Simulating Incubation Effects Using the Explicit Implicit Interaction with Bayes Factor (EII-BF) Model

Sebastien Helie, Ron Sun

***Abstract*—Most psychological theories of problem solving have focused on modeling explicit processes that gradually bring the solver closer to the solution in a mostly explicit and deliberative way. This approach to problem solving is typically inefficient when the problem is too complex, ill-understood, or ambiguous. In such a case, a 'creative' approach to problem solving might be more appropriate. In the present paper, we propose a neural-network-based computational model implementing the Explicit-Implicit Interaction theory of creative problem solving that involves alternating between implicit and explicit processing. In the present, the new model is used to simulate the incubation process in lexical decision and free recall.**

I. Introduction

According to Wallas, *a single achievement of thought* is a four stage process [1]. The first stage, preparation, is the accumulation of knowledge that allows directly solving the problem at hand (i.e., in a sequential, logical way). However, this Cartesian approach to problem solving sometimes fails to provide a satisfying solution and an impasse is reached. Surprisingly, if the problem solver puts the problem aside and directs his/her attention to some other task, s/he is more likely to later find the correct solution [2]-[3]. This phenomenon, which has been the subject of countless anecdotes in the history of science, is thought to result from Wallas' second stage: incubation. Because the solution often comes as a surprise to the solver, the stage following the incubation period was called illumination (or insight), during which the problem solver has the impression that the solution completely elucidates the problem. However, the quality of the solution still remains to be evaluated by the fourth stage: the validation process.

The focus of this paper is Wallas' second stage of creative problem solving, namely incubation. In the present, we propose a new model of creative problem solving based on the *Explicit-Implicit Interaction* theory [4]-[5] and focus on the explanation and simulation of the incubation stage.

Manuscript received March 10, 2009. This research was supported by a postdoctoral research fellowship from *Le Fonds Québecois de la Recherche sur la Nature et les Technologies* granted to the first author and research grants provided by the *Army Research Institute*, contracts DASW01-00-K-0012 and W74V8H-04-K-0002, granted to the second author.

S. Helie is with the Psychology Department, University of California, Santa Barbara, CA 93106-9660 USA (phone: 805-893-7909; fax: 805-893-4303; e-mail: helie@psych.ucsb.edu).

R. Sun is with the Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY 12180-3590 USA (e-mail: rsun@rpi.edu).

II. The Explicit-Implicit Interaction Theory

The Explicit-Implicit Interaction theory [4]-[5] is a psychological theory that relies mainly on the following principles: (1) there are two types of knowledge, namely explicit and implicit (a well established distinction in psychology: see, e.g., [6]-[7]), that are simultaneously involved in most tasks [8]. (2) Explicit and implicit knowledge are often "redundant" [9], and the results of explicit and implicit processing are often integrated to provide a response [8]. In psychology, implicit knowledge refers to knowledge that affects behavior without the awareness of the participant while explicit knowledge may (potentially) come with the accompanying awareness [6]. (For details and justifications of the principles, see [5], [8].)

In the Explicit-Implicit Interaction theory, the preparation and verification stages of problem solving are mostly captured by explicit processing, while implicit processing corresponds mostly to incubation. Wallas' third stage, insight, involves an "explicitation" process and consequently produces a solution in an explicit form. In addition, the Explicit-Implicit Interaction theory assumes the presence of activation thresholds that control the *Internal Confidence Level* (ICL) necessary to come up with a solution to the problem and a negative relationship between the ICL and the reaction time. If a threshold is not crossed after the results of explicit and implicit processing have been integrated, another iteration of processing is performed. This simple yet powerful iterative theory has been formalized into a connectionist model that was used to simulate a variety of tasks. The model, Explicit-Implicit Interaction with Bayes Factor (EII-BF), is described in the following section.

III. The EII-BF Model

The general structure of the model is shown in Fig. 1. The model is composed of two major modules, representing explicit and implicit knowledge respectively. These two modules are connected by using bidirectional associative memories (i.e., the **E** and **F** weight matrices; [10]). In each trial, the task is simultaneously processed in both modules, and the outputs (response activations) are integrated in order to determine a response distribution. Once this distribution is specified, a response is stochastically chosen and the Bayes factor [11] is computed using the two most likely responses. If this measure is higher than a predefined threshold, the chosen response is output and the reaction time is computed; otherwise, another iteration of processing is done in both modules, using the chosen response as the input.

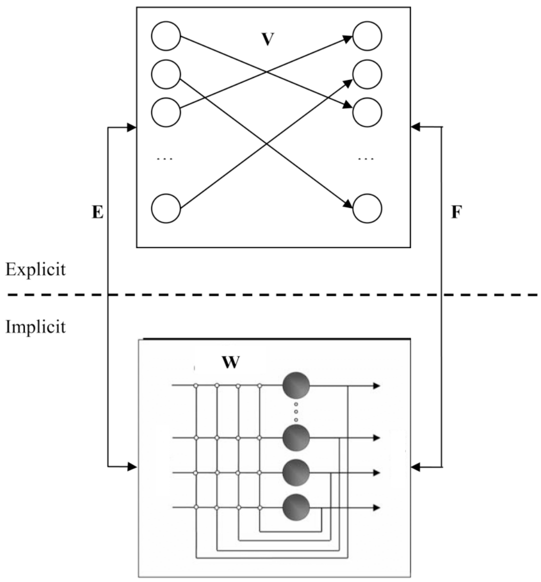

Fig. 1. General architecture of the EII-BF model.

In EII-BF, explicit processing is captured using a two-layer linear connectionist network while implicit processing is captured using a non-linear attractor neural network (*NDRAM*: [12]). The inaccessible nature of implicit knowledge may be captured by distributed representations in an attractor neural network, because units in a distributed representation are capable of accomplishing tasks but are less individually meaningful [6]. This characteristic corresponds well with the relative inaccessibility of implicit knowledge [8]. In contrast, explicit knowledge may be captured in computational modeling by localist representations, because each unit in a localist representation is more easily interpretable and has a clearer conceptual meaning. This characteristic captures the property of explicit knowledge being more accessible and manipulable [8]. This difference in the representation of the two types of knowledge leads to a dual-representation, dual-process model. Theoretical arguments for such models are presented in [6], [8], [13].[1]

Specifically, explicit knowledge is localistically represented in the top level using binary activation. The left layer in Fig. 1 (denoted **x**) is composed of n units while the right layer (denoted **y**) is composed of m units. These layers are connected using the binary weight matrix **V**, and the information is transmitted using the standard dot product (i.e., $\mathbf{y} = \mathbf{V}\mathbf{x}$).[2]

In the bottom level, implicit knowledge is represented using r bipolar units (denoted **z**). Precisely, $\mathbf{z} = \mathbf{t}_1 + \mathbf{t}_2$, where $\mathbf{t}_1$ represents the first s units in **z**, which are connected to the left layer in the top level using the **E** weight matrix. Meanwhile, $\mathbf{t}_2$ represents the remaining $r - s$ units in **z**, which are connected to the right layer in the top level using weight matrix **F**. In words, the **E** and **F** weight matrices are used to 'translate' explicit knowledge into implicit knowledge (i.e., 'implicitation') and vice-versa (i.e., 'explicitation').

Bottom-level activation (**z**) is modified through a settling process using the NDRAM transmission rule [12]:

$$\mathbf{z}_{[t+1]} = f(\mathbf{W}\mathbf{z}_{[t]}), \qquad f(z_i) = \begin{cases} +1 & , z_i > 1 \\ (\delta+1)z_i - \delta z_i^3, & -1 \le z_i \le 1 \\ -1 & , z_i < -1 \end{cases} \tag{1}$$

where $\mathbf{z}_{[t]} = \{z_1, z_2, \ldots, z_r\}$ is the bottom-level activation after t iterations in the network, **W** is the bottom-level weight matrix, and $0 < \delta < 0.5$ is the slope of the transmission function. This settling process amounts to a search through a soft constraint satisfaction process, where each connection represents a constraint and the weights represent the importance of the constraints [15]. Note that it was estimated psychologically that each iteration in this network takes roughly 350 ms of psychological time (for justifications, see [4]).

Once the response activations have been computed in both levels, they are integrated using the *Max* function:

$$o_i = Max\left[y_i,\ \lambda \sum_{j=1}^{r} f_{ji} z_j \right] \tag{2}$$

where $\mathbf{o} = \{o_1, o_2, \ldots, o_m\}$ is the integrated response activation, $\mathbf{y} = \{y_1, y_2, \ldots, y_m\}$ is the result of top-level processing, λ is a scaling parameter, and $\mathbf{F} = [f_{ij}]$ is a weight matrix. The integrated response activation is then transformed into the response distribution:

$$P(o_i) = e^{o_i/\alpha} \left(\sum_j e^{o_j/\alpha} \right)^{-1} \tag{3}$$

where α is a noise parameter (i.e., the temperature). Note that low noise levels tend to exaggerate the probability differences, which lead to a narrow search of possible responses and favors stereotypical responses. In contrast, high noise levels tend to minimize the probability differences, which leads to a more complete search of the response space.

A response is stochastically chosen and the ratio of the probabilities of the two most likely responses is computed (i.e., the Bayes factor). This measure represents the relative support for the most likely response. According to [11], Bayes factors smaller than 3.2 are not worth mentioning, while Bayes factors higher than 10 are considered strong evidence (and Bayes factors higher than 100 are *decisive*). In the current model, the chosen response is output if the Bayes factor is higher than a free parameter (ψ; by default, ψ = 100), and the reaction time of the model is a negative linear function of the Bayes factor. However, if the Bayes factor is

[1] Note that EII-BF is an algorithmic model [14]. Hence, the units do not represent biological neurons but are a constructive proof of the network/representation properties.

[2] In EII-BF, all the weight matrices are learned using Hebbian learning. This type of learning has the advantage of psychological and biological plausibility. The **V**, **E**, and **F** weight matrices were trained using regular Hebbian learning (i.e., the outer matrix product). The bottom-level weight matrix (**W**) was trained using a contrastive Hebbian learning rule [12].

smaller than ψ, the search process continues with a new iteration using the chosen response to activate the left layer ($\mathbf{x} = \mathbf{V}^{\mathrm{T}}\mathbf{o}$; $\mathbf{z} = \mathbf{E}\mathbf{x}$).

The algorithm specifying the complete process is summarized in Table I. Note that in all the simulations, the parameters were set to reasonable values to qualitatively reproduce the human data. No optimization technique was used in this paper.

TABLE I: ALGORITHM OF THE EII-BF MODEL

1. Observe the current state of the environment;
2. Compute the response activations;
3. Compute the integrated response activation and the resulting response distribution;
4. Stochastically choose a response and compute the Bayes factor:
   a. If the Bayes factor is higher than ψ, output the response;
   b. Compute the reaction time.
5. Else, if there is time remaining, go back to step 2.

IV. INCUBATION IN A LEXICAL DECISION TASK

[3] used a rare-word association task and a lexical decision task to test the unconscious work theory of incubation [2], [16]. These tasks are detailed below.

A. Experiment

In [3], each trial was initiated by the presentation of a definition to the participant, who had fifteen seconds to find the associated word (the rare-word association task). If the participant was able to produce the associated word, the lexical decision task started. If the participant did not find the associated word, s/he was asked to estimate his/her *Feeling Of Knowing* (FOK) prior to starting the lexical decision task. A flow chart describing a trial is shown in Fig. 2. A concrete example of a trial can be found in the *Appendix*.

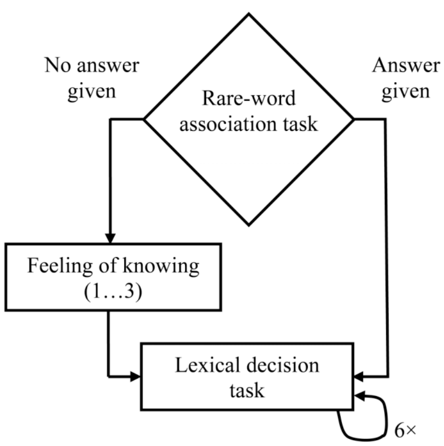

Fig. 2. Flow chart representing a trial in [3].

In the lexical decision task, the participant's job was to identify strings of letters as 'words' or 'non-words'. Each rare-word association trial was followed by a block of six lexical decision trials. The block was composed of three types of strings: unrelated words (distractors), non-words, and the response to the rare-word association trial (the target word). The prediction was that the participants who were unable to provide an answer in the rare-word association task but had a high FOK would be primed for the target word in the lexical decision task (leading to faster response times), but not those who had a low FOK. Likewise, correct responses in the rare-word association task would prime the target but incorrect responses would not. If these were the cases, one might interpret these results as an indication that incubation is a form of "unconscious" processing and that incomplete ("unconscious") processing might be sufficient to prime a target word (despite the failure to produce the target word) [5].

B. Empirical Results

44 participants were tested in 52 rare-word association trials and associated $52 \times 6 = 312$ lexical decision trials. The results of interest were those obtained in the lexical decision task, factorized by the performance in the rare-word association task.

As predicted, correct responses in the rare-word association task primed the target word in the lexical decision task $t(2100) = 8.5$, $p < .01$. In contrast, incorrect responses in the rare-word association trial did not affect the performance of the subsequent lexical decision task (i.e., no priming for the target word; $t(2100) = 0.7$, $p > .2$). In trials in which no response was given in the preceding rare-word association task, analyses of the response times showed a significant interaction between the FOK and the type of stimuli (targets vs. distractors; $t(1648) = 2.28$, $p < .05$). Gamma correlation coefficients (provided in [3]) suggested that targets were faster than distractors when the FOK was high; this relation was reversed when FOK was low (see Fig. 3).

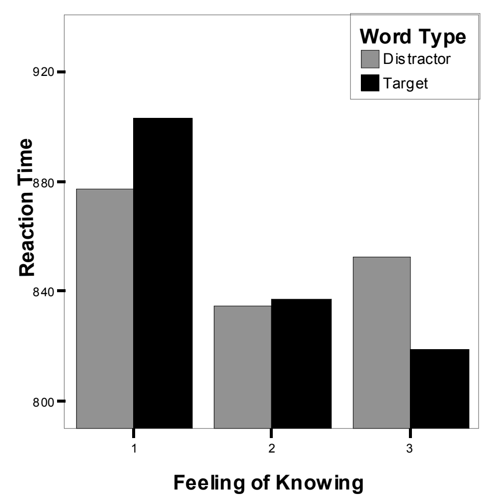

Fig. 3. Human response times in the lexical decision task when no answer was given in the rare-word association task (from [3]).

C. Simulation

In the top level of EII-BF, the left layer was used to represent the words while the right layer represented the definitions (see Fig. 1). Each word and each definition was represented by a different node (i.e., using localist representations; $n = m = 52$), and each word was associated to its definition by a link within the top level. In the bottom level of EII-BF, half of the nodes were used to represent the words while the remaining was used to represent the definitions (both with distributed representations; $r = 200$, $s = 100$). Each word/definition was represented by randomly generated activation patterns in the bottom level. The bottom-level network was pre-trained to encode the associations between the words and their corresponding definitions ($\eta = 0.001$, $\zeta = 0.9999$, $p = 3$, $\delta = 0.4$; for an explanation of the parameters, see [12]).

To simulate a rare-word association trial, a stimulus activated the right (definition) layer in the top level and the corresponding representation in the bottom level.[3] Explicit rules were applied in the top level (in this case, amounting to retrieving definition-word associations), and the information in the bottom level was processed for 42 spins (with roughly 350 ms per spin as hypothesized earlier), approximating the fact that human participants had 15 s. Following this processing, the outputs from both levels were integrated and transformed into a Boltzmann distribution (which served as the activations of the nodes in the left layer of the top level; $\lambda = 1.05$ and $\alpha = 0.15$). The Bayes factor of the distribution was used to estimate the Internal Confidence Level. Because there was no time for further iterations, a response was output if the ICL was higher than $\psi = 150$. Otherwise, no response was provided and the FOK was estimated using the ICL (as in the human experiment). If the ICL was higher than 110, the FOK was estimated as 'high' and if the ICL was lower than 90, the FOK was estimated as 'low'. The remaining range was rated as 'medium'.

For simulating the lexical decision task, three types of stimuli had to be represented. The target was the same word used in the corresponding rare-word association trial, and the distractors were the words used in other trials of the rare-word association task. Non-words used randomly generated representations (real values within [0, 1]). Note that words (either distractors or targets) were represented explicitly in the top level (in the left layer) whereas non-words were not.

Following each rare-word association trial, six lexical decision trials were conducted. Because the stimuli were presented rapidly, a normally distributed noise pattern (a noise vector) was added to each stimulus ($\mu = 0$, $\sigma = 0.25$). The information was transmitted within EII-BF as follow. First, a stimulus activated a node in the left layer of the top level of EII-BF and the corresponding implicit (bottom-level) representation; however, residual activations were present in the bottom level, which remained in the state at the end of the rare-word association trial. These two sources of information were summed. Activations were transmitted simultaneously in the top level and the bottom level. The bottom level underwent six spins, as human participants had a maximum of two seconds ($6 \times 350 = 2100$ ms). The output from the bottom level was integrated with the activations in the right layer of the top level using the *Max* function and transformed into a Boltzmann distribution (which served as activations for the nodes in the right layer of the top level). A response was stochastically selected and the ICL was computed (as explained before) and used to estimate the response time of the model (with *intercept* = 870 and *slope* = 0.2). (Note that no threshold was used on the ICL because a response had to be output.)

44 simulations were run (one for each participant in the corresponding human experiment), each containing 52 rare-word association trials (each stimulus was seen once during the rare-word association trials), each followed by six lexical decision trials, exactly as in the human experiment.

D. Simulation Results

The results in the lexical decision task showed that the mean response time to target words (851 ms) was smaller than the response times to the distractors (859 ms) and the non-words (869 ms). This difference is highly reliable, $F(2, 13722) = 87.38$, $p < .01$, and *post hoc* Tukey analyses showed that all the word types were different ($p < .01$). It is well documented that non-words are usually slower than words in lexical decision tasks [17]. However, the difference between distractors and targets suggest the presence of priming of the targets. Fig. 4 shows the response times in the lexical decision trials split by performance in the rare-word association trials (correct vs. incorrect) and stimulus type. As can be seen, target recognition was significantly faster than distractor recognition when a correct response was given in the rare-word association task $t(5007) = 4.20$, $p < .01$, as in [3].[4] As in [3], this difference between targets and distractors was not statistically significant when an incorrect response was given in the rare-word association task $t(452) = 1.72$, $p > .05$. These simulation results above are all in line with the results in [3].

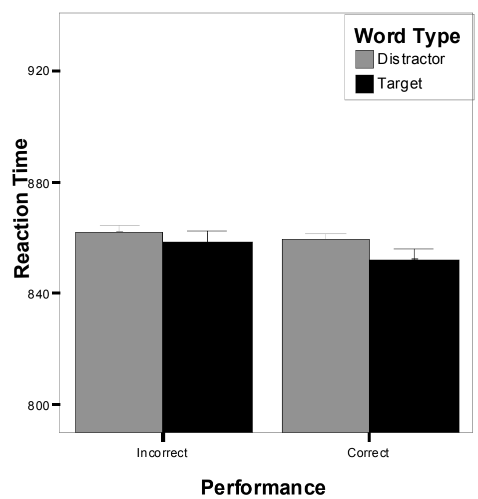

Fig. 4. Simulated response times in the lexical decision task when an answer was given in the rare-word association task.

[3] Processing from the right layer to the left layer of EII-BF uses the same equations as processing from the left layer to the right layer (as described in *Section III*). In the top level, $\mathbf{x} = \mathbf{V}^T\mathbf{y}$, and substitute $\mathbf{F} \Leftrightarrow \mathbf{E}$ and $\mathbf{y} \Leftrightarrow \mathbf{x}$ in (2). All the remaining equations were exactly the same.

[4] While this difference is small, it is still statistically significant because the standard errors were 0.82 ms and 1.81 ms for distractors and targets respectively. This difference in reaction times could have been made bigger by increasing the slope of the reaction time function (0.2). However, because this transformation is linear, the test statistics would not have been affected.

Of more interest are the trials in which no response was given in the rare-word association task. Fig. 5 shows the response times in the lexical decision trials split by FOK and stimulus type. As can be seen, the FOK (from the corresponding rare-word association trial) had a strong effect on the difference between response times to targets and distractors. The interaction between these two factors was significant $F(2, 3697) = 3.13$, $p < .05$. Further decomposition of the analysis showed that targets were faster than distractors when the participants rated their FOK as high $F(1, 3697) = 7.94$, $p < .01$; this difference was smaller for medium feeling of knowing $F(1, 3697) = 5.21$, $p < .05$, and disappeared when the participants rated their feeling of knowing as low $F(1, 3697) = 2.08$, $p > .05$. All these results are in line with [3] (see also Fig. 3).

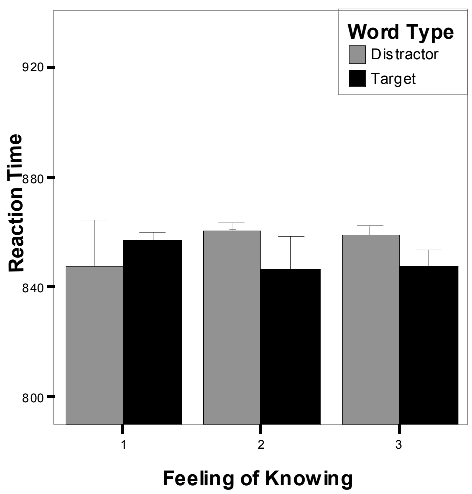

Fig. 5. Simulated response times in the lexical decision task when no response was given in the rare-word association task.

E. Discussion

The simulation results obtained with EII-BF matched well the human data in [3]. The reproduction of these qualitative and quantitative results support the psychological plausibility of the proposed model and the adequacy of the Explicit-Implicit Interaction theory. Several effects were simultaneously reproduced without varying the free parameters across conditions. The model was not designed specifically to simulate this task, but well supported by fundamental theoretical considerations [5], [8]-[9]. Overall, EII-BF captured and mechanistically explained the human data demonstrating the effect of incubation in a lexical decision task.

V. Incubation in a Free Recall Task

A. Experiment

[18] studied the effect of incubation on the number of new words recalled during a second free recall test in a two-phased free recall experiment. This measure is referred to as 'reminiscence' in experimental psychology.

The participants had five minutes to memorize 50 line drawings. Following this study phase, the participants took part in the first free recall test, which lasted 1, 2, or 4 minutes. Once the first free recall test was completed, the participants had a 0-, 1-, 5-, or 10-minute break (which constituted the incubation phase). After the incubation phase, all the participants took part in a second free recall test. The length of the second free recall test was the same as the first (and based on the same set of line drawings seen earlier, without re-studying them). 221 participants were tested in this 3 × 4 design and the dependant variable was reminiscence.

B. Empirical Results

A Test Duration × Incubation Interval ANOVA was performed on reminiscence (see Fig. 6). There was no effect of Test Duration, $F(2, 209) = .27$, $p > .7$, but Incubation Interval had a significant effect on reminiscence, $F(3, 209) = 9.4$, $p < .01$. The mean reminiscence scores for each incubation interval were 2.90, 3.15, 3.72, and 5.00. *Post hoc* tests ($\alpha = .05$) showed that the first two incubation intervals (0 and 1 minute) yielded similar reminiscence scores, and that these scores were smaller than those obtained for longer incubation intervals (5 and 10 minutes respectively, which did not differed statistically). Subsequent experiments showed that the effect of Incubation Interval on reminiscence was significant only during the first minute of the second free recall test [18].

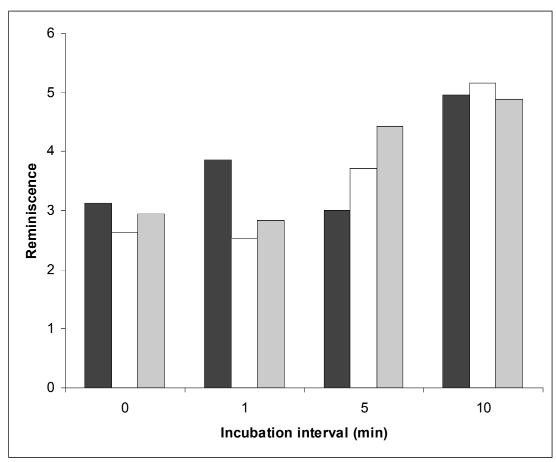

Fig. 6. Reminiscence effect found in [18]. The black bars represent 1 minute tests, the white bars represent 2 minute tests, and the grey bars represent 4 minute tests.

C. Simulation

To simulate this task, only the right layer was used in the top level of EII-BF (see Fig. 1) and each node represented a different line drawing (word).[5] In the bottom level of EII-BF, all the concepts (each represented by a top-level node) were encoded with a common pool of nodes using distributed representations ($r = 500$ and $s = 0$), and a

[5] This can be accomplished in the EII-BF model by setting the number of nodes in the left layer to zero (i.e., $n = 0$ and $m = 50$) and the knowledge integration parameter to a large value (causing top-level rules to be ignored; e.g., $\lambda = 50$).

different bottom-level distributed representation was randomly generated for each line drawing (word).[6]

To simulate the first recall test, a random pattern activated the bottom level, and the activations were propagated within the bottom level until convergence of the attractor neural network. The resulting stable activation state activated the top-level right-layer nodes through bottom-up "explicitation", and the top-level activations were transformed into a Boltzmann distribution (as activations of the corresponding top-level right-layer nodes; $\alpha = 0.15$). The Bayes factor of the distribution was used to estimate the ICL, which was compared with the default threshold value (ψ = 100). If the ICL was sufficiently high, a word (node) was stochastically chosen for recall. Otherwise, no response was output, and a new random pattern was used to activate the bottom level to start the process again. As in all other simulations (*Section IV*; see also [4]), each spin in the bottom level took 350 ms of psychological time. Hence, the durations of recall were 171, 343, and 686 spins (for 1, 2, and 4 minutes respectively).

Simulation of the incubation period was basically the same as that of the first recall test except for the following: If an item was recalled, it was stored in a buffer memory [19]. The incubation intervals were 0, 171, 857, and 1714 spins (for 0, 1, 5, and 10 minutes respectively).

The second free recall test was identical to the first, except that items in the buffer memory were output at the beginning of this period. This represented the fact that in [18], most words were recalled during the first minute of the second test (as mentioned earlier).

D. Simulation Results

216 simulations were run (18 in each of the 12 conditions, approximately the same number as in the human experiment). A Test Duration × Incubation Interval ANOVA was performed on the reminiscence scores (see Fig. 7). As in the analysis of the human data, only the incubation length had a significant effect on reminiscence $F(3, 204) = 56.49$, $p < .01$. The mean reminiscence scores were 1.70, 3.24, 5.11, and 5.56 for the 0-, 1-, 5-, and 10-minute incubation intervals respectively, which is similar to the human data. *Post hoc* Tukey analyses showed that there was a significant difference between the no incubation group and the 1-minute incubation group ($p < .05$). These two groups had a lower reminiscence score than the 5-minute incubation group and the 10-minute incubation group ($p < .05$), which did not significantly differ ($p > .05$), as in the human data.

E. Discussion

EII-BF was able to reproduce the qualitative effect of incubation on reminiscence reported in [18]. The main difference between the human data and the simulation data was the within-condition variance in the simulation data, which was smaller (thus explaining the statistically reliable difference between the no incubation and the 1-minute incubation groups). This difference could have been fixed by optimizing the value of the noise-level parameter. However, the focus of this simulation was not to capture the minute details of the free recall data but to show that a general integrative model could account for the qualitative patterns in the human data. Overall, EII-BF was successful in achieving this goal of capturing the data concerning the effect of incubation in a free recall experiment.

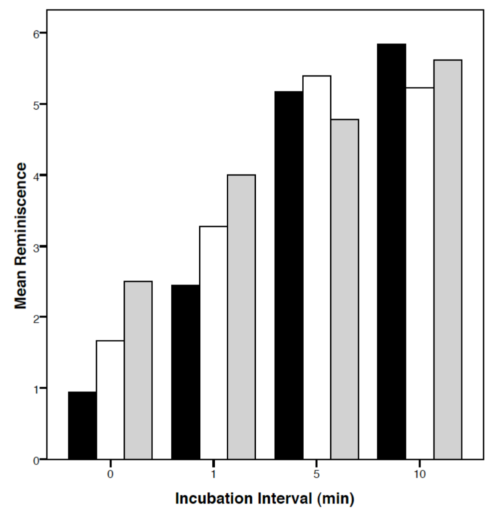

Fig. 7. Simulated reminiscence effect. The black bars represent 1 minute tests, the white bars represent 2 minute tests, and the grey bars represent 4 minute tests.

VI. Conclusion

In this paper, a new connectionist model of creative problem solving has been proposed (inspired by CLARION, but different; see [8]). EII-BF is simple and yet powerful enough to capture psychological data related to incubation in lexical decision [3] and free recall [18]. These simulations suggest that, in line with existing psychological theories and human data, the performance in different psychological tasks is affected by implicit processing while attention is diverted.

Other simulations supporting the presence of other stages of Wallas' analysis of creative problem solving [1] have been run using the EII-BF model (e.g., [4]). The results have been promising. Future work should be devoted to the simulation of many more such tasks/problems, as well as the simulation of regular problem solving to further substantiate the model.

Appendix

This section presents a concrete example of a trial in [3]. First, the participants took part in the rare-word association task:

> "A navigational instrument used in measuring angular distance, e.g., the altitude of the sun, moon and star at sea."

The participants had 15 s to find the associated word (i.e., a sextant). If no solution was produced by the participants, they were asked to rate their FOK and then started six lexical decision trials; otherwise, if a response

[6] The bottom level was pre-trained to encode the random representations with: $\eta = 0.001$, $\xi = 0.9999$, $\delta = 0.4$, and $p = 10$. For an explanation of the parameters, see [12].

was produced, the participants directly moved on to the six lexical decision trials (notwithstanding the correctness of the produced solution).

In the lexical decision task, a string of letter was presented and the participants had to decide whether the string of letters was a word or not. Three types of strings were used: 1) non-words (e.g., "DASCRIBE"), 2) distractors (i.e., a word that was not the response to the rare-word association trial, e.g., "SPENDING") and, 3) targets (i.e., the response to the rare-word association trial, here "SEXTANT"). The strings were shown one at a time. One of the six strings was always the target word. A flowchart representing a trial is shown in Fig. 2.

REFERENCES

[1] G. Wallas, *The Art of Thought*. New York: Franklin Watts, 1926.

[2] J. Dorfman, V. A. Shames, and J. F. Kihlstrom, "Intuition, incubation, and insight: Implicit cognition in problem solving," in *Implicit Cognition*, G. Underwood, Ed. New York: Oxford University Press, 1996, pp. 257-296.

[3] I. Yaniv and D. 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*, vol. 13, pp. 187-205, 1987.

[4] S. Helie and R. Sun, "Knowledge integration in creative problem solving," in *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*, B. C. Love, K. McRae, and V. M. Sloutsky, Eds. Austin, TX: Cognitive Science Society, 2008, pp. 1681-1686.

[5] S. Helie and R. Sun, "Incubation, insight, and creative problem solving: A unified theory and a connectionist model," unpublished.

[6] A. S. Reber, "Implicit learning and tacit knowledge," *Journal of Experimental Psychology: General*, vol. 118, pp. 219-235, 1989.

[7] C. Seger, "Implicit learning," *Psychological Bulletin*, vol. 115, pp. 163-196, 1994.

[8] R. Sun, P. Slusarz, and C. Terry, "The interaction of the explicit and the implicit in skill learning: A dual-process approach," *Psychological Review*, vol. 112, pp. 159-192, 2005.

[9] R. Sun, E. Merrill, and T. Peterson, "From implicit skills to explicit knowledge: A bottom-up model of skill learning," *Cognitive Science*, vol. 25, pp. 203-244, 2001.

[10] B. Kosko, "Bidirectional associative memories", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, pp. 49-60, 1988.

[11] R. E. Kass and A. E. Raftery, "Bayes factors," *Journal of the American Statistical Association*, vol. 90, pp. 773-795, 1995.

[12] S. Chartier and R. Proulx, "NDRAM: Nonlinear Dynamic Recurrent Associative Memory for learning bipolar and non-bipolar correlated patterns," *IEEE Transactions on Neural Networks*, vol. 16, pp. 1393-1400, 2005.

[13] J. B. T. Evans, "The heuristic-analytic theory of reasoning: Extension and evaluation," *Psychonomic Bulletin & Review*, vol. 13, pp. 378-395, 2006.

[14] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman, 1982.

[15] J. J. Hopfield and D. W. Tank, ""Neural" computation of decisions in optimization problems", *Biological Cybernetics*, vol. 52, pp. 141-152, 1985.

[16] S. M. Smith and R. A. Dodds, "Incubation," in *Encyclopedia of Creativity*, vol. 2, M. A. Runco and S. R. Pritzker, eds. San Diego, CA: Academic, 1999, pp. 39-43.

[17] D. E. Meyer and R. W. Schvaneveldt, "Meaning, memory structure, and mental processes," *Science*, vol. 192, pp. 27-33, 1971.

[18] S. M. Smith and E. Vela, "Incubated reminiscence effects," *Memory & Cognition*, vol. 19, pp. 168-176, 1991.

[19] J. R. Anderson and R. Milson, "Human memory: An adaptive perspective," *Psychological Review*, vol. 96, pp. 703-719, 1989.