REPLY

Hidden Processes in Structural Representations: A Reply to Abbott, Austerweil, and Griffiths (2015)

Michael N. Jones
Indiana University, Bloomington

Thomas T. Hills
University of Warwick

Peter M. Todd
Indiana University, Bloomington

In recent work exploring the semantic fluency task, we found evidence indicative of optimal foraging policies in memory search that mirror search in physical environments. We determined that a 2-stage cue-switching model applied to a memory representation from a semantic space model best explained the human data. Abbott, Austerweil, and Griffiths demonstrate how these patterns could also emerge from a random walk applied to a network representation of memory based on human free-association norms. However, a major representational issue limits any conclusions that can be drawn about the process model comparison: Our process model operated on a memory space constructed from a learning model, whereas their model used human behavioral data from a task that is quite similar to the behavior they attempt to explain. Predicting semantic fluency (e.g., how likely it is to say *cat* after *dog* in a sequence of animals) from free association (how likely it is to say *cat* when given *dog* as a cue) should be possible with a relatively simple retrieval mechanism. The 2 tasks both tap memory, but they also share a common process of retrieval. Assuming that semantic memory is a network from free-association behavior embeds variance due to the shared retrieval process directly into the representation. A simple process mechanism is then sufficient to simulate semantic fluency because much of the requisite process complexity may already be hidden in the representation.

Keywords: semantic fluency, free association, optimal foraging, cognitive model

The semantic fluency task (SFT; e.g., "name all the animals you can in a minute") is widely used[1] to study memory retrieval in both experimental and clinical settings. A key finding in SFT is that participants tend to produce temporal bursts of semantically related items with longer lags between bursts. This "patchy" response pattern has led to the proposal that memory retrieval in the task is the product of two distinct processes: a local search process that generates a series of related items based on interitem similarity and a global process that moves between local regions of the search space when these regions become depleted (e.g., Gronlund & Shiffrin, 1986; Troyer, Moscovitch, & Winocur, 1997).

In previous work (Hills, Jones, & Todd, 2012), we compared the response patterns in SFT to patterns seen when animals are foraging for food and found evidence that humans searching memory in SFT produced statistical signatures that are characteristic of optimal foraging policies observed in animals searching for resources in physical space (Charnov, 1976). This correspondence suggests that the mechanisms for searching memory may have been exapted from mechanisms that evolved to forage for resources in the physical world (Hills, 2006; Hills et al., 2015). However, alternate explanations may also account for the behavioral patterns: it may simply be the case that memory is organized in such a patchy fashion that a model randomly producing items in relation to their similarity to previously produced items would naturally generate data that appear indicative of optimal foraging.

We tested a variety of models of search, based on classic cue-combination memory search models (Anderson, 1990; Raaijmakers & Shiffrin, 1981) applied to a spatial representation of semantic memory constructed by a corpus-based distributional

Michael N. Jones, Cognitive Science Program and Department of Psychological and Brain Sciences, Indiana University, Bloomington; Thomas T. Hills, Department of Psychology, University of Warwick; Peter M. Todd, Cognitive Science Program and Department of Psychological and Brain Sciences, Indiana University, Bloomington.

Correspondence concerning this article should be addressed to Michael N. Jones, Department of Psychological and Brain Sciences, 1101 East 10th Street, Indiana University, Bloomington, IN 47404. E-mail: jonesmn@indiana.edu

[1] A PubMed search for semantic/category/verbal fluency returned 5,194 papers that have used the task.

mechanism (BEAGLE; Jones & Mewhort, 2007). The model that best explained the human data was one that was able to switch between two cues: local similarity was used to generate items until no other item was close enough to pass a threshold, and then the model switched to a global frequency cue to select the next item (and search reverted again to local similarity).

Abbott, Austerweil, and Griffiths (2015; hereafter AAG) point out historical issues of model "mimicry" that may well be at play in our analysis. Our findings may be dependent on the memory representation we used.[2] Furthermore, our rejection of a one-stage random selection model, in favor of a two-stage local versus global switching model, is potentially at odds with the success of random walk models in other areas of psychology (e.g., Griffiths, Steyvers, & Firl, 2007). To illustrate the issue of model mimicry, AAG first apply a random walk model to a network representation constructed from our BEAGLE space and replicate our finding that such a model is insufficient to simulate the human data. They then apply a random walk to a network representation constructed from free-association norms (Nelson, McEvoy, & Schreiber, 2004) and find that this model reproduces the behavioral markers indicative of optimal foraging that our local/global cue-switching model is able to. They propose that a random walk over a free-association network is an alternative account of SFT that is equally plausible to our cue-switching model over a semantic space.

AAG analyzed the representations used by each model (BEAGLE space vs. free-association network) and determined that the behavioral bursts in SFT were better represented as clusters in the association network than they were in the BEAGLE space, and this embedded structure is what allowed their random walk model to mimic the optimal foraging predictions of our cue-switching model. They conclude, "A random walk on a semantic network produces optimal foraging behavior, but a random walk on a corresponding spatial representation does not. Consequently, it seems that there is something special about the semantic network representation that allows the simple random walk to appear similar to optimal foraging" (Abbott et al., 2015).

We certainly concur with AAG that this is a research area ripe for model comparison and that constructing competing models of SFT is important to help design new experiments that constrain theoretical accounts and home in on the true generating process. However, the competing accounts being compared must be equivalent, or no firm conclusions can be drawn, and we risk rejecting what may well be an excellent model of the processes and representations that humans actually use. For example, the casual reader may conclude from the AAG article that a network representation of memory is more like what humans have in their heads than a spatial model, and Occam's razor then favors the more parsimonious random walk as the process that humans use when retrieving items in SFT over our local/global cue-switching model. The free-association network is a model of memory structure that better represents the behavioral clusters seen in SFT, and this allows a simple random walk to successfully produce patterns consistent with optimal foraging.

However, there is a deep and broadly important issue concerning the structural representations of memory and the "something special" that AAG suggest about the semantic network used in their model compared to the spatial representation used in ours. Simply put, the spatial representation was based on a statistical model of learning, whereas the network representation was quite

literally built from human behavioral data on a task similar to SFT. The process used in SFT is already partially embedded in AAG's network representation.

Hidden Process Complexity in the Structural Representation

A model of a cognitive phenomenon requires an account of both structure and process and how the two interact (Estes, 1975). Too often in cognitive modeling, the structural representation is ignored in favor of the process account (Johns & Jones, 2010). AAG are explicit in their model that memory is an associative network over which a random walk operates. Similarly, we were explicit that memory in our model is a spatial representation over which a local/global search process operates. However, these two proposals differ in a crucial way.

BEAGLE (Jones, Kintsch, & Mewhort, 2006; Jones & Mewhort, 2007) is a mechanistic model of how humans construct a semantic space from statistical redundancies in the environment (trained on a natural language corpus). In contrast, free association is a behavioral task. The free-association norms (Nelson et al., 2004), used by AAG to construct their network, are aggregated human responses to cue words—behavioral data. BEAGLE explains, for instance, how the pair *dog-cat* become similar to one another in memory by applying a learning mechanism to environmental regularities that humans learn from. Indeed, in early language development, the structure of actual language is a better predictor of word learning than are free-association norms (Hills, Maouene, Riordan, & Smith, 2010; Hills, 2013).

Free-association norms are simply a tabulation of how often humans give, for instance, *cat* as a first response to *dog* as a cue. Free association is not a pure readout of the *structure* of semantic memory—it also incorporates the *process* of memory retrieval. A concern, then, is that AAG's "representation" of memory is based on human responses from an association task that is already very similar to the SFT that the model is meant to explain. Hence, much of the requisite variance needed to account for the process of SFT may already be embedded in the representation. As we describe in the next section, the problem here—proposing that memory is represented as what people do in a task very similar to SFT and then concluding that a simple process model is sufficient to simulate the SFT behavioral data—is an instance of what we refer to as a *Turk problem* in cognitive modeling. Turk problems are so named in reference to the 18th-century chess-playing automaton, and we use it to express the "man in the machine" issue of using human behavioral data as a cognitive model's mental representation. Rather than explaining anything, this approach primarily predicts one kind of behavior (here SFT) from another very similar behavior (free association).

---

[2] We have also noted this in prior work, as individual differences and personal experience can also influence the structure of semantic representation (Hills, Mata, Wilke, Samanez-Larkin, 2013; Hills & Pachur, 2012). Indeed, in those two studies and in Hills et al. (2012), multiple representations were compared against one another, including frequency, hypergraphs (categorical structures), and proximity measures (such as semantic space and social proximity). In all cases, results favored a two-stage model over a one-stage model, with search in a proximity-based local representation interspersed with occasional long-distance transitions using an alternate representation (e.g., frequency).

Free association is a dependent variable—it is something that we need to explain, not a pure measure of how memory is organized. Nelson and McEvoy have noted this in virtually all of their publications on free association, but the practice of using their association norms as a proxy for human memory is still widespread. For example, Nelson et al. (2004) originally emphasized the potential of using free association norms "for evaluating models of semantic representation" (p. 403) and noted that free association is a good predictor of many memory tasks because it uses some of the same retrieval processes used in those tasks, as well as tapping some commonalities of semantic memory. Although free association was intended as a dependent variable, Nelson, McEvoy, and Dennis (2000) note that its primary use is as a yardstick for stimulus control—but very little work has been done to investigate the yardstick itself: "Unlike printed frequency norms in which words are counted across samples of text, words are produced in free association, and we know little about the representations and processes involved" (p. 888).

The tasks of free association and SFT are extremely similar, not only because they both tap semantic memory representations but also because they draw on similar retrieval processes. The similarity is clear in AAG's analysis showing that the clusters in the free-association network are already very similar to the SFT response clusters. SFT can be thought of as a continuous version of the free-association task constrained to a given category (animals). In fact, Bousfield and Sedgewick (1944) originally referred to the task as "sequences of restricted associative responses." In introducing the task, they note that it "is essentially an extension of the classic word association" task (p. 149).

Consider the argument ad absurdum: an even simpler zero-parameter model that perfectly accounts for the SFT data can be built based on the assumption that human memory is represented as an associative chain using data from SFT. This is obviously absurd because it suggests that no process is required and just simulates SFT data from SFT data. Although it is a very parsimonious and predictive model, it has no explanatory power. Predicting behavior from similar behavior without generating new understanding is the problem underlying the idea that SFT can be simulated from free-association data.

Hence, we do not believe that the critical difference between our model and the AAG model is a space versus a network or a random walk versus a cue-switching model. From a graph-theoretic perspective, spaces are networks—fundamentally, both are matrices. Moreover, as noted earlier, both models use random walks over local proximity-based representations. The real issue is, in AAG's words, the following: "A random walk on a semantic network derived from free association produces phenomena suggestive of optimal foraging, while a random walk on a spatial representation generated by BEAGLE does not. This raises the natural question: Why? What is the critical difference between these two representations?" (Abbott et al., 2015). Our response is simple: much of the variance required to simulate the SFT *process* is already partially embedded in the free-association representation.

BEAGLE may very well be wrong as a model of human semantic representation. But a much more informative model comparison would be to pit BEAGLE (a spatial model) against a network created with a Topic model (Griffiths, Steyvers, & Tenenbaum, 2007)—Topic models have been used to generate networks and are a natural contender. Then, both learning models could be applied to the same experience (a text corpus), and the resulting network and space could be fully crossed with a random walk or cue-switching model.

Notably, the problem of inferring process without attending to the influence of the search environment has also led to controversy in the ecological search literature that inspired this cognitive work. Lévy flights have been proposed to parsimoniously model animal foraging because, not surprisingly, sampling from a power-law distribution of path lengths produces the power-law distribution of path lengths that animals often produce when foraging for food. However, closer inspection of animal behavior and the resource environments in which they forage has found that animals change their behavior to increase time in resource-rich environments while minimizing time in resource-poor environments. This behavior (often called area-restricted search) also produces power-law distributed path lengths (Benhamou, 2007; Hills, Kalff, & Wiener, 2013; Plank & James, 2008) and is more optimal than Lévy flights (Ferreira, Raposo, Viswanathan, & da Luz, 2012; Plank & James, 2008). Hence, assumptions that overly simplify the search environment are likely to get the search process wrong.

The Representational Turk Problem in Cognitive Modeling

As indicated earlier, the practice of using free-association norms as a proxy for human semantic memory has a considerable history in cognitive modeling (e.g., De Deyne, Navarro, & Storms, 2013; Steyvers, Shiffrin, & Nelson, 2004). However, it is important to remember that any representation based on behavioral data also contains variance from the process used to produce the response. Oddly, the opposite problem is seen in the semantic representation literature: Complex learning models are used to generate estimates of memory structure, but then a remarkably simple metric (e.g., vector cosine) is used to predict human data (see Jones, Willits, & Dennis, in press, for a review). However, a cosine is certainly not a process model. Hence, the semantic representation literature needs to likewise consider process accounts (retrieval, similarity judgments, etc.) when fitting representational models to human data.

Examples of embedding human behavior in a model to then predict similar human behavior are easy to find across the field. For example, Glenberg and Robertson (2000) demonstrated the superiority of their indexical hypothesis over latent semantic analysis (LSA; Landauer & Dumais, 1997) at explaining human comprehension of sentence affordances. They found that LSA could not distinguish between afforded sentences (using a newspaper to protect one's face from the wind) and nonafforded sentences (using a matchbook to protect one's face from the wind), but this distinction was naturally predicted by their indexical hypothesis. Humans rated the afforded sentences as easier to "envision" than the nonafforded sentences. However, the affordance predictions of Glenberg and Robertson's account were generated from sensibility ratings (virtual nonsense to completely sensible) made by the same subjects who produced the envisioning ratings (impossible to imagine to easy to imagine) that were used as the dependent variable. Not only were the two ratings on judgments almost identical, but they were also from the same participants. Hence, it is both impossible for a learning model such as LSA to outperform

humans and questionable what explanatory power comes from such an exercise.[3]

Certainly, the use of norms as a proxy for human representation has a very fruitful history in cognitive modeling (e.g., McRae, de Sa, & Seidenberg, 1997; Smith, Shoben, & Rips, 1974). However, the representational Turk problem is becoming more common in cognitive modeling as we strive for parsimony under threat of Occam's razor. We have formal methods for penalizing models based on parametric complexity (Lewandowsky & Farrell, 2010). However, we only penalize the process account; the field does not yet have a good idea how to "tax" a model based on its representational assumptions. The representation is becoming the Cayman Islands of modeling, where complexity may be hidden from such taxation to support a simpler process mechanism, often unintentionally. Hummel and Holyoak (2003) have gone so far as to suggest that representational complexity may be the single most serious problem facing cognitive modeling as a scientific enterprise. Particularly as we see more large-scale models making use of naturalistic "big data" in their representations (Griffiths, 2015; Jones, 2015), it is important to be vigilant about process variance hiding in the representation.

---

[3] A rebuttal by Burgess (2000) summarized the flaws that made the Glenberg and Robertson (2000) study an unfair evaluation of LSA. Although Burgess's rebuttal appeared immediately in tandem in the same issue as the Glenberg and Robertson article, the rebuttal has been cited only 19 times compared to 485 citations for the Glenberg and Robertson article.

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. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Benhamou, S. (2007). How many animals really do the Lévy walk? *Ecology, 88,* 1962–1969. http://dx.doi.org/10.1890/06-1769.1

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology, 30,* 149–165. http://dx.doi.org/10.1080/00221309.1944.10544467

Burgess, C. (2000). Theory and operational definitions in computational memory models: A response to Glenberg and Robertson. *Journal of Memory and Language, 43,* 402–408. http://dx.doi.org/10.1006/jmla.2000.2715

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology, 9,* 129–136. http://dx.doi.org/10.1016/0040-5809(76)90040-X

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior Research Methods, 45,* 480–498. http://dx.doi.org/10.3758/s13428-012-0260-7

Estes, W. K. (1975). Some targets for mathematical psychology. *Journal of Mathematical Psychology, 12,* 263–282. http://dx.doi.org/10.1016/0022-2496(75)90025-5

Ferreira, A. S., Raposo, E. P., Viswanathan, G. M., & da Luz, M. G. E. (2012). The influence of the environment on Lévy random search efficiency: Fractality and memory effects. *Physica A: Statistical Mechanics and Its Applications, 391,* 3234–3246.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43,* 379–401. http://dx.doi.org/10.1006/jmla.2000.2714

Griffiths, T. L. (2015). Manifesto for a new (computational) cognitive revolution. *Cognition, 135,* 21–23. http://dx.doi.org/10.1016/j.cognition.2014.11.026

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

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114,* 211–244. http://dx.doi.org/10.1037/0033-295X.114.2.211

Gronlund, S. D., & Shiffrin, R. M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 550–561. http://dx.doi.org/10.1037/0278-7393.12.4.550

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cognitive Science, 30,* 3–41. http://dx.doi.org/10.1207/s15516709cog0000_50

Hills, T. (2013). The company that words keep: Comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language, 40,* 586–604. http://dx.doi.org/10.1017/S0305000912000165

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review, 119,* 431–440. http://dx.doi.org/10.1037/a0027373

Hills, T. T., Kalff, C., & Wiener, J. M. (2013). Adaptive Lévy processes and area-restricted search in human foraging. *PLoS ONE, 8,* e60488. http://dx.doi.org/10.1371/journal.pone.0060488

Hills, T., Maouene, J., Riordan, B., & Smith, L. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language, 63,* 259–273. http://dx.doi.org/10.1016/j.jml.2010.06.002

Hills, T. T., Mata, R., Wilke, A., & Samanez-Larkin, G. R. (2013). Mechanisms of age-related decline in memory search across the adult life span. *Developmental Psychology, 49,* 2396–2404. http://dx.doi.org/10.1037/a0032272

Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 218–228. http://dx.doi.org/10.1037/a0025161

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

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110,* 220–264. http://dx.doi.org/10.1037/0033-295X.110.2.220

Johns, B. T., & Jones, M. N. (2010). Evaluating the random representation assumption of lexical semantics in cognitive models. *Psychonomic Bulletin & Review, 17,* 662–672. http://dx.doi.org/10.3758/PBR.17.5.662

Jones, M. N. (Ed.). (2015). *Big data in cognitive science: From methods to insights.* Sussex, England: Psychology Press.

Jones, M. N., Kintsch, W., & Mewhort, D. J. K. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language, 55,* 534–552. http://dx.doi.org/10.1016/j.jml.2006.07.003

Jones, M. N., & Mewhort, D. J. K. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review, 114,* 1–37. http://dx.doi.org/10.1037/0033-295X.114.1.1

Jones, M. N., Willits, J., & Dennis, S. (in press). Models of semantic memory. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *Oxford handbook of computational and mathematical psychology.* Oxford, UK: Oxford University Press.

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240. http://dx.doi.org/10.1037/0033-295X.104.2.211

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99–130. http://dx.doi.org/10.1037/0096-3445.126.2.99

Nelson, D. L., McEvoy, C. L., & Dennis, S. (2000). What is free association and what does it measure? *Memory & Cognition, 28,* 887–899. http://dx.doi.org/10.3758/BF03209337

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

Plank, M. J., & James, A. (2008). Optimal foraging: Lévy pattern or process? *Journal of the Royal Society, Interface, 5,* 1077–1086. http://dx.doi.org/10.1098/rsif.2008.0006

Raaijmakers, J., & Shiffrin, R. (1981). Search of associative memory. *Psychological Review, 88,* 93–134. http://dx.doi.org/10.1037/0033-295X.88.2.93

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241. http://dx.doi.org/10.1037/h0036351

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

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11,* 138–146. http://dx.doi.org/10.1037/0894-4105.11.1.138