The Interaction of Explicit and Implicit Learning: An Integrated Model

Paul Slusarz (kmicic@ideaworks.com)
Ron Sun (rsun@cecs.missouri.edu)
Department of CECS
University of Missouri-Columbia
Columbia, MO 65211, USA

Abstract

This paper explicates the interaction between the implicit and explicit learning processes in skill acquisition, contrary to the common tendency in the literature of studying each type of learning in isolation. It highlights the interaction between the two types of processes and its various effects on learning, including the synergy effect. This work advocates an integrated model of skill learning that takes into account both implicit and explicit processes; moreover, it embodies a bottom-up approach (first learning implicit knowledge and then explicit knowledge on its basis) towards skill learning. The paper shows that this approach accounts for various effects in the process control task data, in addition to accounting for other data reported elsewhere.

Introduction

The role of implicit learning in skill acquisition and the distinction between implicit and explicit learning have been widely recognized in recent years (see, e.g., Reber 1989, Stanley et al 1989, Willingham et al 1989, Proctor and Dutta 1995, Anderson 1993), Although implicit learning has been actively investigated, complex and multifaceted interaction between the implicit and the explicit and the importance of this interaction have not been universally recognized; to a large extent, such interaction has been downplayed or ignored, with only a few notable exceptions. Research has been focused on showing the *lack* of explicit learning in various learning settings (see especially Lewicki et al 1987) and on the controversies stemming from such claims. Similar oversight is also evident in computational simulation models of implicit learning (with few exceptions such as Cleeremans 1994).

Despite the lack of studies of interaction, it has been gaining recognition that it is difficult, if not impossible, to find a situation in which only one type of learning is engaged (Reber 1989, Seger 1994, but see Lewicki et al 1987). Our review of existing data (see Sun et al 2001) has indicated that, while one can manipulate conditions to emphasize one or the other type, in most situations, both types of learning are involved, with varying amounts of contributions from each (see, e.g., Sun et al 2001; see also Stanley et al 1989, Willingham et al 1989).

Likewise, in the development of cognitive architectures (e.g., Rosenbloom et al 1993, Anderson 1993), the distinction between procedural and declarative knowledge has been proposed for a long time, and advocated or adopted by many in the field (see especially Anderson 1993). The distinction maps roughly onto the distinction between the explicit and implicit knowledge, because procedural knowledge is generally inaccessible while declarative knowledge is generally accessible and thus explicit. However, in work on cognitive architectures, focus has been almost exclusively on "top-down" models (that is, learning first explicit knowledge and then implicit knowledge on the basis of the former), the bottom-up direction (that is, learning first implicit knowledge and then explicit knowledge, or learning both in parallel) has been largely ignored, paralleling and reflecting the related neglect of the interaction of explicit and implicit processes in the skill learning literature. However, there are a few scattered pieces of work that did demonstrate the parallel development of the two types of knowledge or the extraction of explicit knowledge from implicit knowledge (e.g, Rabinowitz and Goldberg 1995, Willingham et al 1989, Stanley et al 1989), contrary to usual top-down approaches in developing cognitive architectures.

Many issues arise with regard to the interaction between implicit and explicit processes: (1) How can we best capture implicit and explicit processes computationally? (2) How do the two types of knowledge develop along side each other and influence each other's development? (3) How is bottom-up learning possible and how can it be realized computationally? (4) How do the two types of knowledge interact during skilled performance and what is the impact of that interaction on performance? For example, the synergy of the two may be produced, as in Sun et al (2001). In this paper, we will focus on the interaction and the synergy resulting from the interaction.

A Model

Let us look into a model that incorporates both implicit and explicit processes.

Representation. The inaccessible nature of implicit knowledge may be captured by subsymbolic distributed representations provided by a backpropagation network (Rumelhart et al 1986). This is because representational units in a distributed representation are capable of accomplishing tasks but are subsymbolic and generally not individually meaningful (see Rumelhart et al 1986, Sun 1995); that is, they generally do not have an associated semantic label. This characteristic of distributed representation accords well with the inaccessibility of implicit knowledge. [1] In contrast, explicit knowledge may be captured in computational modeling by a symbolic or localist representations (Clark and Karmiloff-Smith 1993), in which each unit is easily interpretable and has a clear conceptual meaning, i.e., a semantic label. This characteristic captures the property of explicit knowledge being accessible and manipulable (Smolensky 1988, Sun 1995). This radical difference in the representations of the two types of knowledge leads to a two-level model CLARION (which stands for *Connectionist Learning with Adaptive Rule Induction ON-line*; proposed in Sun 1997), whereby each level using one kind of representation captures one corresponding type of process (either implicit or explicit). [2]

Learning. The learning of implicit action-centered knowledge at the bottom level can be done in a variety of ways consistent with the nature of distributed representations. In the learning settings where correct input/output mappings are available, straight backpropagation (a supervised learning algorithm) can be used for the network (Rumelhart et al 1986). Such supervised learning procedures require the a priori determination of a uniquely correct output for each input. In the learning settings where there is no input/output mapping externally provided, reinforcement learning can be used (Watkins 1989), especially Q-learning (Watkins 1989) implemented using backpropagation networks. Such learning methods are cognitively justified: e.g., Shanks (1993) showed that human instrumental conditioning (a simple type of skill learning) was best captured by associative models (i.e., neural networks), when compared with a variety of rule-based models. Cleeremans (1997) argued that implicit learning could not be captured by symbolic models.

Specifically, $Q(x,a)$ is the "quality value" of action a in state x, output from a backpropagation network. Actions can be selected based on Q values, for example, using the Boltzmann distribution (Watkins 1989).

We learn the Q value function as follows:

$$\Delta Q(x,a) = \alpha(r \quad \gamma\max_b Q(y,b) - Q(x,a)) = \alpha(r - Q(x,a))$$

where x is the current state, a is one of the action. r is the immediate reward, and $\gamma\max_b Q(y,b)$ is set to zero for the process control task we tackle in this paper, because we rely on immediate reward in this particular task (details below). $\Delta Q(x,a)$ provides the error signal needed by the backpropagation algorithm and then backpropagation takes place. That is, learning is based on minimizing the following error at each step:

$$err_i = \begin{cases} r - Q(x,a) & \text{if } a_i = a \\ 0 & \text{otherwise} \end{cases}$$

where i is the index for an output node representing the action a_i. Based on the above error measure, the backpropagation algorithm is applied to adjust internal weights (which are randomly initialized before training).

The action-centered explicit knowledge at the top level can also be learned in a variety of ways in accordance with the localist representations used. Because of the representational characteristics, one-shot learning based on hypothesis testing (Nosofsky et al 1994, Sun 1997) is needed. With such learning, individuals explore the world, and dynamically acquire representations and modify them as needed, reflecting the dynamic (on-going) nature of skill learning (Sun 1997, Sun et al 2001). The implicit knowledge already acquired in the bottom level can be utilized in learning explicit knowledge (through *bottom-up* learning; Sun et al 2001).

Initially, we hypothesize rules of a certain form to be tested (Dienes and Fahey 1995, Nosofsky et al 1994). When a measure of a rule (the IG measure) falls below the deletion threshold, we delete the rule. Whenever all the rules of a certain form are deleted, a new set of rules of a different form are hypothesized, and the cycle repeats itself. In hypothesizing rules, we progress from the simplest rule form to the most complex, in the order as shown in Figure 1, in accordance with those numerical relations used in human experiments (Berry and Broadbent 1988, Stanley et al 1989). (Other rule forms can be easily added to the hypothesis testing process. Since rules are tested in a parallel fashion, adding more rules will not drastically change the working of the model.)

The IG measure of a rule is calculated (in this process control task) based on the immediate reward at every step when the rule is applied. The inequality, $r > threshold$, determines the positivity/negativity of a step and of the rule matching this step. [3] Then, PM (positive match) and NM (negative match) counts of the matching rules are updated. IG is then calculated based on PM and NM:

$$IG(C) = log_2 \frac{PM(C) \quad c_1}{PM(C) \quad NM(C) \quad c_2}$$

where C is the current rule and c_1 and c_2 (where $2 \quad c_1 = c_2$) are Laplace estimation parameters. Thus, IG essentially measures the positive match ratio of a rule.

Simulation of human skill learning data

Simulation Focus. A number of well known skill learning tasks that involve both implicit and explicit processes were chosen to be simulated that span the spectrum ranging from simple reactive skills to more complex cognitive skills. The tasks include serial reaction time tasks,

[1] However, it is generally not the case that distributed representations are not accessible at all but they are definitely less accessible, not as direct and immediate as localist representations. Distributed representations may be accessed through indirect, transformational processes.

[2] Sun (1995, 1997), and Smolensky (1988) contain more theoretical arguments for such two-level models (which we will not get into here).

[3] In the process control task, $r = 1$ if *process-outcome* $=$ *target+/-1* and $r = 0$ otherwise, and *threshold* $= 0.9$.

$$
\begin{array}{|l|}
\hline
P = aW \quad b \\
P = aW_1 \quad b \\
P = aW \quad cP_1 \\
P = aW_1 \quad bP_2 \\
\hline
\end{array}
$$

Figure 1: The order of rules to be tested. $a = 1, 2$, $b = -1, -2, 0, 1, 2$, $c = -1, -2, 1, 2$, P is the desired system output level (the goal), W is the current input to the system (to be determined), W_1 is the previous input to the system, P_1 is the previous system output level (under W_1), and P_2 is the system output level at the time step before P_1.

process control tasks, the Tower of Hanoi task, and the minefield navigation task.

We focus on simulating process control tasks in this paper. We are especially interested in capturing the interaction of the two levels in the human data, whereby the respective contributions of the two levels are discernible through various experimental manipulations of learning settings that place differential emphases on the two levels. These data can be captured using the two-level interactive perspective.

We aim to capture (1) the verbalization effect, (2) the explicit (how-to) instruction effect, and (3) the explicit search effect. Through the simulations, it will be shown that the division of labor between, and the interaction of, the two levels is important.

To capture each individual manipulation, we do the following: (1) The explicit (how-to) instructions condition is modeled using the explicit encoding of the given knowledge at the top level (prior to training). (2) The verbalization condition (in which subjects are asked to explain their thinking while or between performing the task) is captured in simulation through changes in parameter values that encourage more top-level activities, consistent with the existing understanding of the effect of verbalization (that is, subjects become more explicit; Stanley et al 1989, Sun et al 1998). (3) The explicit search condition (in which subjects are told to perform an explicit search for regularities in stimuli) is captured through relying more on the (increased) top-level rule learning, in correspondence with what we normally observe in subjects under the kind of instruction. (4) Many of these afore-enumerated manipulations lead to what we called the synergy effect between implicit and explicit processes: that is, the co-existence and interaction of the two types of processes leads to better performance than either one alone (Sun et al 2001). By modeling these manipulations, we at the same time capture the synergy effect as well.

General Model Setup. Many parameters in the model were set uniformly as follows: Network weights were randomly initialized between -0.01 and 0.01. Percentage combination of the two levels (through a weighted sum) is used: that is, if the top level indicates that action a has an activation value l_a (which should be 0 or 1 as rules are binary) and the bottom level indicates that a has an activation value q_a (the Q-value), then the final outcome is $v_a = w_1 \quad l_a \quad w_2 \quad q_a$. The combination weights of the two levels were set at $w_1 = 0.2$ and $w_2 = 0.8$. Stochastic decision making with the Boltzmann distribution (based on the weighted sums) is then performed to select an action out of all the possible actions. The Boltzmann distribution is as follows:

$$p(a \ x) = \frac{e^{v_a \ \alpha}}{\sum_i e^{v_{a_i} \ \alpha}}$$

Here α controls the degree of randomness (temperature) of the decision-making process. It was set at 0.01. (This method is also known as Luce's choice axiom.) Other parameters include numbers of input, output, and hidden units, the external reward, the rule deletion threshold, the backpropagation learning rate, and the momentum. Most of these parameters were not free parameters, because they were set in an a priori manner (based on our previous work), and not varied to match the human data.

For modeling each of these manipulations, usually only one or a few parameter values are changed. These parameters are changed as follows. To capture the verbalization effect, we raise the rule deletion threshold at the top level. The hypothesis is that, as explained earlier, verbalization tends to increase top-level activities, especially rule learning activities. To capture the explicit search effect, we increase the weighting of the top level in addition to raising the rule deletion threshold. The hypothesis is that explicit search instructions tend to increase the reliance on top-level rule learning. To capture the explicit instruction effect, we simply wire up explicit a priori knowledge at the top level.

Simulating Stanley et al (1989)

The task. Two versions of the process control task were used in Stanley et al (1989). In the "person" version, subjects were to interact with a computer simulated "person" whose behavior ranged from "very rude" to "loving" (over a total of 12 levels) and the task was to maintain the behavior at "very friendly" by controlling his/her own behavior (which could also range over the 12 levels, from "very rude" to "loving"). In the sugar production factory version, subjects were to interact with a simulated factory to maintain a particular production level (out of a total of 12 possible production levels), through adjusting the size of the workforce (which has 12 levels). In either case, the behavior of the simulated system was determined by $P = 2 \quad W - P_1 \quad N$, where P was the current system output, P_1 was the previous system output, W was the subjects' input to the system, and N was noise. Noise (N) was added to the output of the system, so that there was a chance of being up or down one level (a 33% chance respectively).

There were four groups of subjects. The control group was not given any explicit how-to instruction and not asked to verbalize. The "original" group was required to verbalize: Subjects were asked to verbalize after each block of 10 trials. Other groups of subjects were human data

	sugar task	person task
control	1.97	2.85
original	2.57	3.75
memory training	4.63	5.33
simple rule	4.00	5.91

Figure 2: The human data for the process control task from Stanley et al (1989).

model data

	sugar task	person task
control	2.276	2.610
original	2.952	4.187
memory training	4.089	5.425
simple rule	4.073	5.073

Figure 3: The model data for the task of Stanley et al (1989).

given explicit instructions in various forms, for example, "memory training", in which a series of 12 correct input/output pairs was presented to subjects, or "simple rules", in which a simple heuristic rule ("always select the response level half way between the current production level and the target level") was given to subjects. The numbers of subjects varied across groups. 12 to 31 subjects were tested in each group. All the subjects were trained for 200 trials (20 blocks of 10 trials).

The data. The exact target value plus/minus one level (that is, "friendly", "very friendly", or "affectionate") was considered on target. The mean scores (numbers of on-target responses) per trial block for all groups were calculated. Analysis showed the verbalization effect: The score for the original group was significantly higher than the control group ($F(1,73) = 5.20, p < 0.05$). Analysis also showed the explicit instruction effect: The scores for the memory training group and for the simple rule group were also significantly higher than the control group. See Figure 2.

The model setup. The model was set up as described earlier. We used 168 input units, 40 hidden units, and 12 output units. There were 7 groups of input units, each for a particular (past) time step, constituting a moving time window. Each group of input units contained 24 units, in which half of them encoded 12 system output levels and the other half encoded 12 system input levels at a particular step. The 12 output units indicated 12 levels of subjects' input to the system. The learning rate was 0.1. The momentum was 0.1.

The rule deletion threshold was set at 0.15 for simulating control subjects. To capture the verbalization condition, the rule deletion threshold was raised to 0.35 (to encourage more rule learning activities). To capture the explicit instruction conditions, in the "memory training" condition, each of the 12 examples was wired up at the top level as simple rules (in the form of $P_1 -\quad W$); in the "simple rule" condition, the simple rule (as described earlier) was wired up at the top level. A reward of 1 was given when the system output was within the target range. In simulating the person task (a common, everyday task), we used pre-training of 10 blocks before data collection, to capture prior knowledge subjects likely had in this type of task.

The match. Our simulation captured the verbalization effect in the human data well. See Figures 2 and 3. We used a t test to compare the "original" group with the control group in the model data, which showed a significant improvement of the original group over the control group ($p <\ 01$), the same as the human data.

Our simulation also captured the explicit instruction effect, as shown in Figure 3. We used pair-wise t tests to compare the "memory training" and "simple rule" groups with the control group in the model data, which showed significant improvements of these two groups over the control group, respectively ($p <\ 01$).

Both effects point to the positive role of the top level. When the top level is enhanced, either through verbalization or through externally given explicit instructions, performance is improved, although such improvement is not universal (Sun et al 2001). They both showed synergy between the top-level explicit processes and the bottom-level implicit processes.

Simulating Berry and Broadbent (1988)

The task. The task was similar to the computer "person" task in Stanley et al (1989). Subjects were to interact with a computer simulated "person" whose behavior ranged from "very rude" to "loving" and the task was to maintain the behavior at "very friendly" by controlling his/her own behavior (which could also range from "very rude" to "loving"). In the salient version of the task, the behavior of the computer "person" was determined by the immediately preceding input of the subject: It was usually two levels lower than the input ($P = W - 2 \quad N$). In the non-salient version, it was determined by the input before that and was again two levels lower than that input ($P = W_1 - 2 \quad N$). Noise (N) was added to the output of the computer "person" so that there was a chance of being up or down one level (a 33% chance respectively).

Four groups of subjects were used: salient experimental, salient control, non-salient experimental, and non-salient control. The experimental groups were given explicit search instructions after the first set of 20 trials, and after the second set of 20 trials were given explicit instructions in the form of indicating the relevant input that determined the computer responses (W or W_1). 12 subjects per group were tested.

The data. The exact target value plus/minus one level (that is, "friendly", "very friendly", or "affectionate") was considered on target. The average number of trials on target was recorded for each subject for each set of 20 trials. Figure 4 shows the data for the four groups of subjects for the three sets of trials. Analysis showed that on the first set, neither of the two experimental groups differed significantly from their respective control groups.

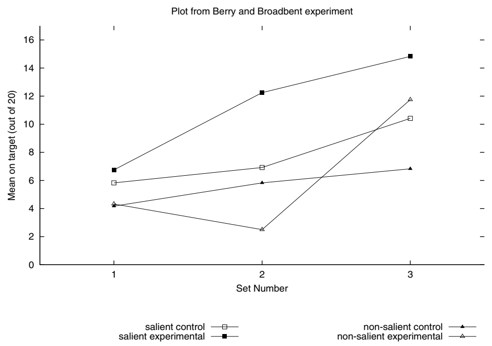

Figure 4: The data of Berry and Broadbent (1988).

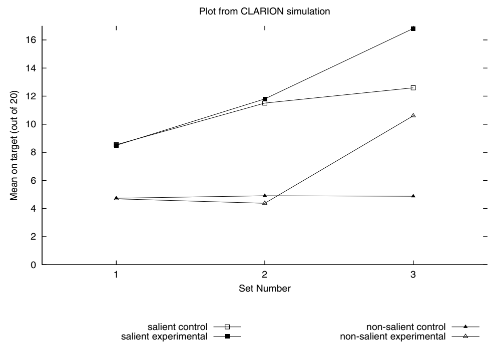

Figure 5: The simulation of Berry and Broadbent (1988).

However, on the second set, the salient experimental group scored significantly higher than the salient control group ($p < 0.01$), but the non-salient experimental group scored significantly less than the non-salient control group ($p < 0.05$). On the third set, both experimental groups scored significantly higher than their respective control groups ($p < 0.01$). The data clearly showed (1) the explicit search effect: improving performance in the salient condition and worsening performance in the non-salient condition; (2) the explicit instruction effect: improving performance in all conditions; as well as (3) the salience difference effect (during the 2nd set, under the explicit search condition).

The model setup. The model was set up similarly as described earlier for simulating Stanley et al (1989), except the following differences. The rule deletion threshold was set at 0.1 initially. To capture the explicit search effect (during the second training set), the rule deletion threshold was raised to 0.5 (for increased learning activities in the top level), and the weighting of the two levels was changed to 0.5/0.5 (for more reliance on the top level). To capture the explicit instructions given in this task (during the third training set), only rules that related the given critical variable to the system output were hypothesized and tested at the top level thereafter, in correspondence with the instructions (that is, $P = aW \quad b$, where W is the critical variable indicated by the instructions). The learning rate was 0.04. The momentum was 0.

The match. We captured in our simulation of this task the following effects exhibited in the human data: the salience difference effect, the explicit search effect, and the explicit instruction effect. The results of the simulation are shown in Figure 5. On the first set, neither of the two experimental groups differed significantly from their respective control groups; however, on the second set, the salient experimental group scored slightly higher than the salient control group, but the non-salient experimental group scored slightly less than the non-salient control group. On the third set, both experimental groups scored significantly higher than their respective control groups ($p < 0.01$).

The data demonstrated clearly the explicit instruction effect (improving performance in all conditions), and showed to some extent the explicit search effect (improving performance in the salient condition and worsening performance in the non-salient condition), as well as the salience difference effect along with the explicit search effect. The data showed the extent and the limit of the synergy effect (in that the non-salient condition discouraged synergy).

General Discussions

Although implicit learning is a controversial topic, the existence of implicit processes in skill learning is not in question — what is in question is their extent and importance. We allow for the possibility that both types of processes and both types of knowledge coexist and interact with each other to shape learning and performance, so we go beyond the controversies and the studies that focused mostly on the minute details of implicit learning (Gibson et al 1997).

The incorporation of both processes allows us to ask the question of how synergy is generated between the two separate, interacting components of the mind (the two types of processes). The model may shed some light on this issue. Sun and Peterson (1998) did a thorough computational analysis of the source of the synergy between the two levels of CLARION in learning and in performance. The conclusion, based on the systematic anaylsis, was that the explanation of the synergy between the two levels rests on the following factors: (1) the complementary representations of the two levels: discrete vs. continuous; (2) the complementary learning processes: one-shot rule learning vs. gradual Q-value approximation; and (3) the bottom-up rule learning criterion used in CLARION. [4] It is very likely, in view of the match between the model and human data as detailed in this paper, that the corresponding synergy in human performance results also from these same factors (in the main).

[4] Due to lengths, we will not repeat the analysis here. See Sun and Peterson (1998) for details.

As a result of its distinct emphasis, CLARION is clearly distinguishable from existing unified theories/architectures of cognition, such as SOAR, ACT, and EPIC. For example, SOAR (Rosenbloom et al 1993) is different from CLARION, because SOAR makes no distinction between explicit and implicit learning, and is based on specialization, using only symbolic forms of knowledge. Although ACT (Anderson 1993) makes the distinction, it is different from CLARION because traditionally it focuses mainly on top-down learning (from declarative to procedural knowledge).

Concluding Remarks

This work highlights the importance of the interaction of implicit and explicit processes in skill learning. It captures the interaction through a model that includes both types of processes. This modeling work reveals something new in the existing data (cf. Gibson et al 1997, Lebiere et al 1998). The contribution of this model lies in capturing human data in skill learning through the interaction of the two types of processes, and also in demonstrating the computational feasibility and psychological plausibility of bottom-up learning (Sun et al 2001).

References

M. Ahlum-Heath and F. DiVesta, (1986). The effect of conscious controlled verbalization of a cognitive strategy on transfer in problem solving. *Memory and Cognition*. 14, 281-285.

J. R. Anderson, (1993). *Rules of the Mind*. Lawrence Erlbaum Associates. Hillsdale, NJ.

D. Berry and D. Broadbent, (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*. 79, 251-272.

A. Clark and A. Karmiloff-Smith, (1993). The cognizer's innards: a psychological and philosophical perspective on the development of thought. *Mind and Language*. 8 (4), 487-519.

A. Cleeremans, (1994). Attention and awareness in sequence learning. *Proc. of Cogntiive Science Society Annual Conference*, 330-335.

Z. Dienes and R. Fahey, (1995). The role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 21, 848-862.

F. Gibson, M. Fichman, and D. Plaut, (1997). Learning in dynamic decision tasks: computational model and empirical evidence. *Organizational Behavior and Human Decision Processes*, 71 (1), 1-35.

A. Karmiloff-Smith, (1986). From meta-processes to conscious access: evidence from children's metalinguistic and repair data. *Cognition*. 23. 95-147.

S. Keele, R. Ivry, E. Hazeltine, U. Mayr, and H. Heuer, (1998). The cognitive and neural architecture of sequence representation. Technical report No.98-03, University of Oregon.

C. Lebiere, D. Wallach, and N. Taatgen, (1998). Implicit and explicit learning in ACT-R. *Proc. of ECCM'98*, pp.183-189. Nottingham University Press.

P. Lewicki, M. Czyzewska, and H. Hoffman, (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 13 (4), 523-530.

R. Nosofsky, T. Palmeri, and S. McKinley, (1994). Rule-plus-exception model of classification learning. *Psychological Review*. 101 (1), 53-79.

M. Rabinowitz and N. Goldberg, (1995). Evaluating the structure-process hypothesis. In: F. Weinert and W. Schneider, (eds.) *Memory Performance and Competencies*. Lawrence Erlbaum, Hillsdale, NJ.

A. Reber, (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*. 118 (3), 219-235.

D. Rumelhart, J. McClelland and the PDP Research Group, (1986). *Parallel Distributed Processing*. MIT Press, Cambridge, MA.

P. Rosenbloom, J. Laird, and A. Newell, (1993). *The SOAR papers: Research on Integrated Intelligence*. MIT Press, Cambridge, MA.

C. Seger, (1994). Implicit learning. *Psychological Bulletin*. 115 (2), 163-196.

P. Smolensky, (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences*, 11 (1), 1-74.

W. Stanley, R. Mathews, R. Buss, and S. Kotler-Cope, (1989). Insight without awareness. *Quarterly Journal of Experimental Psychology*. 41A (3), 553-577.

R. Sun, (1995). Robust reasoning: integrating rule-based and similarity-based reasoning. *Artificial Intelligence*. 75, 2. 241-296.

R. Sun, (1997). Learning, action, and consciousness: a hybrid approach towards modeling consciousness. *Neural Networks*, special issue on consciousness. 10 (7), pp.1317-1331.

R. Sun and T. Peterson, (1998). Autonomous learning of sequential tasks: experiments and analyses. *IEEE Transactions on Neural Networks*, Vol.9, No.6, pp.1217-1234.

R. Sun, E. Merrill, and T. Peterson, (2001). From implicit skill to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science*.

C. Watkins, (1989). *Learning with Delayed Rewards*. Ph.D Thesis, Cambridge University, Cambridge, UK.

D. Willingham, M. Nissen, and P. Bullemer, (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 15, 1047-1060.