judgment-related concepts) have been shown to play a key role in probability judgment, largely by guiding automatic 'intuitive' responses. According to this insight, a recent work has used distances in semantic spaces (as in e.g. Figure 1) to predict associations, and in turn, participant probability assignments for naturalistic events [39••]. Similar methods have also been used to predict factual judgment, both in the lab and in the real-world [39••,40]. Interestingly, associations derived from DS models predict both when participants are likely to give correct judgments and when they are likely to make mistakes (such as when the most associated response is incorrect).

Another area in which associative processing plays a fundamental role is social judgment, and DS models have recently been shown to encode the types of associative biases responsible for social stereotypes and prejudices [41••,42–44]. Interestingly, this work finds the parameters of the semantic models that are best at representing social information, are also the ones that lead to the most detrimental social biases [44]. As DS models are used for a variety of artificial intelligence applications, researchers have also been developing techniques to debias learnt representations [43].

The fact that DS representations encode human-like associations has one intriguing implication: by training DS representations on different corpora we can infer the associations that would be possessed by groups of people differentially exposed to – or responsible for producing – those corpora, and by doing so, infer differences in high-level judgments across these groups. Drawing on this insight, recent applications of DS models have attempted to study differences in social, political, and moral associations pertaining to media bias and social media structure [45–51,52•]. Other work has trained DS models on historical language data, to study changing gender, class, and ethnic associations over time [53••,54,55]. It is important to note that much of this analysis would not be possible using standard techniques in the behavioral sciences, that rely on explicit human responses (e.g. it is impossible to go back in time to survey gender associations in the early 20th century).

A final application of DS models to human judgment involves the analysis of complex judgments that are not necessarily associative. These types of judgments cannot easily be predicted using measurements of distance in semantic space, as in Figure 1. However, as these judgments nonetheless rely on semantic representations, it is possible to learn a mapping from the high-dimensional semantic space to the judgment domain in consideration. With training data, this approach is able to generate highly accurate predictions for different types of complex judgments, including judgments of desirability for multi-attribute choice items [56] and perceptions of risk for naturalistic hazards [57•].

Related and future work

Although we have focused on human judgment, DS models have important applications in many areas of cognitive and behavioral science. As in the DS studies of complex judgments just mentioned [57•], psycholinguistic properties of words, including age of acquisition, concreteness, and valence, have been predicted from DS models with a high degree of accuracy by regressing psycholinguistic ratings directly on word vectors [58–61]. Word vectors have even been used to uncover the neural maps of semantic representations by linking word vectors to brain signal elicited by those words [62–64].

While, as described above, DS models have been quite successful at modeling cognition and behavior, there are (at least) three major directions of future research that will be relevant to human judgment. The first is finding ways to use distributional semantics to represent the meaning of not just single words or frozen phrases, but novel multiword phrases, sentences, paragraphs, and documents. This is the challenge of so-called 'compositional distributional semantics' [65–68, also see Refs. 13 and 14, which offer powerful pre-trained DS representations for multiword texts]. While effective and scalable methods for delivering representations at the phrase level and above are an active area of research in computational linguistics, such representations will also be very useful for modeling human evaluations and assessments of complex propositions, and eventually, human reasoning and problem solving.

The second avenue of research is combining distributed word representations with perceptual data. In one approach to this problem, perceptual/visual information is collected in parallel to corpus-based word representations, and then both pieces of information are independently used to predict and explain cognitive and behavioral phenomena [69–71]. Applying these techniques to the study of human judgment will be critical, as knowledge of judgment entities is obtained not just through linguistic experience, but also through direct experience with these entities in the world.

It is also important to note that there are other approaches to building models of semantic representation, including approaches that use other types of data (e.g. word association norms elicited from experimental participants or links between Wikipedia articles, instead of natural language), and approaches that rely on alternate mathematical structures (e.g. networks instead of spaces) [72–74]. Different approaches have different strengths, and the best models of knowledge and association in human judgment are likely to be those that rely on multiple sources of data and permit multiple types of representations.

Conclusion

In summary, distributional semantics (DS) offers a powerful new technique for uncovering the knowledge representations at play in human cognition and behavior. These representations can be used to predict semantic and linguistic judgment (e.g. judgments of relatedness, and in some cases, judgments of similarity and analogy), as well as high-level judgment (e.g. judgments of probabilities, forecasts for events, social associations, stereotypes and prejudice, consumer choice, and risk perception). Other applications of DS include measuring the psychological properties of text, estimating historical associations and judgments, estimating psycholinguistic properties of words, and specifying the neural underpinnings of semantic cognition. The near future will see the development of richer DS representations that are obtained from perceptual and lexical data sources, and are capable of representing complex words and phrases. These representations will in turn facilitate the development of powerful, accurate, and domain-general computational models of cognition and behavior, capable of making human-like judgments for a wide array of naturalistic judgment targets.

Funding

Funding was received from the National Science Foundation grant SES-1626825, and from the Alfred P. Sloan Foundation.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Shepard RN: **Representation of structure in similarity data: problems and prospects**. *Psychometrika* 1974, **39**:373-421.
2. McRae K, Cree GS, Seidenberg MS, McNorgan C: **Semantic feature production norms for a large set of living and nonliving things**. *Behav Res Methods* 2005, **37**:547-559.
3. Lenci A: **Distributional models of word meaning**. *Ann Rev Linguist* 2018, **4**:151-171.
4. Jones MN, Willits JA, Dennis S: **Models of semantic memory**. In *Oxford Handbook of Computational and Mathematical Psychology.* Edited by Busemeyer JR, Wang Z, Townsend JT, Eidels A. Oxford University Press; 2015:232-254.
5. Landauer TK, Dumais ST: **A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge**. *Psychol Rev* 1997, **104**:211-240.
6. Griffiths TL, Steyvers M, Tenenbaum JB: **Topics in semantic representation**. *Psychol Rev* 2007, **114**:211-244.
7. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J: **Distributed representations of words and phrases and their compositionality**. *Adv Neural Inf Process Syst* 2013:3111-3119.
8. Pennington J, Socher R, Manning CD: **Glove: global vectors for word representation**. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. 2014:1532-1543.
9. Jones MN, Mewhort DJ: **Representing word meaning and order information in a composite holographic lexicon**. *Psychol Rev* 2007, **114**:1-37.
10. Levy O, Goldberg Y, Dagan I: **Improving distributional similarity with lessons learned from word embeddings**. *Trans Assoc Comput Linguist* 2015, **3**:211-225.
11. • Mandera P, Keuleers E, Brysbaert M: **Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: a review and empirical validation**. *J Mem Lang* 2017, **92**:57-78.

The authors evaluate various DS models on key psycholinguistic tasks. They also discuss how the assumptions underlying DS models relate to psychologically plausible learning principles.

12. • Pereira F, Gershman S, Ritter S, Botvinick M: **A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data**. *Cogn Neuropsychol* 2016, **33**:175-190.
The authors provide a comparison of various pretrained distributional semantic models in terms of predictive performance on various semantic judgment and association tasks.

13. Devlin J, Chang MW, Lee K, Toutanova K: *Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding*. 2018. arXiv preprint arXiv:1810.04805.

14. Peters ME, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, Zettlemoyer L: *Deep Contextualized Word Representations*. 2018. arXiv preprint arXiv:1802.05365.

15. Harris ZS: **Distributional structure**. *Word* 1954, **2**:146-162.

16. Levy J, Bullinaria J, McCormick S: **Semantic vector evaluation and human performance on a new vocabulary MCQ test**. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. 2017:2549-2554.

17. Hill F, Reichart R, Korhonen A: **Simlex-999: evaluating semantic models with (genuine) similarity estimation**. *Comput Linguist* 2015, **41**:665-695.

18. Bruni E, Tran NK, Baroni M: **Multimodal distributional semantics**. *J Artif Intell Res* 2014, **49**:1-47.

19. Hofmann MJ, Biemann C, Westbury C, Murusidze M, Conrad M, Jacobs AM: **Simple Co-occurrence statistics reproducibly predict association ratings**. *Cogn Sci* 2018, **42**:2287-2312.

20. Jones MN, Kintsch W, Mewhort DJK: **High-dimensional semantic space accounts of priming**. *J Mem Lang* 2006, **55**: 534-552.

21. Johns BT, Dye M, Jones MN: **The influence of contextual diversity on word learning**. *Psychon Bull Rev* 2016, **23**: 1214-1220.

22. Johns BT, Sheppard CL, Jones MN, Taler V: **The role of semantic diversity in word recognition across aging and bilingualism**. *Front Psychol* 2016, **7**:703-714.

23. Auguste J, Rey A, Favre B: **Evaluation of word embeddings against cognitive processes: primed reaction times in lexical decision and naming tasks**. In *Proceedings of the 2nd Workshop on Evaluating Vector Space Representations for NLP*. 2017:21-26.

24. Ettinger A, Linzen T: **Evaluating vector space models using human semantic priming results**. In *Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP*. 2016:72-77.

25. Günther F, Dudschig C, Kaup B: **Latent semantic analysis cosines as a cognitive similarity measure: evidence from priming studies**. *Q J Exp Psychol* 2016, **69**:626-653.

26. Marelli M, Amenta S: **A database of orthography-semantics consistency (OSC) estimates for 15,017 English words**. *Behav Res Methods* 2018, **50**:1482-0171495.

27. Steyvers M, Shiffrin RM, Nelson DL: **Word association spaces for predicting semantic similarity effects in episodic memory**. In *Experimental Cognitive Psychology and Its Applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer.* Edited by Healy AF. 2004:237-249.

28. Bhatia S: **Semantic processes in preferential decision making**. *J Exp Psychol Learn Mem Cogn* 2018. [in press].

29. Hills TT, Jones MN, Todd PM: **Optimal foraging in semantic memory**. *Psychol Rev* 2012, **119**:431-440.

30. Johns BT, Taler V, Pisoni DB, Farlow MR, Hake AM, Kareken DA, Unverzagt FW, Jones MN: **Cognitive modeling as an interface between brain and behavior: measuring the semantic decline in mild cognitive impairment**. *Can J Exp Psychol* 2018, **72**:117-126.

31. Healey MK, Kahana MJ: **A four-component model of age-related memory change**. *Psychol Rev* 2016, **123**:23-69.

32. Klein RA, Vianello M, Hasselman F, Adams BG, Adams RB Jr, Alper S, Aveyard M, Axt JR, Babalola MT, Bahník *et al.*: *Many Labs 2: Investigating Variation in Replicability Across Sample and Setting*. 2018 http://dx.doi.org/10.31234/osf.io/9654g. November 19.

33. Nematzadeh A, Meylan SC, Griffiths TL: **Evaluating vector-space models of word representation, or, the unreasonable effectiveness of counting words near other words**. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*. 2017:859-864.

34. Kintsch W: **Similarity as a function of semantic distance and amount of knowledge**. *Psychol Rev* 2014, **121**:559-561.

35. Jones MN, Gruenenfelder TM, Recchia G: **In defense of spatial models of semantic representation**. *New Ideas Psychol* 2018, **50**:54-60.

36. Gerz D, Vuli I, Hill F, Reichart R, Korhonen A: **SimVerb-3500: a large-scale evaluation set of verb similarity**. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 2016.

37. Mohammad SM, Dorr BJ, Hirst G, Turney PD: **Computing lexical contrast**. *Comput Linguist* 2013, **39**:1-60.

38. Shwartz V, Santus E, Schlechtweg D: **Hypernyms under siege: linguistically-motivated artillery for hypernymy detection**. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. 2017:65-75.

39. •• Bhatia S: **Associative judgment and vector space semantics**. *Psychol Rev* 2017, **124**:1-20.
The author shows how measures of association, obtained from DS models, capture the associations at play in high-level judgment, and subsequently predict probability assessments, event forecasts, and responses to trivia quizzes.

40. Bhatia S, Walasek L: **Association and response accuracy in the wild**. *Mem Cogn* 2018. [in press].

41. •• Caliskan A, Bryson JJ, Narayanan A: **Semantics derived automatically from language corpora contain human-like biases**. *Science* 2017, **356**:183-186.
The authors show associations derived from DS models predict classic Implicit Association Test results of gender, racial, and age bias.

42. Lenton AP, Sedikides C, Bruder M: **A latent semantic analysis of gender stereotype-consistency and narrowness in American English**. *Sex Roles* 2009, **60**:269-278.

43. Bolukbasi T, Chang KW, Zou JY, Saligrama V, Kalai AT: **Man is to computer programmer as woman is to homemaker? Debiasing word embeddings**. *Adv Neural Inf Process Syst* 2016, **29**:4349-4357.

44. Bhatia S: **The semantic representation of prejudice and stereotypes**. *Cognition* 2017, **164**:46-60.

45. Bhatia S, Goodwin G, Walasek L: **Trait associations for Hillary Clinton and Donald Trump in news media: a computational analysis**. *Soc Psychol Pers Sci* 2018, **9**:123-130.

46. Holtzman NS, Schott JP, Jones MN, Balota DA, Yarkoni T: **Exploring media bias with semantic analysis tools: validation of the Contrast Analysis of Semantic Similarity (CASS)**. *Behav Res Methods* 2011, **43**:193-200.

47. Mooijman M, Hoover J, Lin Y, Ji H, Dehghani M: **Moralization in social networks and the emergence of violence during protests**. *Nat Hum Behav* 2018, **2**:389-396.

48. Hopkins DJ: **The exaggerated life of death panels? The limited but real influence of elite rhetoric in the 2009–2010 health care debate**. *Polit Behav* 2018, **40**:681-709.

49. Li P, Schloss B, Follmer DJ: **Speaking two “languages” in America: a semantic space analysis of how presidential candidates and their supporters represent abstract political concepts differently**. *Behav Res Methods* 2017, **49**:1668-1685.

50. Dehghani M, Johnson K, Hoover J, Sagi E, Garten J, Parmar NJ, Vaisey S, Iliev R, Graham J: **Purity homophily in social networks**. *J Exp Psychol Gen* 2016, **145**:366-375.

51. Hoover J, Johnson K, Boghrati R, Graham J, Dehghani M: **Moral framing and charitable donation: integrating exploratory social media analyses and confirmatory experimentation**. *Collabra: Psychology* 2018, **4**:9.

52. Garten J, Hoover J, Johnson KM, Boghrati R, Iskiwitch C,
• Dehghani M: **Dictionaries and distributions: combining expert knowledge and large scale textual data content analysis**. *Behav Res Methods* 2018, **50**:344-361.
The authors propose a technique for building psychological dictionaries and determining the psychological properties of text using DS models.

53. Garg N, Schiebinger L, Jurafsky D, Zou J: **Word embeddings**
•• **quantify 100 years of gender and ethnic stereotypes**. *PNAS* 2018, **115**:E3635-E3644.
The authors use historical word vectors to investigate 20th century trends in gender and ethnic associations, and show how occupational and personality stereotypes for women and ethnic minorities evolved over time.

54. Kozlowski AC, Taddy M, Evans JA: *The Geometry of Culture: Analyzing Meaning through Word Embeddings*. 2018. arXiv:1803.09288.

55. Hamilton WL, Leskovec J, Jurafsky D: **Diachronic word embeddings reveal statistical laws of semantic change**. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*. 2016:1489-1501.

56. Bhatia S, Stewart N: **Naturalistic multiattribute choice**. *Cognition* 2018, **179**:71-88.

57. Bhatia S: **Predicting risk perception: new insights from data**
• **science**. *Manag Sci* 2018. [in press].
The author shows that risk judgments about a range of risk sources can be predicted with a high degree of accuracy by regressing risk ratings directly on word vectors obtained from DS models.

58. Paetzold G, Specia L: **Inferring psycholinguistic properties of words**. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2016:435-440.

59. Hollis G, Westbury C, Lefsrud L: **Extrapolating human judgments from skip-gram vector representations of word meaning**. *Q J Exp Psychol* 2017, **70**:1603-1619.

60. Van Rensbergen B, De Deyne S, Storms G: **Estimating affective word covariates using word association data**. *Behav Res Methods* 2016, **48**:1644-1652.

61. Sedoc J, Preoţiuc-Pietro D, Ungar L: **Predicting emotional word ratings using distributional representations and signed clustering**. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*. 2017:564-571.

62. Mitchell TM, Shinkareva SV, Carlson A, Chang KM, Malave VL, Mason RA, Just MA: **Predicting human brain activity associated with the meanings of nouns**. *Science* 2008, **320**:1191-1195.

63. Pereira F, Lou B, Pritchett B, Ritter S, Gershman SJ, Kanwisher N, Botvinick M, Fedorenko E: **Toward a universal decoder of linguistic meaning from brain activation**. *Nat Commun* 2018, **9**:963.

64. Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL: **Natural speech reveals the semantic maps that tile human cerebral cortex**. *Nature* 2016, **532**:453-458.

65. Mitchell J, Lapata M: **Composition in distributional models of semantics**. *Cogn Sci* 2010, **34**:1388-1429.

66. Le Q, Mikolov T: **Distributed representations of sentences and documents**. *International Conference on Machine Learning*. 2014:1188-1196.

67. Socher R, Perelygin A, Wu J, Chuang J, Manning CD, Ng A, Potts C: **Recursive deep models for semantic compositionality over a sentiment treebank**. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. 2013:1631-1642.

68. Beltagy I, Roller S, Cheng P, Erk K, Mooney RJ: **Representing meaning with a combination of logical and distributional models**. *Comput Linguist* 2016, **42**:763-808.

69. Anderson AJ, Bruni E, Bordignon U, Poesio M, Baroni M: **Of words, eyes and brains: correlating image-based distributional semantic models with neural representations of concepts**. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. 2013:1960-1970.

70. Anderson AJ, Kiela D, Clark S, Poesio M: **Visually grounded and textual semantic models differentially decode brain activity associated with concrete and abstract nouns**. *Trans Assoc Comput Linguist* 2017, **5**:17-30.

71. Hoffman P, McClelland JL, Ralph L, Matthew S: **Concepts, control, and context: a connectionist account of normal and disordered semantic cognition**. *Psychol Rev* 2018, **125**:293-328.

72. De Deyne S, Navarro DJ, Perfors A, Brysbaert M, Storms G: **The "small world of words" English word association norms for over 12,000 cue words**. *Behav Res Methods* 2018:1-20.

73. Thompson GW, Kello C: **Walking across Wikipedia: a scale-free network model of semantic memory retrieval**. *Front Psychol* 2014, **5**:86.

74. Kenett YN, Levi E, Anaki D, Faust M: **The semantic distance task: quantifying semantic distance with semantic network path length**. *J Exp Psychol Learn Mem Cogn* 2017, **43**:1470-1489.