UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A non-Bayesian Account of the "Causal Reasoning" in Sobel, Tenenbaum & Gopnik (2004)

Permalink

https://escholarship.org/uc/item/4j60b10w

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Nakagawa, Masanori
Sakamoto, Kayo
Terai, Asuka

Publication Date

2005

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A non-Bayesian Account of the "Causal Reasoning" in Sobel, Tenenbaum & Gopnik (2004)

Serban C. Musca (smusca@upmf-grenoble.fr)
Psychology and NeuroCognition Lab - CNRS UMR 5105, University Pierre Mendes France (Grenoble 2)
1251 Avenue Centrale, BP 47, 38040 Grenoble cedex 9, France

Gautam Vallabha (vallabha@cnbc.cmu.edu)
Center for the Neural Basis of Cognition, Carnegie Mellon University, 4400 Forbes Avenue
Pittsburgh, PA 15213, USA

Abstract

Sobel, Tenenbaum & Gopnik (2004) investigated the development of causal inferences in preschoolers in three experiments with tasks adapted from conditioning literature (backwards blocking and screening-off) and concluded from this indirect evidence that children develop a mechanism for Bayesian structure learning. It is proposed that (a) the differential performances in the two tasks are more likely due to differential memory demands, and (b) the observed developmental differences between 3½ and 4½-year old children may be due to maturation of the memory system, with higher retroactive interference in younger children and lower retroactive interference in older children. This account is supported by simulations with Ans & Rousset's (1997, 2000) memory self-refreshing neural networks architecture. The implications of the account proposed here on a theory of causal relation learning are discussed.

Keywords: causal inference; retroactive interference; backwards blocking; screening-off; memory limitations; preschoolers; developmental maturation; memory self-refreshing; artificial neural networks.

Introduction

Early knowledge of the causal structure of the world is thought to result from innate abilities or from interactions with the environment during early childhood. In the latter framework, a learning mechanism must be specified in order to delineate a theory of causal relation learning. Sobel, Tenenbaum & Gopnik (2004) have recently proposed such a theory, suggesting that children construct "a 'causal graph' – an abstract representation of the causal structure of a set of variables – based on evidence about the conditional probability of those variables" (Sobel *et al*., 2004, p. 306). In particular, they proposed that children use Bayesian reasoning to construct the causal graph, and tested these claims in two experiments. In both, children were told that only certain objects (called *blickets*) cause a device (a *blicket-detector*) to be activated. In the "indirect screening-off" task, the children were shown that the detector is activated when two objects (A and B) are placed on it, and that it does not activate when object A is placed on it by itself. Then, they were asked if each object was a blicket. In the "backwards blocking" task, the detector is activated when two objects (A and B) are placed on it, and also when object A is placed on it by itself. Sobel *et al.* found that 4½-year old children, and to a lesser extent, 3½-year old children, were both able to make the expected inferences, that is that object B is a blicket in the indirect screening-off task, but not in the backwards-blocking task. Sobel *et al.* (2004) used these results to argue that children's responses are based on Bayesian structure learning rather than on learning of cause-consequence associations.

Though attractive and nicely formalized, the Bayesian account has two problematic limitations. First, there is a conceptual problem. A Bayesian inference structure cannot operate without an initial core of knowledge. If the theory aims to explain the *origins* of this core of knowledge, one is faced with a chicken and egg problem: children are supposed to apply their statistical knowledge in order to enhance some *pre-existing* knowledge "database", but this cannot explain where the initial core of knowledge comes from. Second, the tasks used by Sobel *et al.* (2004) are adapted from conditioning literature, and are designed to tap into the memory system. To use these tasks as measures of causal reasoning, one has to assume that memory demands are the same for both groups of children. This assumption is erroneous: Both the indirect screening-off and backwards-blocking results may be shown to be memory artifacts rather than instances of causal reasoning. Further, the performance difference between the 4½ and 3½-year-olds can be explained as a maturation of the memory system rather than the development of a Bayesian mechanism. Our critique is based on the simulation of memory as a "self-refreshing neural network".

Memory as a self-refreshing neural network

A common problem with neural network models of memory is that of catastrophic forgetting. The memory of a neural network resides in connection weights that are adjusted to improve the network's performance on the current training set. Consequently, training on a new set S_2 tends to overwrite the effects of prior training on set S_1 (McCloskey & Cohen, 1989; Ratcliff, 1990). The problem can be avoided if *sequential learning* (*i.e.* first S_1 then S_2) is transformed into *concurrent learning* (S_1 and S_2 trained together). As concurrency is implausible for sequential learning (*e.g.* Blackmon *et al.*, in press), it can be approximated by having the network "internally" generate "pseudo-exemplars" representative of its previous training exemplars and intersperse these with exemplars from S_2. This approach, introduced by Robins (1995), can be implemented in a model with two complementary networks, NET1 and NET2 (Ans & Rousset, 1997, 2000; Ans, Rousset, French & Musca, 2002, 2004; French, 1997). Each network consists of an input layer connected both to an auto-associative target and to a hetero-associative target.

Ans & Rousset's (1997, 2000) dual-reverberated self-refreshing network (DRSR) is used here. This model works as follows. (1) NET1 is trained on S_1 exemplars to criterion. The auto-associative part learns the structure of the inputs, while the hetero-associative part learns the mapping to the outputs. Training ends when the criterion has been reached on both parts. (2) A random input is presented to NET1's input layer, cycled R times through the auto-associative part, then propagated through the network and mapped to a hetero-associative output. This input-output pair is called a "reverberated pseudo-pattern" (PP). Many such PPs are generated and used to train NET2; thus NET2 learns a generalization of the auto- and hetero-associative map learnt by NET1. (3) When NET1 has to learn S_2, its training involves exemplars from S_2 interspersed with PPs generated by NET2. In effect, NET1 learns new items from the environment and the old items from NET2. As a result, NET1 learns a combination of S_1 and S_2. A *memory consolidation parameter*, C, determines the relative frequency of PPs to new exemplars in the NET1 training. If C is low, learning is biased toward the new material (causing more retroactive interference, RI); if C is high, learning is biased toward the old material (causing more proactive interference).

Simulations

In the DRSR model used in our simulations, NET1 and NET2 were feedforward backpropagation networks trained with a gradient descent algorithm that minimizes the cross-entropy error function (Hinton, 1989). Both NET1 and NET2 had 20 input nodes, 8 hidden nodes, and 21 output nodes (20 auto-associative nodes and 1 hetero-associative node)[1] and were trained with a learning rate of 0.01 and a momentum of 0.7. After each phase of training (described below), 10^4 PPs were generated by NET1 and used to train NET2. In all the simulations presented below $R = 5$, and C was manipulated. C was set to 0.25 (low parameter value, high RI) to simulate the performance of 3½-year-olds, and to 10 (high parameter value, low RI) to simulate the performance of 4½-year-olds.

Simulation 1 and Simulation 2

These simulations correspond to the "indirect screening-off" and the "backwards blocking" conditions of Experiment 1 in Sobel *et al.* (2004). In the experiment, the following protocol was used:

(1) The children were familiarized with the task of using novel names for objects; they were shown a knob (or a tee-joint) and told that it was a "dax" (or a "wug").

(2) The blicket-detector was demonstrated to the children using two blocks. Each was placed on the detector, and the child's attention was drawn to the fact that one block (a "blicket") activated the detector, while the other (a non-blicket) did not.

(3) Two training trials ensured that the children understood the function of the detector; each block was placed on the detector and the child was asked whether it was a blicket or not.

(4) The experimental phase consisted of two trials using two *new* objects, A and B. This consisted of either the indirect screening-off or the backwards-blocking condition.

The stimuli for the simulations were created to match the above experimental conditions. Sixteen 10-bit binary vectors were generated, each of which was made up of five *zeros* and five *ones* (randomly placed within the 10 dimensions of the vector). From these, 10 input stimuli – 4 demonstration items, 4 training items, and 2 experimental items – were selected at random. A single stimulus (*e.g. A*) always consisted of the concatenation of a 10-bit nil vector and of the 10-bit vector for stimulus A. A compound stimulus (*e.g. AB*) consisted of the concatenation of the 10-bit vector for stimulus A and the 10-bit vector for stimulus B (note that the leftmost 10 bits of the input layer are nonzero only when there is a compound stimulus). Finally, four 10-bit vectors (two with four *ones* and two with three *ones*) were generated; these corresponded to the knobs and tee-joints in the familiarization phase of the experiment.

The training of the network proceeded in four phases. At the end of each phase, PPs were generated using NET1, and NET2 was trained on these PPs. Except for the first phase, NET1 was trained on the current inputs and the PPs generated by NET2 (one may think of NET2 as providing a summary of the training history for the ongoing learning by NET1).

Phase 1: The "knob" and "tee-joint" inputs were presented to NET1 and the auto-associative part of the network was trained on them. This "familiarized" the network with the input space (however, this phase is not essential for the overall results).

Phase 2: The "demonstration" and "training" items were presented to the network (in the experiment, the "training" trials were interactive, consisting of responses and feedback. In the simulation, these were treated as additional demonstration trials). There were 8 items in all (4 blickets and 4 non-blickets); each was individually presented to the network with the hetero-associative target set to 1 for a blicket and 0 for a non-blicket.

[1] Unless otherwise stated, training ends when the criterion has been reached on *both* the auto-associative and the hetero-associative parts of a DRSR network.

Phase 3: This was the first "experimental" trial. The network was simultaneously presented with A and B and a target output of 1. This input-output pairing is henceforth denoted as AB → 1.

Phase 4: This was the second "experimental" trial. In the "indirect screening-off" condition (Simulation 1), the network was presented with only A and a target of 1 (A → 1); in the "backwards blocking" condition (Simulation 2), the network was presented with A → 0. Other than this difference in Phase 4, Simulations 1 and 2 were the identical in all other respects.

Training proceeded until the RMS error for the (new) training exemplars fell under 0.1 or for a maximum of 10^4 epochs. The training exemplars were presented in random order. Each PP from NET2 was generated "on the fly" and used only once.

After the completion of Phase 4, the network was separately presented with the A and B inputs, and the activation of the output unit was taken to index a "blicket" response (possibly based on a task-specific decision threshold). Finally, the results for each simulation were averaged over 16 replications (on each replication, the 10 input stimuli were drawn at random from the pool of 16 10-bit vectors).

Figure 1 shows the results for Simulation 1, and Figure 2 the results for Simulation 2 (note that the vertical scale in Figure 2b goes from 0.7 to 1.0). Both simulations qualitatively match the behavioral results. Simulation 1 results are free from the ceiling effect that is observed on stimulus *B*, and of the floor effect observed on stimulus *A* at age 4½.

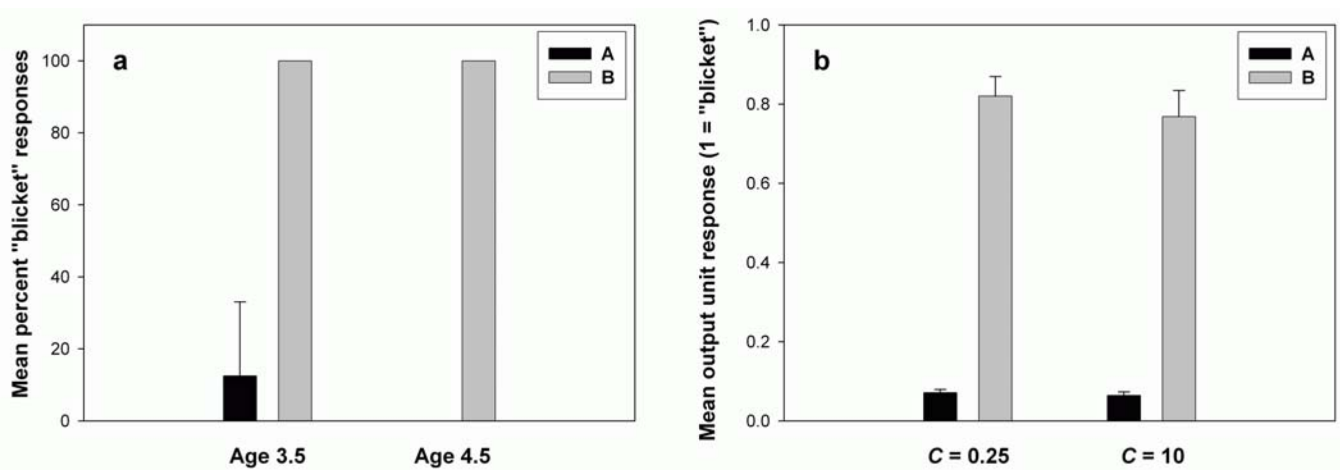

Figure 1: Indirect screening-off condition: a) behavioral results (from Table 1 in Sobel *et al.*, 2004, p. 312); b) simulations with DRSR, with the memory consolidation parameter (*C*) manipulated.

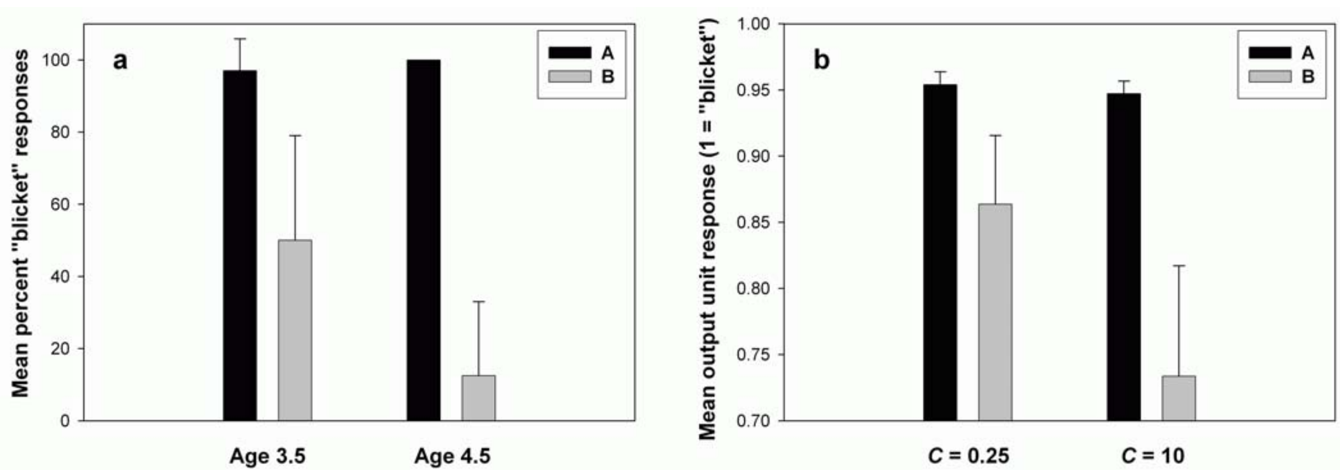

Figure 2: Backwards blocking condition: a) behavioral results (from Table 1 in Sobel *et al.*, 2004, p. 312); b) simulations with DRSR, with the memory consolidation parameter (*C*) manipulated.

Simulation 3

The simulation corresponds to Experiment 3 in Sobel *et al.* (2004)[2]. Children were given a close variant of the backwards blocking task in Experiment 1, but preceded by a phase where the frequency of the blickets was manipulated. In this phase, a child was presented with 10 different blocks of which (a) 9 were identified as blickets and 1 as a non-blicket (the "common" condition), or (b) 1 was identified as a blicket and 9 as non-blickets (the "rare" condition). Children in the "common" condition were more likely to say that B was also a blicket.

The stimuli were those from Simulation 2, except for two additional vectors made up of five *zeros* and five *ones*. Out of these 18 10-bit vectors, 16 input stimuli (2 demonstration items, 2 training items, 10 frequency-training items, 2 experimental items) were selected at random. The same Phase 1 procedure as in Simulation 2 was used. Phase 2 consisted of the presentation of "demonstration" items A′ and B′ (the network was presented $\{A'B' \rightarrow 1\}$, then $\{A' \rightarrow 1\}$ and then $\{B' \rightarrow 0\}$) and of "training" items A″ and B″ (the network was presented $\{A''B'' \rightarrow 1\}$, then $\{A'' \rightarrow 1\}$ and then $\{B'' \rightarrow 0\}$). This was followed by additional training that manipulated the blicket frequency. The network was presented with 10 new input stimuli with 9 identified as blickets and 1 as a non-blicket (the "common" condition), or with 1 identified as a blicket and 9 as non-blickets (the "rare" condition). After this training, Phases 3 and 4 proceeded as before. Finally, the results for each simulation (the activation of the output unit in response to a B input) were averaged over 25 replications (on each replication, the 16 input stimuli were drawn at random from the pool of 18 10-bit vectors).

The results presented in Figure 3 concern the response to stimulus *B* only. The simulation results match very well the behavioral results, except for the "rare" conditions in 3½-year-olds (note that the vertical scale in Figure 3b goes from 0.3 to 1.0). However, it is surprising that when blickets are rare stimulus *B* is called *more often* a blicket than when blickets are neither common nor rare (81% blicket response for stimulus *B* in the "rare" condition in 3½-year-olds in Experiment 3 as compared to 50% blicket response for stimulus *B* in 3½-year-olds in Experiment 2). We thus propose that the discrepancy arises from a bias for responding *blicket* in the "rare" condition of Experiment 3 in 3½-year-olds.

Discussion

Considering that only one parameter was manipulated and the networks were initialized with random weights, the behavioral results of Sobel *et al.* (2004) were simulated surprisingly well. The reasons for the model's success are quite straightforward – the self-refreshing mechanism tends to transform sequential learning into concurrent learning. With this simplification, two-network self-refreshing memory may be understood by analogy to a three unit network with two input units (X and Y) connected to an output unit (O), with a linear activation function and weight update governed by the delta rule (Widrow & Hoff, 1960), $dw_{ij} = \varepsilon \cdot act_i \cdot (target - actual)$. Note that the delta rule imposes a fixed point – if the net input is less than the target, then weights from the active input units are increased; if the net input is greater, the weights are decreased. The "A" and "B" inputs are presented by clamping input units X and Y (respectively) to 1.0.

For the indirect screening-off condition, imagine that the 3-unit network is trained on $\{AB \rightarrow 1\}$ until the weights stabilize at 0.5 each. The second, new trial, is $\{A \rightarrow 0\}$. In accordance with idea of a self-refreshing memory, the 3-unit network would be trained on both $\{AB \rightarrow 1\}$ and $\{A \rightarrow 0\}$. Each time the network is presented $\{A \rightarrow 0\}$, the weight from X is decremented slightly; when $\{AB \rightarrow 1\}$ is presented, both weights increase. Over several training epochs, the weight from X slowly goes to 0 and that from Y goes to 1.0. With $C < 1$, $\{A \rightarrow 0\}$ is presented more often, so the X-weight decreases faster and the Y-weight increases to compensate. Hence, in Figure 1b there is a slight advantage for the B responses with $C = 0.25$ over $C = 10$.

For the backwards-blocking condition, again imagine that the 3-unit network is trained on $\{AB \rightarrow 1\}$ until the weights stabilize at 0.5. Then the network would be trained on both $\{AB \rightarrow 1\}$ and $\{A \rightarrow 1\}$. When $\{A \rightarrow 1\}$ is presented, the weight from X increases slightly. As a result, when $\{AB \rightarrow 1\}$ is presented, the net input is greater than the target output of 1.0. Because of the delta rule, weights from both X and Y are slightly decremented. If $\{AB \rightarrow 1\}$ is presented less often ($C < 1$), the weight from Y decreases slowly. If $\{AB \rightarrow 1\}$ is presented more often ($C > 1$), the weight from Y decreases much more quickly. Hence, in Figure 2b, there is a substantial advantage for the B responses with $C = 0.25$ over $C = 10$.

The results of Simulation 3 may be understood by assuming that the 3-unit network has a "bias" unit that is connected to the output unit and whose activity is fixed at 1.0. If there are a variety of inputs and the target output is 1.0 most of the time (as in the training for the "common" condition) the bias weight will be positive and relatively large. When the network is subsequently trained with $\{AB \rightarrow 1\}$ and $\{A \rightarrow 1\}$, it is already predisposed to respond "1", so the error will be small and the weights from X and Y would not change much. The "common" conditions in Figure 3c (for both C = 0.25 and C = 10) reflect this predisposition of the network to respond "1". In the "rare" condition, the target output during the frequency-training is usually 0.0, so the weight from the bias unit would either be negative or close to zero. Consequently, the changes in the weights from X and Y would be very similar to those in the Simulation 2, viz. there would be a substantial advantage for the B responses with $C = 0.25$ over $C = 10$.

[2] Experiment 2 does not investigate the difference in performance between 3½-year-olds and 4½-year-olds, but replicates the results found in Experiment 1 with 4½-year-olds. For this reason, no additional simulation has been conducted for Experiment 2.

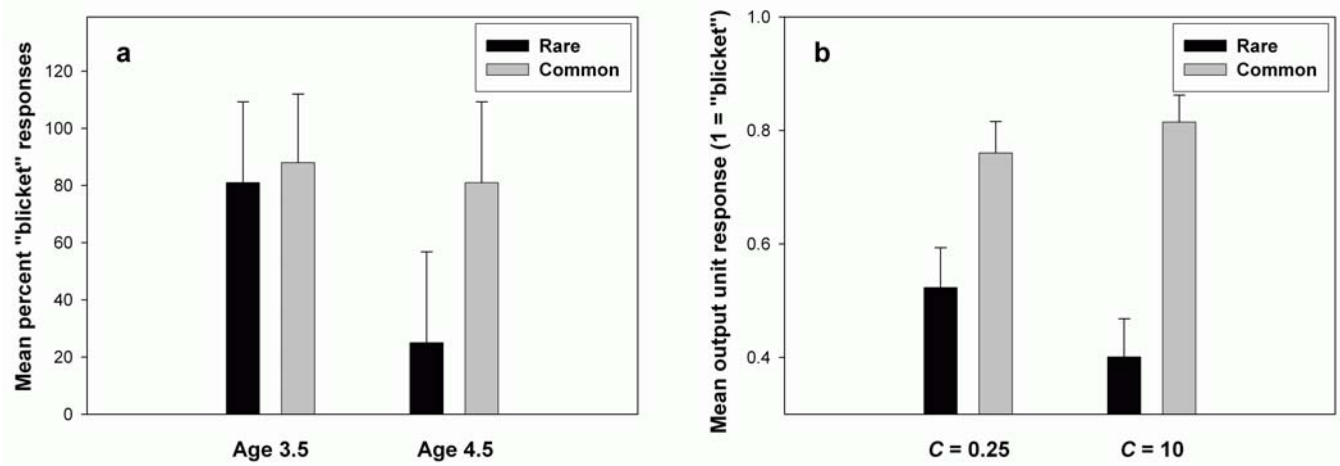

Figure 3: Backwards blocking preceded by a manipulation of blickets' frequency: a) behavioral results (from Table 5 in Sobel *et al.*, 2004, p. 325); b) simulations with DRSR, with the memory consolidation parameter (C) manipulated.

The main import of the above discussion is that the network's ability to match the data is not due to a complex associative architecture. It stems from a straightforward combination of a Rescorla-Wagner-type learning rule with interleaved exposure. The DRSR architecture is a sophisticated and general way of managing the interleaving in the absence of the old training exemplars, but it is not essential for the current purposes. Critically, children's developmentally different performances were simulated by manipulating the memory consolidation parameter that affects the level of retroactive interference. This suggests that the results presented by Sobel *et al.* (2004) in support of the idea of a Bayesian structure learning mechanism may be explained by a memory limitation in young children that fade away as they grow older. The network model also allows us to make more fine-grained predictions regarding the children's performance. One consequence of a distributed representation for the inputs is that learning cannot be separated from generalization. If all the items shown to the network as "blickets" in Phase 1 and 2 share certain properties (*e.g.* color, gross shape), then the network would be more susceptible to classifying similar objects (and more resistant to classifying dissimilar objects) as blickets in future trials. This can be tested in the experiment: say all the objects classified as blickets in the demonstration and training trials are bright red, and in the trials for indirect screening-off (where $\{A \rightarrow 0\}$) the "A" object is also bright red. This might bias the child to respond that "A" is a blicket, even though the causal reasoning may suggest otherwise. In general, it is not clear whether children classify "blickets" as an abstract category (determined purely by the detector) because this would go against the everyday experience in which objects belonging to a category usually have common features or at least have "family resemblances". Thus there may be exposure-dependent or similarity-based learning occurring in this experiment which a purely "causal reasoning" account may not capture.

Finally, we must note one point of concern regarding rate of learning. In the simulation, the model received extensive, repeated training on the same two patterns, while in the experiment the children are able to perform the task after seeing the patterns only once. However, there may have been "internal" training or repetition as part of a cognitive rehearsal process. Such a rehearsal is needed even in a symbolic process (*e.g.* Bayesian inference over a causal graph), since the cognitive system needs some way of sustaining the symbols in memory as the symbolic processes unfold. Therefore we feel that the learning rate issue does not by itself disqualify the model as an account of the children's cognitive performance.

Conclusion

The main conclusion from our simulation is not a critique of Bayesian causal reasoning *per se*, but whether such reasoning is necessary to account for the children's performance in Sobel et al. (2004). Our simulations have shown that pure "memory operations" are sufficient to account for indirect screening-off, backwards blocking, and the effect of prior frequency. This suggests that for young children (and maybe in general) the memory and reasoning systems are intertwined in highly complex ways. In particular, the "nodes" in the causal graph, which Sobel *et al.* take as a starting point for reasoning, may actually be the end result of highly sophisticated cognitive processing to isolate and abstract certain parts of the world (analogous to how our concepts of "circle" and "equilateral triangle", while starting points for geometric reasoning, are actually the result of sophisticated cognitive abstraction). Possibly, as children get older, they get socialized into a similarly sophisticated cause and effect view. In other words, we suggest that a child's developmental interaction with the environment gradually coalesces as a core of knowledge. The Bayesian structure learning proposed by Sobel *et al.* (2004) may be appropriate for describing the operation and development of rational processes that operate on this core knowledge, but it does not appear to explain how that knowledge arises to begin with.

Though we did not specifically address this question here, we would like to speculate on a possible explanation of the origins of causal reasoning. It is our sentiment that associative learning of cause-consequence patterns is at the root of this ability. However, contrary to a model of associative learning of the kind proposed by Rescorla & Wagner (1972), we suggest that the learning of cause-consequence patterns is achieved very gradually in a memory system where a memory self-refreshing mechanism becomes efficient through a developmental maturation process during early childhood.

Acknowledgments

SCM was supported in part by the European Commission (research grant HPRN-CT-1999-00065) and by the French government (CNRS UMR 5105). GV was supported by the National Institutes of Health, USA (Grant MH64445). The authors thank James McClelland for helpful discussions.

References

Ans, B., & Rousset, S. (1997). Avoiding catastrophic forgetting by coupling two reverberating neural networks. *Comptes Rendus de l'Académie des Sciences Paris, Life Sciences*, *320*, 989-997.

Ans, B., & Rousset, S. (2000). Neural networks with a self-refreshing memory: Knowledge transfer in sequential learning tasks without catastrophic forgetting. *Connection Science*, *12*, 1-19.

Ans, B., Rousset, S., French, R. M., & Musca, S. C. (2002). Preventing catastrophic interference in multiple-sequence learning using coupled reverberating Elman networks. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ans, B., Rousset, S., French, R. M., & Musca, S. C. (2004). Self-refreshing memory in artificial neural networks: learning temporal sequences without catastrophic forgetting. *Connection Science*, 16: 71-99.

Blackmon, J., Byrd, D., Cummins, R., Poirier, P., & Roth, M. (in press). Atomistic learning in non-modular systems. *Philosophical Psychology*.

French, R. M. (1997). Pseudo-recurrent connectionist networks: an approach to the 'sensitivity-stability' dilemma. *Connection Science*, *9*, 353–379.

Hinton, G. E. (1989) Connectionist learning procedures. *Artificial Intelligence*, *40*, 185–234.

McCloskey, M., and Cohen, N. J., (1989). Catastrophic interference in connectionist networks: the sequential learning problem. In H. G. Bower (Ed.), *The Psychology of Learning and Motivation, Vol. 24* (pp. 109–165). New York: Academic Press.

Ratcliff, R. (1990). Connectionist models of recognition and memory: constraints imposed by learning and forgetting functions. *Psychological Review*, *97*, 285–308.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton-Century-Crofts.

Robins, A. (1995). Catastrophic forgetting, rehearsal and pseudorehearsal. *Connection Science, 7*(2), 123-146.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, 28, 303-333.

Widrow, G., & Hoff, M. E. (1960). *Adaptative switching circuits.* Paper presented at the, Western Electronic Show and Convention, Institute of Radio Engeneers (IRE), New York.