UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Conceptual Expansion During Divergent Thinking

Permalink

https://escholarship.org/uc/item/3w98s6rc

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Author

Hass, Richard W.

Publication Date

2016

Peer reviewed

Conceptual Expansion During Divergent Thinking

Richard W. Hass (hassr@philau.edu)
College of Science, Health, and the Liberal Arts, Philadelphia University, 4201 Henry Avenue,
Philadelphia, PA, 19144, USA

Abstract

Recent research on creative thinking has implicated conceptual expansion as potential cognitive underpinnings. These theories were examined within the context of a laboratory study using two divergent thinking prompts. Participants generated alternative/creative uses for a brick and for a glass bottle (separately) for two minutes and responses were time-stamped using a Matlab GUI. Semantic distances between responses and conceptual representations of the DT prompts were computed using latent semantic analysis. Results showed that semantic distance increased as responding progressed, with significant differences between the two tasks, and intraparticipant variation. Results have implications for theories of creative thinking and represent methodological and analytic advances in the study of divergent thinking.

Keywords: creativity; semantic distance; latent semantic analysis; divergent thinking; conceptual expansion

Conceptual Expansion During Divergent Thinking

Divergent thinking is a problematic topic in the study of creativity for many reasons. One issue is that divergent thinking (DT) refers to both a psychometric construct—thinking in multiple directions—and the set of tasks used to quantify the construct (for a full discussion see Hass, in revision). Perhaps because of DT's psychometric roots, cognitive analysis of creative thinking often omits reference to DT studies (e.g., Finke, Ward, & Smith, 1992; Weisberg, 2006). However, recently there has been a surge of neuroscientific studies using DT as a proxy for creative thinking, citing among other points that DT tests have some predictive validity for real world creative success (Kim, 2006). Despite criticisms of DT as a means to assess creativity (e.g., Weisberg, 2006, Ch. 9), if they are to be used in neuroscientific studies, then cognitive theories must be developed to explain what is transpiring *during* DT. This paper represents one part of a larger project to try to do just that. With processing data in hand, interpretation of neuroscientific studies of DT will become much more straightforward and useful for a cognitive science of creative thinking.

Cognitive explanations for DT performance

Many neuroscientific studies using DT as a proxy have shown that originality on DT tasks is related both to the brain's cognitive control and default mode networks (for a review see Beaty, Benedek, Kaufman, & Silvia, 2015). The main conclusion drawn from these studies is that better control of self-directed thought defines improved performance on DT tasks. However, most studies are correlational in that they do not ask how cognitive control operates during DT, rather, cognitive control is assessed on a separate task and correlated to DT performance. For example, Zabelina, Saporta, and Beeman (2015) showed that DT performance was positively related to how well participants overcame an invalid cue that preceded 20% of trials on the Navon (1977) Local-Global Letter Task. However, there was no relation between DT performance and attention filtering (assessed in terms of a congruency effect on the Letter Sets Task). Moreover, these DT-attention relationships did not match effects relating individual differences in attention to individual differences in real-world creative achievement. The main question asked by the current analysis is whether tracking cognition during DT response generation can shed more light on these kinds of conflicting results.

Only a single study attempting to track cognition during DT exists. Gilhooly, Fioratu, Anthony, and Wynn (2007) took verbal protocols from participants and found that they often invoked distinct strategies during DT. For example, in thinking of alternative uses for a shoe, many participants engaged in self-cuing (repeating the word "shoe"), and reconstructed the problem representation by mentally disassembling the object (i.e., using only the laces of the shoe). So it seems that there may be several levels of cognitive processing operating during DT, and it is imperative that we move toward studies that quantify those processes, rather than rely purely on correlational data. Though the current analysis did not involve verbal protocols, the next section outlines a cognitive framework for the kind of data that were colletted.

Conceptual expansion. Ward (2008) described conceptual expansion as the formation of novel exemplars of a concept during [creative] problem solving. Indeed, Abraham (2014)—in her theoretical examination of the conceptual expansion hypothesis—used a DT task (think of alternative / creative uses for a shoe) as her primary example of conceptual expansion. She argued that envisioning the use of a shoe as a plant pot or as a pencil holder, by definition, expands upon the canonical concept of shoe. Abraham and colleagues (2012) showed evidence of differential brain activity when participants generated common versus unusual responses to DT prompts (see also Chrysikou & Thompson-Schill, 2011). Given Ward's definition of conceptual expansion, they reasoned that this additional activation was evidence of a conceptual expansion processes during "unusual" idea generation.

There are many reasons that conceptual expansion represents a good theoretical framework for creative thinking. Particularly, it allows for research to focus on specific questions regarding the process. For example, does conceptual expansion unfold in a linear fashion? Do people actively monitor the amount of expansion in the responding? Is expansion related to processing speed? Is expansion another way to describe analogical transfer?

Before answering these questions it is important to settle on an operational definition of conceptual expansion. In this analysis, conceptual expansion was operationally defined as the degree of semantic distance between DT responses and the prompt (e.g., think of alternative uses for a brick). Semantic distance was derived from cosine similarity scores obtained via latent semantic analysis (LSA, e.g., Landauer & Dumais, 1997). Though LSA is not a one-to-one mapping of conceptual expansion, it is of interest to examine relationships among semantic distance and response order and inter-response time.

Serial order, response time, and semantic distance

The so-called serial order effect has been described in many studies showing that generally people provide more creative responses to DT prompts later in response array's (e.g., Beaty & Silvia, 2012; Christensen, Guildford, & Wilson, 1957). Beaty and Silvia found that the originality of DT responses (scored with a subjective system) increased as a function of response order, but that participants with higher fluid intelligence scores began with more creative responses during DT than participants with lower fluid intelligence scores, and showed less of an increase. Hass (in revision) replicated the analysis using semantic distance and growth-curve modeling and showed that high fluid intelligence scores related to higher initial semantic distance during DT.

Though the serial order effect seems to be well established, a cognitive explanation is less clear. If it is the case that associative processes spur recall of the concepts that map onto DT responses, then response latencies should be related to the distance between the conceptual content in the response and the conceptual content in the DT prompt (e.g., Kahana, 1996). This hypothesis is directly tested in the current study.

Is LSA a valid means of measuring conceptual expansion?

Before describing the method and results, it is important to discuss the validity of LSA-derived semantic metrics in DT studies. Hass (in revision) provided a discussion of the use of LSA-derived distances in scoring DT responses as opposed to other semantic methods (see also Harbison & Haarmann, 2014). The crux of the argument was that if the distances are culled from comparisons between each response and a fixed conceptual representation of the DT prompt (e.g., *brick*), then the metric has both construct validity and convergent validity with subjective scores. This approach is similar to more traditional DT flexibility

scoring[1] (e.g., Madore, Addis, & Schacter, 2015), but with the added benefits of a continuous scale of measurement and the availability of computational models. Both techniques target persisting themes in cognitive theories of creative thinking: remote association (e.g., Mednick, 1962) and conceptual expansion. However, since flexibility scores rely on the creation of ad-hoc categories after data is collected, the system is potentially biased and also provides low-resolution information regarding the graded structure of categories (cf., Gabora, Rosch, & Aerts, 2008).

Also, unlike the current study, prior analyses of DT data with LSA have seemed focused on *replacing* subjective scoring with semantic scoring, which essentially keeps DT tied to the psychometric "summary score" approach. For example, Forster and Dunbar (2009) showed that LSA-derived semantic distance scores from DT data were correlated with originality ratings, and that since distances are objectively calculated, they may be preferred to subjective scoring. Harbison and Haarmann (2014) similarly showed that subjective scores and distances correlated, though they also showed that another natural language processing procedure (point-wise mutual information) was more highly correlated with subjective scores.

Rather than persist with the summary score approach, this analysis used Growth Curve Modeling (e.g., Mirman, Dixon, & Magnuson, 2007) to examine individual differences in the serial-order effect, and to examine relationships between serial order, inter-response time, and semantic distance (as a proxy for conceptual expansion). Variations in semantic distance within individuals were also examined across two oft-used DT prompts: think of creative alternative uses for a brick, and for a glass bottle. It was expected that responses would increase in distance as a function of response order, and also that there would be a linear relationship between distance and IRTs. Prior analyses also revealed differing levels of semantic distance across DT prompts (Hass, in revision), so that analysis was also performed.

Method

Participants

Sixty participants (18 females) were recruited from the participant pool at a large state college in New Jersey. The average age of participants was 19.45 years (SD = 1.46). All participants were given partial course credit for participating. Participants provided informed consent prior to participation. Time-stamp malfunctions led to the elimination of data from three participants.

Materials

All materials were presented on Lenovo ThinkVision monitors. Participants typed responses on computer keyboards. The experiment was automatically administered

---

[1] Flexibility is defined as the number of category switches in a response array.

using a custom Matlab GUI which provided an editable response field for participants to enter responses. Matlab timestamped both the initial keypress for each response and the final return key. Pressing return cleared the response from the response field and pasted it below to keep a running log of the participants' responses so he or she would be encouraged to continue producing novel responses. Prompts appeared in 50-point font and were visible throughout response generation. Responses appeared in 36-point font.

Procedure

Participants were greeted by an experimenter and filled out a demographic survey while the experimenter initiated Matlab. Instructions appeared on the screen and the experimenter read aloud to the participant the initial instructions regarding using the keyboard to enter responses. Participants then engaged in a short category generation task (30s of naming colors) to grow accustomed to the experimental setting. After that, participants were randomly assigned to two un-related task conditions that lasted 5 minutes[2]. Finally, participants were presented with the instructions for the DT tasks. Participants were told to think of creative uses for common objects that would be presented in text on the screen. They were told that there would be two such tasks and that they would have 2 minutes to complete each task.

The task prompt then appeared above the response field, with the order of the two prompts (brick, glass bottle) randomized by Matlab. The prompts read "Think of uses for a Brick *besides* building a wall" and "Think of uses for a Glass Bottle *besides* holding liquid". These instructions were designed to increase the validity of the semantic analysis using the canonical concept of brick as a building material and bottle as a liquid holder. Participants were instructed to continue responding until time had expired. When the two minutes per task expired, Matlab displayed a message to indicate that the next task was loading. The inter-task time was 10 seconds to allow for a brief break. After completion of the second task, a thank you message appeared on the screen.

Results

Data preparation and semantic analysis

LSA was performed using the tools available at lsa.colorado.edu. Analysis was performed using the data from the TASA corpus, compiled to represent general semantic knowledge gained from primary school through the first year in college. Three hundred factors were used, in keeping with prior analyses that used this tool (e.g., Forster & Dunbar, 2009). Prior to LSA, all responses were spell-checked, and a set of stopwords was removed using functions from the *tm* package (Feinerer, Hornik, & Meyer,

2008) in the R Statistical Programing Environment (R Core Team, 2015).

The "one-to-many" LSA tool was used to compare each DT response from the dataset to a target phrase—a composite description of the DT prompt compiled from Merriam Webster Dictionary entries (see Hass, under review) in document space. The phrase representing the *brick* concept was, "a small, hard block of baked clay that is used to build structures such as houses and sometimes to make streets and paths" (see http://www.merriam-webster.com/dictionary/brick). A similar phrase was used for the glass bottle comparison. LSA represents phrases as the centroid of the word vectors contained in the phrase. The centroid is essentially a vector average, and thus represents a sort of blend of the meanings of the words in each response. This method of representation has been shown to work well for long-passages of text such as student essay responses (e.g., Rehder, et al., 1998).

For each response, the LSA tool computed the cosine of the angle between the vector representing the target (the DT prompt) and the vector representing the response. This represents the *similarity* of two vectors, such that the cosine of the angle between two identical vectors is 1, the cosine of two orthogonal (i.e. unrelated) vectors is 0, and the cosine of two vectors pointing in opposite directions is -1. The cosine similarity values were then transformed into to *distances* by subtracting each from 1 (e.g., Prabhakaran, Green, & Gray, 2013).

Table 1: Descriptive statistics by DT prompt. Inter-response times, and distances were analyzed at the level of response. Fluency was analyzed at the level of participant.

Prompt	Variable	M	SD	s.e(M)
Brick (n_{resp} = 402)	IRT	13.48	12.73	0.64
	Distance	0.88	0.14	0.01
	Fluency	7.04	3.09	0.40
Bottle (n_{resp} = 393)	IRT	13.62	12.31	0.62
	Distance	0.78	0.17	0.01
	Fluency	6.89	2.88	0.38

Inter-response time (IRT) was calculated as the difference in end-of-response time stamps between adjacent pairs of responses. IRT for the first response was defined simply as the time stamp of the first response. Table 1 provides descriptive statistics for IRTs and distances along with average fluency counts for each task.

Statistical Analysis

Inter-response time. Before examining a multilevel model for semantic distance, the relationship between IRTs and response order was examined with a simple correlation. The correlation was small but significant ($r(793) = .18$, $p < .001$). In addition to showing that participants took more time to respond as their 2 minutes on task elapsed, the small magnitude of the correlation means that response order and

---

[2] Antecedent task condition had no effect on of the results reported in this paper

IRT can be used in a linear model for semantic distance without collinearity issues.

Semantic Distance. A multilevel model for semantic distance was assembled because of the variation in fluency across participants, and to test for possible variation in the relationship between response order and distance across participants. Model testing followed procedures given by Mirman, and colleagues (2007). The significance of predictors and random effects was determined by comparing nested models with a likelihood ratio. For all models, semantic distance was the dependent variable, with response order, and IRT as level-1 predictors. Response order was rescaled with zero as the first response so that the intercept, and IRT was rescaled in grand-mean-deviation form. DT prompt was entered as a level-2 predictor.

Table 2 summarizes the various models compared in terms of model deviance (Mirman, et al., 2007), with significant differences identified as statistically significant likelihood ratios. Model 1 is a baseline linear growth model. The response-order coefficient was significant[3] ($\beta_{11} = 0.01$, 95%CI = (0.006, 0.013)), confirming an overall linear serial order effect. Model 2 examined potential nonlinearity in the response order effect. The comparison narrowly missed significance, suggesting that there was an inverted-U trend to the data, but that this did not explain much more of the variance in distances across responses than the linear response-order predictor. The addition of the IRT variable also did not improve the fit (comparison 3). So the best level-1 growth model for semantic distance is defined with a linear response order predictor, and a random intercept per participant.

Table 2: Results of Semantic distance model testing. Model 1 is nested in Model 2. Models 3 and 4 are nested in Model 5.

Comparison	logLik	ΔD	p
1. Order (linear)	349.13	-	-
2. Order (quadratic) v. 1	350.89	3.53	.06
3. IRT v. 1	350.55	2.84	.09
4. Prompt (intercept) v. 1	405.61	112.96	< .001
5. Prompt (slope) v. 4	455.24	99.26	< .001
6. Prompt x Order v. 5	459.61	8.74	.003

Comparisons 4 through 6 in Table 2 represent tests to determine whether participants' distances varied according to the DT prompt (brick v. glass bottle), and whether the serial order effect varied across prompts and across individuals. These comparisons represent the key contribution of growth-modeling as they allow for examination of variations across individuals.

Comparison 4 shows that there was a significant difference in the average semantic distance of first responses given by participants across prompts. The
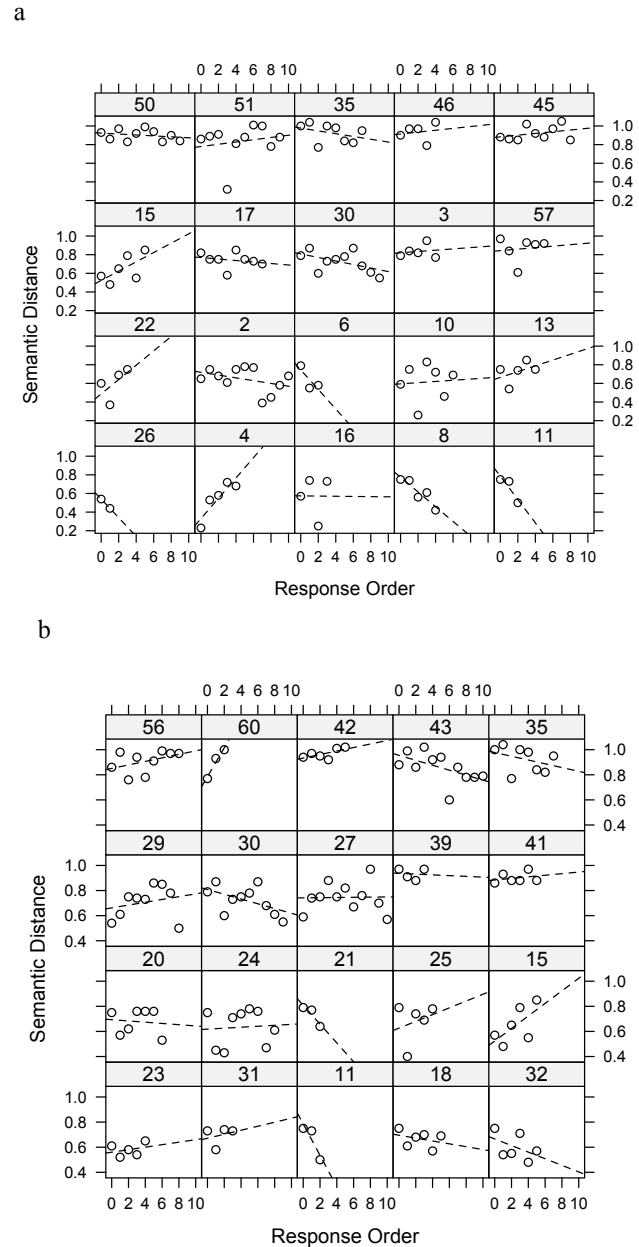
a

b

Figure 1: Variations in the serial-order effect across 20 randomly sampled participants. (a) responses to the Alternative Uses for a Brick prompt only, (b) responses to the Alternative uses for a Glass Bottle prompt only. Dotted lines represent OLS regression within participant.

semantic distances of initial responses were significantly lower on the bottle task compared to the brick task ($\gamma_{10} = -0.11$, 95%CI = (-0.13, -0.09)). This suggests that semantic distance is context dependent and so conceptual expansion may not be a general ability.

Finally, Figure 1 illustrates the results of comparisons 5 and 6, for 20 randomly sampled participants. There was both significant variation participants' changes in distance across the prompts, and a significant amount of variation in

---

[3] The use of *p*-values for evaluating coefficients in Multilevel models is controversial so 95% confidence intervals are reported.

participants' serial order effects. Thus, the serial-order effect may not be a universal phenomenon. Rather, semantic distance is a function of the conceptual content in the DT prompt, and though on average, semantic distances between responses and prompts tend to increase, inter-participant variability remains to be explained.

Discussion

The central aim of this paper was to examine conceptual expansion during divergent thinking and relate it to response order and IRT. Conceptual expansion was operationally defined in terms of the semantic distance between the concept represented in the DT prompt and a particular response. Several interesting results emerged, which in turn lead to new questions about DT, creative thinking, conceptual expansion, and response latency.

First, response latency did not directly relate to conceptual expansion. Rather, the degree of conceptual expansion shown by participants was more dependent upon both the concept represented by the DT prompt (brick v. glass bottle) and likely individual differences in semantic memory organization. This is consistent with other evidence that individual differences in semantic memory organization relate to individual differences in creative thinking and creative accomplishments (e.g., Kenett, Anaki, & Faust, 2014; Kenett, Beaty, Silvia, & Anaki, 2016). In those studies, network analysis was applied to category fluency responses from participants, rather than to DT responses, but the interconnectedness and flexibility of participants' semantic networks did indeed correlate with DT performance. Taken together with the verbal protocol analysis performed by Gilhooly and colleagues (2007), it seems that DT performance varies along with variations in participants' semantic processing, and likely according to their retrieval and cuing strategies, though verbal protocols were not taken from these participants.

Interestingly, the serial order effect does not seem to be a cognitive universal, nor does it seem that participants always need more time to come up with more distant responses. This is somewhat inconsistent with the remote-association account of creative thinking forwarded by Mednick (1962, see also Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014). According to Mednick, more creative people should generate more ideas when prompted, and the associations among ideas should be looser than less creative people (see also Wallach & Kogan, 1965). There are many problematic aspects of the theory for which the current study has implications. First, the variations in distances across prompts and within participants suggests that associative processes vary substantially according to the conceptual context and individual knowledge. That is, we should not assume that people should approach all creative idea generation tasks with the same amount of knowledge, or the same potential to expand on such knowledge.

Second, there may be two conceptual expansion processes operating during DT. In an analogous study, Smith, Vul, and Huber (2013) used LSA-derived semantic similarity to show that adjacent responses in a modified remote associates task (RAT) were semantically dependent. Performance on RAT items are often used to simulate creative insight (e.g., Kounios and Beeman, 2009). Smith and colleagues also argued that the search process is conscious given that responses are sequentially dependent. Though potential local dependencies were not examined in this paper, it is very likely that semantic distances between adjacent responses will illustrate some degree of dependence. If so, it may be evidence of both global and local conceptual expansion processes operating during DT. A global process might monitor the overall conceptual expansion with the DT prompt as the basis for comparison, while a local process might monitor the expansion needed for the next iteration compared with the previous response iteration. These data are amenable to such analysis, and I encourage others to investigate these local-global monitoring phenomena. If they exist, they provide context for the effects described by Zabelina and colleagues (2015) that multiple levels of attention and monitoring are differentially related to creative thinking and creative achievement.

Limitations. It should be noted that the use of LSA-derived distances as a conceptual-expansion metric is limited to the validity of the TASA corpus for representing DT responses. Indeed, there were a few cases in which responses (e.g., smartphone) were not found in the corpus, and responses had to be discarded. Also, though the corpus is able to resolve ambiguity in word meaning through co-occurrence data, there are likely places in which creative wordplay (e.g., use of a brick as "the weight of life") that might yield invalid LSA cosines.

Despite these limitations, this analysis stands as the first step toward understanding how people approach creative thinking tasks like these DT problems from the perspective of cognitive science. Continued examination creative thinking data using semantic distance and other related techniques, couched in growth curve models is highly recommended. Among other issues, this type of analysis is likely to address some of the inconsistencies in creative thinking study results when DT summary scores are correlated with measures of cognitive processing. It is clear that variations in people's semantic knowledge and possibly their ability to monitor progress during creative idea generation is a key factor in explaining how DT unfolds.

Acknowledgments

References

Abraham, A. (2014). Creative thinking as orchestrated by semantic processing vs. cognitive control brain networks, *Frontiers In Human Neuroscience, 8,* 1–6.

Abraham, A., Pieritz, K., Thybusch, K., Rutter, B., Kröger, S., Schweckendiek, J., Stark, R., Windman, S., &

Hermann, C. (2012). Creativity and the brain: Uncovering the neural signature of conceptual expansion. *Neuropsychologia*, *50*(8), 1906–1917.

Beaty, R. E., Benedek, M., Kaufman, S. B., & Silvia, P. J. (2015). Default and executive network coupling supports creative idea production. *Scientific Reports*, 1–14. http://doi.org/10.1038/srep10964

Beaty, R. E., & Silvia, P. J. (2012). Why do ideas get more creative across time? An executive interpretation of the serial order effect in divergent thinking tasks. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(4), 309–319.

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*(7), 1186–1197.

Christensen, P. R., Guilford, J. P., & Wilson, R. C. (1957). Relations of creative responses to working time and instructions. *Journal of Experimental Psychology, 53*, 82–88.

Chrysikou, E. G., & Thompson-Schill, S. L. (2011). Dissociable brain states linked to common and creative object use. *Human Brain Mapping, 32*, 665–675.

Feinerer, I., Hornik, K., and Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software, 25*(5): 1-54.

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications.* Cambridge, MA: MIT Press.

Forster, E. A., & Dunbar, K. N. (2009). Creativity evaluation through latent semantic analysis. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31th annual conference of the cognitive science society* (pp. 602–607). Austin, TX: Cognitive Science Society.

Gabora, L., Rosch, E., & Aerts, D. (2008). Toward an Ecological Theory of Concepts. *Ecological Psychology*, *20*(1), 84–116.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology*, *98*, 611–625.

Harbison, J. I., & Haarmann, H. (2014). Automated scoring of originality using semantic representations. In *Proceedings of cogsci 2014.* https://mindmodeling.org/cogsci2014/papers/405/index.html.

Hass, R. W. (in revision). Tracking the dynamics of divergent thinking via semantic distance. *Memory & Cognition.*

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109.

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons, 1–16. http://doi.org/10.3389/fnhum.2014.00407/abstract

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*, 1–14.

Kim, K. H. (2006). Can We Trust Creativity Tests? *Creativity Research Journal*, *18*(1), 3–14.

Kounios, J., & Beeman, M. (2009). The Aha! Moment: The Cognitive Neuroscience of Insight. *Current Directions in Psychological Science*, *18*(4), 210–216.

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*(2), 211–240.

Madore, K. P., Addis, D. R., & Schacter, D. L. (2015). Creativity and Memory: Effects of an Episodic-Specificity Induction on Divergent Thinking. *Psychological Science*, *26*(9), 1461–1468.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220-232.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494. http://doi.org/10.1016/j.jml.2007.11.006

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383.

Prabhakaran, R., Green, A. E., & Gray, J. R. (2013). Thin slices of creativity: Using single-word utterances to assess creative cognition. *Behavior Research Methods* (online first) http://doi.org/10.3758/s13428-013-0401-7

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rehder, B., Schreiner, M. E., Wolfe, M. B. W., Laham, D., Landauer, T. K., & Kintsch, W. (1998). Using latent semantic analysis to assess knowledge: Some technical considerations. *Discourse Processes*, *25*(2-3), 337–354.

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the Remote Associates Test. *Cognition*, *128*(1), 64–75.

Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity–intelligence distinction.* New York: Holt, Rinehart, & Winston.

Ward, T. B. (2008). The role of domain knowledge in creative generation. *Learning and Individual Differences*, *18*(4), 363–366.

Weisberg, R. W. (2006). *Creativity: Understanding innovation in problem solving, science, invention, and the arts.* Hoboken, NJ: John Wiley.

Zabelina, D., Saporta, A., & Beeman, M. (2015). Flexible or leaky attention in creative people? Distinct patterns of attention for different types of creative thinking. *Memory & Cognition*, 1–12. http://doi.org/10.3758/s13421-015-0569-4