Neural Network Model of Semantic Processing in the Remote Associates Test

Ivana Kajić
School of Computing
Plymouth University
Plymouth, Drake Circus, PL4 8AA
United Kingdom
ivana.kajic@plymouth.ac.uk

Thomas Wennekers
School of Computing
Plymouth University
Plymouth, Drake Circus, PL4 8AA
United Kingdom
thomas.wennekers@plymouth.ac.uk

Abstract

The ability to generate novel, unique and useful ideas is an important trait of intelligent behaviour. It is also a virtue of a creative individual in many scientific and artistic domains. In this study we are concerned with the Remote Associates Test (RAT), a task widely used in psychology and neuroscience to study insight and creative problem solving. The RAT is used to assess the ability of an individual to generate novel relationships among familiar words. The test consists of word triplets (e.g. cream, water, skate) and the task is to find a unique word associated with all three words. Here, we aim to identify a basic set of computational mechanisms underlying cognitive processes in the RAT solving. To this end, we propose a multi-layer neural network based on biologically and cognitive realistic mechanisms. The search for a solution in a RAT problem is realised by spreading of activity among word associations in a semantic layer, and the selection of a response by a winner-take-all layer. The model yields human-like performance and distinguishes between easy and difficult RAT problems. The modelling findings are consistent with the existing theories in creativity research, confirming that less stereotypical word associations are important for the good performance on the RAT.

1 Introduction

The adjective "creative" is often attributed to new, valuable and surprising ideas and objects [1]. It is seen as a positive quality of an individual. Various tests and questionnaires have been devised in the attempt of measuring different aspects of creative thinking. Of particular interest to neuroscience and cognitive science have been cognitive abilities such as working memory, sustained attention and cognitive flexibility underlying behaviour assessed with creativity tests [2].

Divergent thinking tests measure the ability to generate new, meaningful and relevant ideas. An example of a task in a divergent thinking test is to think of as many possible practical uses of a certain object, such as a cooking pot or a brick [3]. In contrast to divergent thinking tests, convergent thinking tests have a correct, not necessarily unique, solution. A convergent task commonly used to study insight is the "9 dots" puzzle, where nine dots have been arranged in three rows and the task is to connect them with four straight lines. Another common task is the Remote Associates Test (RAT) used to measure the ability to form novel and meaningful word associations. While these tests are broadly used in creativity research, solving them requires a range of functions [2, 4] which can facilitate or hinder the performance and the process of problem solving. Thus, to understand the assessed process of ideation it is important to identify and understand the underlying cognitive processes.

1.1 The Remote Associates Test

In this study we are concerned with the modelling of cognitive processes underlying solving of the Remote Associates Test (RAT). The RAT has been widely used in cognitive neuroscience and psychology [5–7] to study insight, problem solving and creative thinking since it was conceived in 1962 [8]. It was developed to measure the ability of an individual to form new associations among seemingly unrelated words. The test consists of word triplets and the task is to find a word associated with all three words for every triplet. An example of a RAT problem is a word triplet *cream, water, skate* with a solution *ice*. The performance on the RAT is measured as the number of correctly solved items within a given time limit. Mednick's theory on associative hierarchies [8] suggests that highly creative individuals are able to form rare, uncommon and remote word associations. This is in contrast with associations formed by less creative individuals who respond with fewer, more common and stereotypical associations. In the present study we investigate this idea by analysing how different word pair associations influence the test performance. The simplicity of the paradigm, availability of a battery of tests in various languages and the opportunity to administer the test in neuroimaging studies [4, 5] have made the RAT ubiquitous in the study of associative basis of semantic cognition.

An interesting cognitive aspect of the RAT solving is the semantic search process in which an individual attempts to find a solution to a RAT problem. Insight solutions to a RAT problem occur abruptly, without voluntary control of a person solving the task, and are preceded by unconscious processing [5]. In contrast, analytical solutions are derived gradually as the person attempts to solve a task [9]. Analytical solving usually allows more time to respond, which correlates with better test performance [10, 11]. One of the first proposals of semantic processing in the RAT involving subconscious processing is the spreading activation theory [12]: every cue word activates a subset of words in a semantic network. Overlapping activations of a unit crossing the threshold give rise to a solution. However, this theory does not provide an account of mechanisms which underlie the search process nor does it explain the variety in performance. Better understanding of the analytical RAT solving has been provided by studies analysing human responses when participants solve RAT problems by reporting every word they think of [13, 14]. Common findings confirm that the search process is restricted by all three cues, with one primary cue being used to generate a response. Evidence was provided in favor of a local sequential search process in the RAT [13], meaning that the next guess is chosen based on the previous one.

The RAT has been also investigated in various computational studies [15–18]. Statistical approaches using large corpora of text and natural language processing tools have yielded a performance which is comparable to or even better than the human performance on the test [16, 17]. Mednick's theory on associative hierarchies [8] was scrutinised by identifying relevant properties of a semantic network of an individual [18–20]. Semantic networks of individuals scoring well on the RAT satisfy small-world network properties, where every word in the network can be reached in a few steps by following associated words of any other word. Qualitative differences in associative networks between individuals with low and high creative abilities have been described in [19]. A modelling study investigating the generation of ideas in spontaneous thought in a neural network [20] confirmed that scale-free and small-world network properties are important for efficient search in the memory generating conventional and creative ideas. A Markov Chain Monte Carlo model [15] successfully reproduced experimentally derived human response patterns [13] supporting findings on the local search strategy. While providing valuable insights into the understanding of memory search, these models offers a limited explanation of the cognitive mechanisms involved in the semantic processing in the RAT and their relation to the test performance.

Summarising the existing work, the evidence suggests that both the executive cognitive component guiding the search process and the organisation of the semantic network are important for the semantic processing in the RAT. To identify a minimal set of computational primitives underlying these processes, we propose a neural network in which every layer realises a different function interpreted in the theoretical framework of semantic processing

2 The Model

We propose a neural network model of the RAT solving based on the spreading activation [12] and a winner-take-all (WTA) mechanism. Solving of a RAT problem is realised as a search for a

Algorithm 1 Search by spreading activation and a winner-take-all function

```
nr_visited ← 0
visited = [ ]
for i ← 1, N do
    activations_i ← 0
for all c in cues do
    I_c ← 1
j ← −1
while nr_visited < max_nodes do
    for all neighbours i of the current node j do
        if w_ij > ϑ_s then
            activations_i ← activations_i + w_ij activations_j
    activations_i ← activations_i + I_i
    visited ← visited + j
    nr_visited ← nr_visited + 1
    if j == target then
        break
    j ← WTA(activations)
```

solution in a semantic network. To better understand the workings of the neural network, we first present the search algorithm which is then implemented by the network. The equations of the neural network model are presented together with an explanation of how algorithm components map to computations performed by neuron-like units.

2.1 Search Algorithm

The search for a solution is done separately for every RAT problem consisting of three cue words and a target solution. For every problem, the word cues are used to initialise the search process. In every step of the search process, a word is selected as a candidate solution to a problem. Every word is represented as a node in a graph and every node has a level of activity which changes as the search unfolds. Thus, the search for a solution to a RAT problem is modelled as a graph traversal and the direction of the search is determined by the levels of node activities. If the selected node matches the target word, or a certain number of words in the search has been exhausted, the search process terminates.[1] The activity of neighbours of a currently visited node is elevated by the amount proportional to the connection strength between the node and its neighbours. Connection strengths for $N = 5018$ words are derived from a freely available database of word association strengths acquired in a free association task [21]. In the task, the participants were instructed to write a single word which first came to their mind when prompted with a cue word. To obtain the associative strength between word pairs, the number of participants who responded with a specific word to a cue was divided by the number of participants performing the same task. Only responses which were given by at least two participants were considered.

After elevating the activities of neighbouring nodes, a new node will be selected by the WTA function and the old node will be annotated as visited. In the algorithm, the WTA function sorts node indices based on their activities and returns the index of a node with the highest activity level which has not been visited before. In this way, words which have already been considered as a solution will not appear again in the search process. If two nodes have the same level of activity, the first node returned by the sort algorithm will be selected. The search path is created by appending every visited node to the list of previously visited nodes.

In the beginning, the activity of all nodes is set to zero. Every node receives an external input, which is set to one only for the nodes representing the problem cues, and zero for all other nodes. Based on high activity levels, the WTA sequentially selects three problem cues and the activity spreads through the network from the problem cues first. Only after all three cues have been visited, the

[1] Here we do not model the process of determining whether a response is a solution to the RAT problem. Instead, we focus on the process of ideation of possible words, assuming when an individual comes up with a word they will be able to determine if it is a correct solution by comparing it against the task constraints.

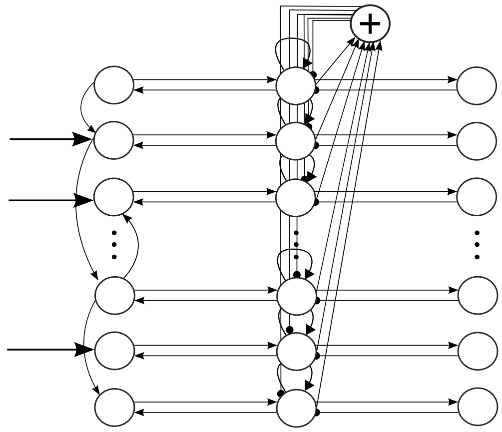

Figure 1: Neural network model of the semantic search in the RAT. The word with the highest activity level in the semantic layer is selected by the WTA layer as a response to a RAT problem. If the selected word does not match the solution the inhibitory layer suppresses the activity of the selected word allowing the WTA to select the next winner.

WTA will select a new node as a candidate solution. The process terminates when the selected response matches the target or when a certain number of nodes has been visited. The pseudocode for the search algorithm is shown in Algorithm 1. In every step the WTA function selects a node j, which is appended to the list of visited nodes after it has been processed. In the first iteration, there is no spreading of the activity from any node as there is no winning node ($j = -1$). This occurs in the second step when the WTA has selected the first problem cue receiving the external input. This condition is handled in the code and omitted here for clarity.

To explore the influence of different word pair association strengths on the RAT performance, we implement spreading of activity only to those neighbours whose connection strength to the processing node is greater than the spreading threshold ϑ_s.[2] Weak connection strengths stand for rare and uncommon associations, which according to Mednick's theory [8] are more likely to be generated by highly creative individuals. Increasing the threshold corresponds to removal of such association pairs, resulting in longer number of steps between two nodes or, for very high thresholds, inability to reach a node.

2.2 Neural Network

A three-layer neural network model is used to simulate solving of RAT problems using the search algorithm described in Algorithm 1. The three layers are the semantic layer, winner-take-all (WTA) layer and inhibitory layer. All three layers have distinct functional roles. The model scheme is shown in Figure 1. The semantic layer represents a semantic network, consisting of a vocabulary and the associative relationships between the words in the vocabulary. The WTA layer selects a winning unit simulating a solution guess to a RAT problem. The spread of activity from a winning unit in the semantic layer is done via feedback connections from the WTA layer to the semantic layer. If the selected word is not a solution to the RAT problem, the inhibitory layer suppresses the activity of the winning unit in the WTA layer. This allows the next unit with the highest activity level to win in the next step. The activity of units in all layers is updated in parallel in each time step.

The activity of units in the first layer can be written as:

[2] This is equivalent to rectifying all $w_{ij} < \vartheta_s$ to zero.

$$a_i(t+1) = a_i(t) + \rho_a \Bigg(\sum_{j=1}^{N} W_{ij} z_j(t) a_j(t) + I_i(t) \Bigg) \tag{1}$$

$$z_i(t) = \Theta(w_i(t) - \vartheta_w) \tag{2}$$

$$\|z(t)\| = 1 \tag{3}$$

where $a_i(t)$ describes a non-decaying activity of a unit i at time t where $i = 1, \ldots, N$. W_{ij} is the connection strength from the unit j to the unit i, extracted from the database of free associations [21]. The constant ρ_a is inversely proportional to the stimulus length t_n and allows the analytical derivation of precise values of activity levels (not shown). The binary vector $z(t)$ always has exactly one or zero active elements, with the active element representing the currently visited node (variable j in Algorithm 1) in the semantic layer. The winner is a word considered as a solution to the RAT problem. Consistent with the algorithm, only the activity of nodes whose edges incident to the winning node are greater than the spreading threshold ϑ_s, or the activity of nodes receiving external input, will be elevated. Thresholding is done by setting the weights smaller than ϑ_s in the connection matrix W to zero. The winning node is set based on unit activities in the WTA layer. If a unit in the WTA layer crosses the threshold ϑ_w, the Heaviside step function will toggle the corresponding bit in $z(t)$. Only at the beginning of a simulation, the external input $I(t)$ is used to sequentially increase the activity of cue nodes in the semantic layer. This in turn will elevate the activity of the units in the WTA layer yielding a winner whose feedback connection will activate the spread of activity from a winning unit in the semantic layer. The first three winners are thus going to be the three problem cues. $I_i(t)$ is clamped to one for the duration t_n starting at non-overlapping times t_i for nodes representing problem cues: $t = t_i : I_{cue_i}(t_i) = 1$ where cue_i is the index of a problem cue $i \in \{1, 2, 3\}$. The activity of the units in the WTA layer is described as:

$$w_i(t+1) = w_i(t) + \rho_w \left[c_1 z_i(t) + c_2 \widetilde{a}_i(t) - c_3 y(t) - c_4 r(t) + c_5 \eta_i(t) \right] \tag{4}$$

$$y(t+1) = \sum_{j=1}^{N} z_j(t) \tag{5}$$

Units in this layer receive one-to-one feedforward input $\widetilde{a}_i$ from the units in the first layer. The first WTA unit to cross the threshold ϑ_w will be the one receiving the strongest normalised input $\widetilde{a}_i(t)$.[3]. It will activate the self-excitatory connection of a strength c_1, and the input to the single inhibitory neuron $y(t)$ (a unit with the '+' sign in Figure 1). The inhibitory neuron will project feedback connections to all units in the layer, such that the activity of all the units apart from the winning one will be suppressed and kept below the threshold ϑ_w. The winning neuron will also toggle the corresponding bit in the $z(t)$ as described in equation 2 and elevate activities of its neighbours in the semantic layer. Due to the self-excitation, the activity of a winning unit in the WTA layer will continue to increase until suppressed by a unit in the third layer. The noise term $\eta_i(t)$ randomly drawn from a uniform distribution is added to allow the WTA to select a winning unit when two or more units receive the same input $\widetilde{a}(t)$ from the semantic layer. Finally, after a unit $z_i(t)$ has been active for a certain amount of time t_n it will activate the corresponding inhibitory unit in the third layer:

$$r_i(t+1) = r_i(t) + \Theta\Bigg(\sum_{j=t-t_n}^{t} z_i(t_j) - t_n \Bigg) \tag{6}$$

The feedback connections to the WTA layer will inhibit the activity of the winning unit. As input integration in $r(t)$ is non-decaying, the inhibition will be permanent preventing the winning unit from winning again. A new unit will be selected in the semantic layer based on the highest level of activity $\widetilde{a}_i(t)$.

[3] $\widetilde{a}_i(t) = \frac{a_i(t)}{a_k(t)}$ where $k = \arg\max_j a_j(t)$. The activity has been normalised to bound the input range for the units in the WTA layer. Knowing the input range it is possible to analytically derive the parameter values for which the precision of the competition mechanism can be controlled.

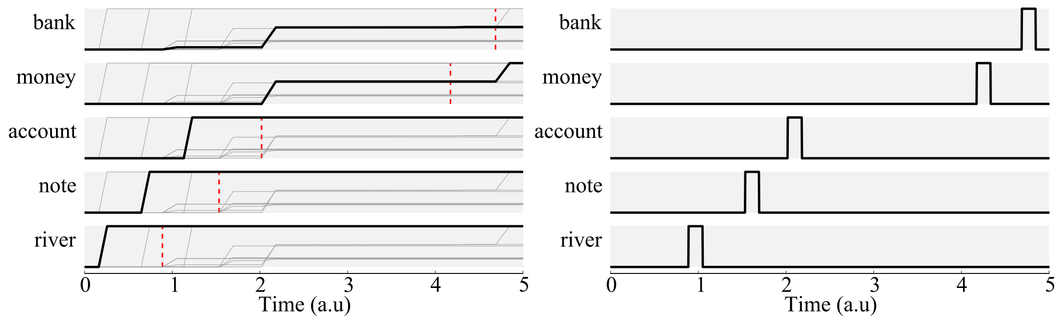

Figure 2: Normalised unit activities in the semantic layer (left) and the winning units (right) for one simulation of a RAT problem with the problem cues: *river, note, account*. The solution *bank* is found as a second response, after *money*. Because the first three winners are the problem cues they are not considered as potential solutions. Red dashed lines in the left plot represent a moment when a winning unit has been selected and determines the onset of spreading activity to its neighbouring units.

Figure 2 shows unit activities in the network over a course of time while solving a RAT problem: *river, note, account*. First three winners in the network are the cues *river, note* and *account* and therefore not considered as solutions. The first response *money* is inhibited as it does not match the solution to this RAT problem. The correct response *bank* is chosen next and the simulation of this RAT problem is terminated. For visual clarity, only a small fraction of activated units in the semantic layer is shown.

3 Results

To test the model performance we use 117 out of 144 RAT items [10] for which the cues and the target are available in the free association database [21]. To obtain problem difficulties, we divide 117 problems into three categories (easy, medium and hard) based on the percentage of participants solving an item in the 15 seconds condition [10]. The percentage of participants solving a problem varies between 0% and 96% and we divide the categories in three equal parts: easy problems are solved by 64%–96% participants (17 problems), medium by 32%–64% (43 problems) and hard problems by 0%–32% participants (58 problems).

The performance on all three categories is tested by varying two model parameters: the number of responses and the spreading threshold. The number of responses is the number of words in the search path, starting from the first word selected by the WTA. If the correct response was not among that predetermined number of responses, the problem is annotated as unsolved. As a first approximation, the number of responses can be related to the time allowed to produce the answer. The basic assumption is that with more time a participant will be able to think of more words. The second parameter is the spreading threshold ϑ_s in the semantic layer. By increasing the threshold we are discarding all word pair associations with the association strength weaker than the threshold. Lower association strengths correspond to word association pairs given by fewer individuals in the free association task [21]. Thus, increasing the thresholding reduces the number of uncommon and rare cue-target associations. Such associations, interpreted in the context of Mednick's theory on associative hierarchies [8], are more likely to be produced by highly creative individuals, as the associations of low creative individuals are characterised by stereotypy and commonality. We explore the relationship between the importance of such associations and the performance on the RAT.

Network simulation results for 117 problems and the three difficulty categories are shown in Figure 3. As expected, the performance on the RAT increases with the number of responses. Compared to the medium and hard items, all easy items are solved with fewer responses (10 responses). Approximately 20% of easy problems are solved with a single response, implying that the solution to easy RAT items is a close association between one of the problem cues and the solution. Close associates are strong word associations, in this case, the word pairs with high association strengths in the free

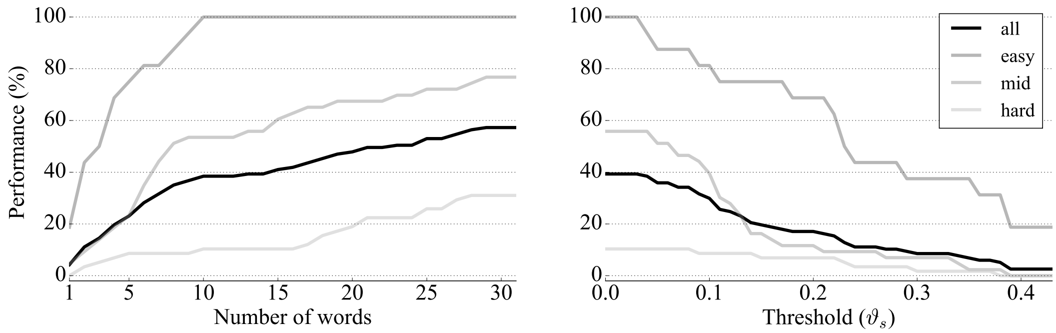

Figure 3: Performance on the easy, medium and difficult RAT items depending on the number of responses in the search process (left) and the spreading threshold (right). Increasing the threshold removes word pair associations with association strengths weaker than the threshold.

association database [21]. Such close associations emerge because many participants in the free association task have responded with the same word to a cue word. A continuous increase in the performance for medium and hard problems is observed when 15 or more responses are considered. For the purposes of current analysis, we have restricted the range of responses to an interval which could be interpreted in the context of known data. Participants instructed to report every word they consider as a solution when trying to solve a RAT problem on average produce eight words within two minutes [13], although there are large variations in the number of responses. As we are not explicitly modelling this process, we take this number as a reference. Therefore, we assume this number to be greater in the model which does not model a specific cognitive strategy that would differentiate between reported and unreported words. The average percentage of solved items for 117 tested problems for humans in 15 sec condition was 28.8%. With six responses the model yields similar performance (28.2%) on the same set of problems.

Right plot in Figure 3 shows how removing word associations has different effects on the test performance depending on the problem difficulty. The weight value of 0.4 is the 98th percentile of all non-zero association strengths in the free association database [21]. Overall, removing word associations impairs the performance on the test for all problem difficulties. However, easier and difficult problems are affected differently. Relative to the performance on the RAT without pruning of associations ($\vartheta_s = 0$), the drop in performance by 50% occurs at lower threshold values for problems of medium difficulty compared to easy problems. For easy items, 50% decrease in performance occurs when all word pairs of association strength $\vartheta_s = 0.23$ (94th percentile) or lower are removed, while for the items of medium difficulty this already occurs for the threshold value of $\vartheta_s = 0.12$ (87th percentile). This effect is also observed for smaller drops in performance when comparing the performance on the RAT problems of medium and hard difficulty with easy RAT problems, and when using a different number of words (here only shown for 15 words). This indicates that less common word pair associations are important for the better performance on the difficult RAT problems.

4 Discussion

With this work we have aimed to identify a basic set of computational mechanisms underlying the semantic processing in the RAT solving. This has been done by devising a neural network model in a way consistent with the existing understanding of human semantic processing. Semantic layer in the model implements a localist representation of lexical knowledge. Memory search in the semantic layer is realised by spreading activation and a WTA network, both cognitively realistic mechanisms. The spreading of activity has been specified in the context of associative hierarchies relevant for characterising creative abilities of an individual [8]. It is assumed to occur at the subconscious level of semantic processing [12, 22]. The WTA mechanism and the inhibitory layer responsible for selecting a single word in the vocabulary are reminiscent of attentional mechanisms mediating cognitive control attributed to the function of the frontal brain areas [4, 5]. The focus of attention is directed towards a single word which is selected among several competing alternatives. Anterior cingulate cortex (ACC) has been shown to play an important role in attentional switches and conflict resolution in case of competing alternatives [23, 24].

The model is able to distinguish between easy and difficult RAT problems in the normative RAT data set [10]: easy items are solved within fewer response attempts and, compared to the more difficult items, are less affected by the removal of unusual word associations. This is in accordance with Mednick's theory on associative hierarchies [8] according to which creative individuals are more likely to produce less stereotypical and uncommon word associations. Thus, lower threshold values in the model might be related to the ability of an individual to consider such associations. Alternatively, and not exclusively, this could be a property of the organisation of an individual's semantic network. The semantic network constructed from a free association database [21] satisfies small-world properties with power-law degree distribution [25] important for a good performance on the RAT [19]. Therefore, semantic networks of individuals scoring lower on the RAT might lack such associations, resulting in compromised small-world network organisation. When interpreted in the context of number of words needed to produce a correct response to a RAT problem, more difficult items have a solution which is more distant from the problem cues in the search path. This would justify the difficult RAT problems as having more "remote" associations.

While the proposed model is based on theories of semantic processing and biologically realistic mechanisms, a detailed theory of different cognitive strategies in the RAT is needed to model the differences in the underlying processes. Different hemispheric contributions in semantic processing have been observed when people solve the RAT by insight and analytically [5]. It is to expect that analytic solving requires greater engagement of semantic and working memory, reasoning and other systems involved in problem solving. Therefore, different and possibly overlapping brain networks would require a more comprehensive, large-scale brain model, simulating interactions among several brain regions realising different cognitive functions. With this work, instead of capturing the processing spanning several brain regions, we have focused on a minimal set of basic neurocomputational mechanisms independently of a cognitive strategy. Future work will address the extension and expansion of the model. One advantage of a large-scale model would be a biologically realistic, distributed representation of sensory information which can better capture associative nature of word representations at the neural level. It remains to be explored how a higher level of biological realism can inform our understanding of associative knowledge and creative problem solving.

5 Conclusion

We have developed and presented a neural network model of semantic processing in the Remote Associates Test, a commonly used paradigm in the research of insight and creative problem solving. The model is based on the theory of spreading activation in semantic processing, and uses neurally realistic computations. We have shown that the model exhibits human like performance on the task, demonstrating that uncommon and less stereotypical associations are important for a good performance on difficult test problems. Finally, the Python scripts used for the processing of the free association data and the complete source code are available online at `http://github.com/ikajic/remote_associates_test`.

Acknowledgments

The authors would like to thank Terry Stewart, Vaibhav Tyagi, and Michael Klein for discussions that helped to improve this paper. They would also like to thank Haline Schendan and Giorgio Ganis for their valuable comments and insights throughout the project. This work has been supported by the Marie Curie Initial Training Network FP7-PEOPLE-2013-ITN, CogNovo, grant number 604764.

References

[1] Margaret A. Boden. *The Creative Mind: Myths and Mechanisms*. Routledge, 2 edition, November 2003.

[2] Dietrich Arne. The Cognitive Neuroscience of Creativity. *Psychonomic Bulletin & Review*, 11(6):1011–1026, 2004.

[3] E. Paul Torrance. *Guiding creative talent*. Prentice-Hall, 1962.

[4] Andreas Fink, Mathias Benedek, Roland H. Grabner, Beate Staudt, and Aljoscha C. Neubauer. Creativity meets neuroscience: Experimental tasks for the neuroscientific study of creative thinking. *Methods*, 42(1):68–76, 2007.

[5] John Kounious and Mark Beeman. The cognitive neuroscience of insight. *Annual Review of Psychology*, 65:71–93, 2014.

[6] Sebastien Helie and Ron Sun. Incubation, insight, and creative problem solving: a unified theory and a connectionist model. *Psychological Review*, 117(3):994–1024, 2010.

[7] Mark Jung-Beeman, Edward M. Bowden, Jason Haberman, Jennifer L. Frymiare, Stella Arambel-Liu, Richard Greenblatt, Paul J. Reber, and John Kounios. Neural Activity When People Solve Verbal Problems with Insight. *PLoS Biology*, 2(4):e97, 04 2004.

[8] Sarnoff A. Mednick. The associative basis of the creative process. *Psychological Review*, 69(3):220–232, 1962.

[9] Roderick W. Smith and John Kounios. Sudden insight: All-or-none processing revealed by speed–accuracy decomposition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6):1443, 1996.

[10] Edward M. Bowden and Mark Jung-Beeman. Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, 35(4):634–639, 2003.

[11] Stephen G. Harkins. Mere effort as the mediator of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, 91(3):436–455, 2006.

[12] Allan M. Collins and Elizabeth F. Loftus. A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407 – 428, 1975.

[13] Kevin A. Smith, David E. Huber, and E Vul. Multiply-constrained semantic search in the remote associates test. *Cognition*, 128:64–75, 2013.

[14] Eddy J. Davelaar. Semantic search in the remote associates test. *Topics in Cognitive Science*, 7(3):494–512, 2015.

[15] David D. Bourgin, Joshua T. Abbott, Thomas L. Griffiths, Kevin A. Smith, and Edward Vul. Empirical Evidence for Markov Chain Monte Carlo in Memory Search. In *Annual Conference of the Cognitive Science Society*, 2014.

[16] Ariel Klein and Toni Badia. The Usual and the Unusual: Solving Remote Associates Test Tasks Using Simple Statistical Natural Language Processing Based on Language Use. *The Journal of Creative Behavior*, 49(1):13–37, 2015.

[17] Hannu Toivonen, Oskar Gross, Jukka M. Toivanen, and Alessandro Valitutti. On Creative Uses of Word Associations. In *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, volume 190 of *Advances in Intelligent Systems and Computing*, pages 17–24. Springer Berlin Heidelberg, 2013.

[18] Padraic Monaghan, Tom Ormerod, and Ut N. Sio. Interactive activation networks for modelling problem solving. In *Computational models of cognitive processes: Proceedings of the 13th Neural Computation and Psychology Workshop*, volume 21, pages 185–195, 2014.

[19] Yoed N. Kenett, David Anaki, and Miriam Faust. Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 8(407), 2014.

[20] Nagendra Marupaka, Laxmi R. Iyer, and Ali A. Minai. Connectivity and thought: the influence of semantic network structure in a neurodynamical model of thinking. *Neural Networks*, 32:147–158, 2012.

[21] Douglas L. Nelson, Cathy L. McEvoy, and Thomas A. Schreiber. The University of South Florida Free Association, Rhyme, and Word Fragment Norms. *Behavior Research Methods, Instruments, & Computers*, 36(3):402–407, 2004.

[22] Ilan Yaniv and David E. Meyer. Activation and metacognition of inaccessible stored information: potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2):187, 1987.

[23] Matthew M. Botvinick, Jonathan D. Cohen, and Cameron S. Carter. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12):539 – 546, 2004.

[24] John G. Kerns, Jonathan D. Cohen, Angus W. MacDonald, Raymond Y. Cho, V. Andrew Stenger, and Cameron S. Carter. Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science*, 303(5660):1023–1026, 2004.

[25] Mark Steyvers and Joshua B Tenenbaum. The Large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, 29(1):41–78, 2005.