Top-Down versus Bottom-Up Learning in Skill Acquisition

Ron Sun (rsun@cecs.missouri.edu)
Xi Zhang (xzf73@mizzou.edu)
Department of CECS, University of Missouri, Columbia, MO 65211, USA

Abstract

This paper studies the interaction between implicit and explicit processes in skill learning, in terms of top-down learning (that is, learning that goes from explicit to implicit knowledge) vs. bottom-up learning (that is, learning that goes from implicit to explicit knowledge). Instead of studying each type of knowledge (implicit or explicit) in isolation, we highlight the interaction between the two types of processes, especially in terms of one type giving rise to another. The work presents an integrated model of skill learning that takes into account both implicit and explicit processes and both top-down and bottom-up learning. We examine and simulate human data in the Tower of Hanoi task. The paper shows how the quantitative data in this task may be captured using either top-down or bottom-up approaches, although top-down learning is a more apt explanation of the human data currently available. The results demonstrate the difference between the two different directions of learning (top-down vs. bottom-up), and also provide a new perspective on skill learning in the Tower of Hanoi task.

Introduction

This paper studies the interaction between the implicit and explicit processes in skill learning. It explores two directions of skill learning: top-down learning and bottom-up learning. Top-down learning goes from explicit knowledge to implicit knowledge, while bottom-up learning goes from implicit knowledge to explicit knowledge. Instead of studying each type of knowledge (implicit or explicit) in isolation, we want to highlight the interaction between the two types of processes, especially in terms of one type giving rise to another.

In this work, we want to test possibilities of bottom-up learning vs. top-down learning. We do so by using the task of Tower of Hanoi, which is arguably a typical benchmark problem in high-level cognitive skill acquisition and has been used in many previous studies of skill acquisition, cognitive modeling, and cognitive architectures (see, e.g., Proctor and Dutta 1995, Anderson 1993, Anderson and Lebiere 1998).

To explore bottom-up and top-down learning, the work presents an integrated model of skill learning that takes into account both implicit and explicit processes and both top-down and bottom-up learning, although the model was initially designed as a purely bottom-up learning model. We examine and simulate human data in the Tower of Hanoi task. The work shows how the quantitative data in this task may be captured using either top-down or bottom-up approaches, although we will show that top-down learning is a more apt explanation of the human data currently available in this task.

Overall, the result of our simulations suggests that both directions are possible in human cognitive skill acquisition, and the actual direction may be either bottom-up or top-down (or a mix of both), depending on task settings, instructions, and other variables. These results demonstrate the two different directions of learning (top-down vs. bottom-up), and also provide a new perspective on skill learning.

Top-Down vs. Bottom-Up: The CLARION Model

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, Anderson 1993, Seger 1994, Proctor and Dutta 1995, Stadler and Frensch 1998). However, 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 (e.g., Mathews et al 1989, Sun et al 2001). Similar oversight is also evident in computational simulation models of implicit learning (with few exceptions such as Cleeremans 1994 and Sun et al 2001).

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, Sun et al 2001). Our review of existing data 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

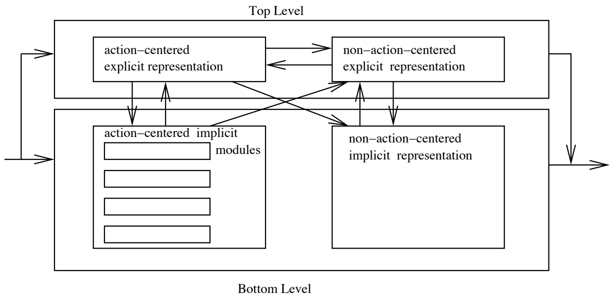

Figure 1: The CLARION architecture.

of contributions from each.

Empirical demonstrations of interaction can be found in Stanley et al (1989), Willingham et al (1989), Bower et al (1990), Wisniewski and Medin (1994), and Sun et al (2001). These demonstrations used a variety of means, including experimental manipulations such as verbalization, explicit instructions, and dual tasks.

Likewise, in the development of cognitive architectures (e.g., Rosenbloom et al 1993, Anderson 1993, Anderson and Lebiere 1998), focus has been mostly 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 implicit learning literature.

However, there are a few 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, Willingham et al 1989, Stanley et al 1989; see also Karmiloff-Smith 1986, Mandler 1992), contrary to usual top-down approaches in developing cognitive architectures.

To tackle these issues, we developed the model CLARION (Sun and Peterson 1998, Sun et al 2001). CLARION is an integrative model with a dual representational structure. It consists of two levels: the top level encodes explicit knowledge and the bottom level encodes implicit knowledge. See Figure 1. In this paper, we will focus only on action-centered components of the model.

Overall Action Decision Making

1. Observe the current state x.
2. Compute in the bottom level the "value" of each of the possible actions (a_i's) in the state x: $Q(x, a_1)$, $Q(x, a_2)$, ..., $Q(x, a_n)$.
3. Find out all the possible actions (b_1, b_2, ..., b_m) at the top level, based on the the current state information x (which goes up from the bottom level) and the existing rules in place at the top level.
4. Choose an appropriate action a, by combining (in some way) the values of a_i's (at the bottom level) and b_j's (which are sent down from the top level).
5. Perform the action a, and observe the next state y and (possibly) the reinforcement r.
6. Update the bottom level in accordance with an appropriate algorithm (to be detailed later), based on the feedback information.
7. Update the top level using an appropriate algorithm (for constructing, refining, and deleting rules, to be detailed later).
8. Go back to Step 1.

The Bottom Level

Representation The input to the bottom level consists of three groups: (1) sensory input, (2) working memory items, (3) the top item of the goal stack. The output of the bottom level is the action choice. It consists of three groups of actions: working memory set/reset actions, goal push/pop actions, and external actions. These three groups are computed by separate networks.

Learning The *Q-learning* algorithm (Watkins 1989) is a reinforcement learning algorithm. In the algorithm, $Q(x, a)$ estimates the maximum (discounted) cumulative reinforcement that can be received from the current state x on. The updating of $Q(x, a)$ is based on:

$$\Delta Q(x, a) = \alpha(r + \gamma e(y) - Q(x, a)) \qquad (1)$$

where γ is a discount factor, y is the new state resulting from action a in state x, and $e(y) = \max_b Q(y, b)$. Note that x and y include sensory inputs (internal and external), working memory items (if any activated), and the current goal (if exists).

Q-learning can be implemented in backpropagation networks (Sun and Peterson 1998). Applying Q-learning, the training of the backpropagation network is based on minimizing the following error at each step:

$$err_i = \begin{cases} r + \gamma e(y) - Q(x, a_i) & \text{if } a_i = a \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

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

The Top Level

Representation At the top level, in contrast to the bottom level (which involves distributed representation due to the use of backpropagation networks), explicit knowledge may be captured in computational modeling by a symbolic or localist representation, 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 (Sun 1995). Explicit knowledge is expressed in the form of rules.

The condition of a rule, similar to the input to the bottom level, consists of three groups: sensory input, working memory items, and the current goal. The output of a rule, similar to the output from the bottom level, is an action choice. It may be one of the three types: working memory actions, goal actions, and external actions.

Bottom-Up Learning The *Rule-Extraction-Refinement* algorithm (RER) learns explicit rules using information in the bottom level (to capture the bottom-up learning process). The basic idea of this algorithm is as follows: If an action decided by the bottom level is successful (i.e., if it satisfies a certain criterion), then the agent extracts a rule (with its action corresponding to that selected by the bottom level and with its condition specifying the current state), and adds the rule to the top-level rule network. Then, in subsequent interactions with the world, the agent refines the extracted rule by considering the outcome of applying the rule: If the outcome is successful, the agent may try to generalize the condition of the rule to make it more universal; if the outcome is not successful, then the condition of the rule should be made more specific and exclusive of the current state.

The way we measure the successfulness of an outcome (which is based on an information gain measure) and the way generalization/specialization is carried out (which is based on adding/removing allowable input values) have been fully described in Sun and Peterson (1998) and Sun et al (2001). Due to lengths, we will not repeat the details here.

Fixed Rules Some of the rules at the top level may be fixed. This type of rule (FR) represents genetic pre-endowment of an agent presumably acquired through evolutionary processes, or prior knowledge acquired from prior experience.

FRs enable top-down learning. With these rules in place, the bottom level learns under the guidance of the FRs. Initially, the agent relies mostly on the FRs in its action decision making. But gradually, when more and more knowledge is acquired by the bottom level through observing actions directed by FRs, the agent becomes more and more reliant on the bottom level (given that the cross-level combination is adaptable). Hence, top-down learning takes place.

Combining the Two Levels

In P_{RER} percent of the steps, if there is at least one RER rule indicating a proper action in the current state, we use the outcome from that rule set; in P_{FR} percent of the steps, if there is at least one fixed rule indicating a proper action in the current state, we use the outcome from that rule set; otherwise, we use the outcome of the bottom level. These probabilities are adaptable based on "probability matching" (with two parameters; Sun and Peterson 1998).

Condition/No. of disks	2	3	4	5
No verbalization	0.0	2.1	4.3	21.2
Verbalization	0.0	0.0	0.9	1.3

Figure 2: The RT data of Gagne and Smith (1962).

When we use the outcome from the top level, we randomly select an action suggested by the matching rules. When we use the outcome from the bottom level, we use the stochastic decision process for selecting an action: $p(a|x) = \frac{e^{Q(x,a)/\alpha}}{\sum_i e^{Q(x,a_i)/\alpha}}$, where x is the current state, a is an action, and α controls the degree of randomness (temperature) of the decision-making process.

Experiments

Tower of Hanoi

In the Tower of Hanoi task of Gagne and Smith (1962), there were three pegs. At the beginning, a stack of disks was stored on one of the pegs. The goal was to move these disks to another (target) peg. Only one disk can be moved at a time from one peg to another. These disks were of different sizes, and larger disks could not be placed on top of smaller disks. Initially, the stack of disks was arranged according to size so that the smallest disk was at the top and the largest was at the bottom.

Subjects were given 3-disk, 4-disk, and 5-disk versions of the task in succession, each version running until a final stable solution was found, and their mean numbers of moves (and excess moves) were recorded. Some subjects were instructed to verbalize: They were asked to explain why each move was made. The performance of the two groups of subjects (verbalization vs. no verbalization) was compared. In this task, we intend to capture the verbalization effect on performance.

Figure 2 shows the performance of the two groups in terms of mean number of excess moves (in excess of the minimum required number of moves in each version). Comparing the verbalization group and the no verbalization group in the figure, the advantage of verbalization is apparent. ANOVA indicated that there was a significant difference between verbalization and no verbalization ($p < 0.01$).

There have also been data concerning the response time of each move made by human subjects in this task. For example, the RT data from Anderson (1993) were obtained under the special instructions to subjects that encouraged a goal recursion approach (Anderson 1993). Data were available for the cases of 2, 3, 4, and 5 disks (Anderson 1993).

Bottom-Up Simulation

The Model Setup. To implement bottom-up simulation, we set up the following: (1) For deciding on each type of action (external, goal stack, or working memory actions), there is a corresponding network and a set of RER rules, respectively. (2) The input to each network is the same, including sensory input, the top goal stack (GS) item, and working memory (WM) items. (3) The outputs of the networks are external actions, GS actions and WM actions, respectively. (4) At each step, if the actions are decided by the top level, we use the existent RER rule set to get three actions—external, GS or WM actions; if the actions are decided by the bottom level, we use Boltzmann distribution to select an action from the output of each network. (5) The chosen action are coordinated and performed, and the top level and the bottom level are updated then.

During the simulation of the verbalization group, we changed the parameters for probability matching in cross-level combination to reflect the heavier reliance on the top level by the verbalization group.

Strictly speaking, GS is not necessary. But we include GS, because of generality, and because it may help learning sometimes (but it may also hamper learning sometimes). The format of GS is not important. For our simulation, each GS item includes both a subtower and a focal disc:

```
DSIZE: Size of SUBTOWER
FROM: Current peg of SUBTOWER
TO: Target peg of SUBTOWER
DSIZE1: Size of FOCAL-DISK
FROM1: Current peg of FOCAL-DISK
TO1: Target peg of FOCAL-DISK
```

A subtower is a set of disks at the top of a peg. The focal-disk is the disk beneath a subtower. Note that this set of information is redundant.

Multiple goal items could be stored in the GS one on top of another. Whenever a goal item is achieved, it will be popped.

A simple set of possible goal recursion rules is as follows (Anderson 1993):

If DSIZE $>$ 0, then push a new goal for moving a subtower of size DSIZE-1 to the spare peg and for moving the disk of size DSIZE to its target peg.

If DSIZE = 0, then make a move of FOCAL-DISK to its target peg.

If LOC(SUBTOWER)=TO and LOC(FOCAL-DISK) $\neq$ TO1, then move FOCAL-DISK to its target peg.

If LOC(SUBTOWER)=TO and LOC(FOCAL-DISK)=TO1, then pop the current goal.

Such a set of rules was hand-coded into the model in the ACT-R simulation of Anderson (1993). However, in this simulation, we did not use such hand-coded, a priori rules in the model. We want the model itself to learn something that has essentially the same effect (in both the bottom level and the top level through bottom-up learning).

The Match. The result of our simulation is shown in Figure 3. 20 runs (simulated subjects) were included in each group. Analogous to the analysis of the human data, ANOVA (number of disks $\times$ verbalization vs. no verbalization) indicated that

Condition/No. of disks	2	3	4	5
No verbalization	0.0	1.6	3.2	10.5
Verbalization	0.0	0.4	0.9	2.5

Figure 3: The bottom-up simulation of Gagne and Smith (1962).

in the model data, there was likewise a significant main effect between verbalization and no verbalization ($p < 0.01$), confirming the verbalization effect we discussed.

We compared this bottom-up simulation with a *bottom-only* simulation. We noticed that the bottom-only simulations consistently failed to learn, even when given 10 times as much training trials. This contrast suggests the importance of top-level explicit knowledge and bottom-up learning. Without them, the task was hard to learn. This fact is consistent with our synergy hypothesis (see Sun and Peterson 1998, Sun et al 2001): The reason why there are these two distinct levels (implicit and explicit) is because of the synergy that may be generated from the interaction of the two levels. The interaction of the two levels helps to improve learning, and facilitate performance and transfer (Sun et al 2001).

However, both the bottom-up and the bottom-only simulation failed to capture the RT data reported earlier.

Top-Down Simulation

The top-down simulation of the Tower of Hanoi task involves the use of fixed rules, along the line of Anderson's (1993) model, but adds the involvement of the bottom level (implicit processes), which may interfere with the top-level fixed rules. Therefore, compared with Anderson's, this is a far more complex simulation, using a more complete model that involves both explicit and implicit knowledge.

The Model Setup. Specifically, in this simulation, fixed rules were used, which implemented Anderson's (1993) analysis of subjects' performance of this task as a subset. That is, we first implemented the previous set of rules (Anderson 1993), as fixed rules at the top level of CLARION. However, this simulation was a lot more complex than top-level only (rule-based only) simulations because we had to deal with the interference from the bottom level, as the bottom level was running in parallel with the top-level rules but might recommend different actions and thus interfere with the top-level goal recursion process. The main change lied in the process of popping a sequence of goals from the GS, when a move made by the bottom level was not consistent with the top goal in the GS. In that case, we kept popping goals until reaching a goal on the GS that was consistent with the move or until the GS was empty. The structure of the GS was the same as before. The implemented set of fixed rules was an extension of the previous set. Due to their lengths, we will not show them here.

In the bottom level, Q-learning was used. Due to the use of fixed rules, Q-learning was under the "guidance" of the top level in this case. Therefore, top-down learning was involved in this case.

For capturing the performance of the verbalization subjects, the parameters for probability matching in cross-level combination were adjusted to reflect their tendencies to rely more heavily on the top level.

The Match. The result, comparing verbalization vs. no verbalization, is shown in Figure 4. 20 runs (simulated subjects) were included in each group. Analogous to the analysis of the human data, ANOVA (number of disks $\times$ verbalization vs. no verbalization) indicated that in the model data, there was likewise a significant main effect between verbalization and no verbalization ($p < 0.01$), confirming again the verbalization effect we discussed.

In this simulation, we further tackled the capturing of the RT data from Anderson (1993), which incidentally included only a portion of the total moves in each case. The data were obtained under the special instructions to subjects that encouraged the goal recursion approach (as embodied by the fixed rules used in the top level of CLARION).

Figure 5 shows the data. The comparisons between the human and the simulation data were presented for the cases of 2, 3, 4, and 5 disks. In the data, there is a regular pattern of RT peaks, which arguably reflect planning periods during which goal recursion (establishing a sequence of subgoals to be accomplished) happens (Anderson 1993).

As demonstrated by Figure 6, it is clear that the response times of the two simulated groups were reasonably close to the human data (where there was no distinction between verbalization and no verbalization). Although the match of both groups were excellent, the match between the simulated verbalization group and the human data were closer.

This particular simulation shows that the CLARION framework can accommodate traditional accounts of human performance in this task (such as Anderson 1993, Anderson and Lebiere 1998). Moreover, it extends such accounts by incorporating implicit processes (at the bottom level) as well as explicit processes (at the top level). The role of the bottom level in this task (and other high-level cognitive skill tasks) is that of "quick-and-dirty" reactions that may lead to bad performance initially due to interference with top-level rule-guided actions, but may also lead to faster and better performance given sufficient training.

The account of human RT data is important, because such an account has been viewed as the hallmark of a successful simulation. We succeeded in showing that the two-level framework of CLARION can capture the essential patterns of the human RT data, which further testifies to the cognitive validity of the model.

Condition/No. of disks	2	3	4	5
No verbalization	0.00	1.50	4.90	12.55
Verbalization	0.00	0.25	0.90	2.65

Figure 4: The top-down simulation of Gagne and Smith (1962).

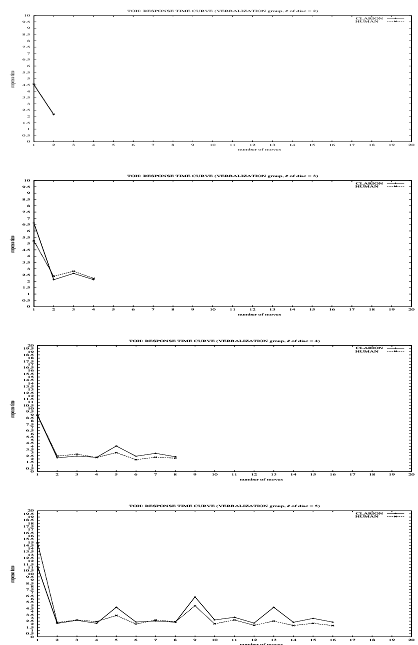

Figure 5: Simulation of the response time data of Tower of Hanoi from Anderson (1993). Four cases are included. The verbalization group is used.

Discussions

Along with the simulations of other tasks (see Sun et al 2001), we fully demonstrated that CLARION is capable of both bottom-up and top-down learning, although it was initially developed as a purely bottom-up learning model. The original reason for developing a bottom-up learning model was that in the existing literature, bottom-up learning has been very much ignored as argued by Sun and Peterson (1998) and Sun et al (2001), and therefore, there is a real need to counter-balance this bias. Our bottom-up learning model, since then, has been successful in accounting for a wide variety of skill learning tasks in a bottom-up fashion, ranging from serial reaction time tasks (sequence learning tasks), to minefield navigation tasks (Sun et al 2001). But one lingering question has been: Can this same model account for top-down learning? The present work answers this question clearly in the affirmative: CLARION can not only account for bottom-up learning data, but also the verbalization group:

	MSE	relative MSE
2-disk	0.002	0.001
3-disk	0.529	0.107
4-disk	0.252	0.098
5-disk	1.555	0.299
overall	0.967	0.200

the non-verbalization group:

	MSE	relative MSE
2-disk	0.222	0.049
3-disk	0.086	0.024
4-disk	0.579	0.109
5-disk	3.271	0.375
overall	1.925	0.236

Figure 6: The MSEs and the relative MSEs of the RT simulations of Tower of Hanoi.

top-down learning ones. And it accounts for top-down learning equally well.

Our experiments in the TOH task showed that top-down learning is a more plausible way of accounting for the existing human data in this task. This does not come as a surprise. The task structure of TOH is highly structured, with inherent recursive embedding, and involves a small number of input/output dimensions. These characteristics naturally lead to explicit processing. This tendency is even further exacerbated by the instructions that explicitly encourage goal recursion strategies.

Acknowledgments

This work was supported in part by Army Research Institute grant DASW01-00-K-0012.

References

E. Altmann and J. Trafton, (2002). Memory for goals: An activation-based model. *Cognitive Science*, 26, 39-83.

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

J. Anderson and C. Lebiere, (1998). *The Atomic Components of Thought*, Lawrence Erlbaum Associates, Mahwah, NJ.

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

R. Gagne and E. Smith, (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology*, 63, 12-18.

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

J. Mandler, (1992). How to build a baby. *Psychology Review.* 99, 4. 587-604.

R. Mathews, R. Buss, W. Stanley, F. Blanchard-Fields, J. Cho, and B. Druhan, (1989). Role of implicit and explicit processes in learning from examples: a synergistic effect. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 15, 1083-1100.

R. Proctor and A. Dutta, (1995). *Skill Acquisition and Human Performance.* Sage Publications, Thousand Oks, CA.

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

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.

M. Stadler and P. Frensch, (1998). *Handbook of Implicit Learning.* Sage Publication, Thousand Oaks, CA.

W. Stanley, R. Mathews, R. Buss, and S. Kotler-Cope, (1989). Insight without awareness: on the interaction of verbalization, instruction and practice in a simulated process control task. *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 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 skills to explicit knowledge: a bottom-up model of skill learning. *Cognitive Science*, 2001.

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.

E. Wisniewski and D. Medin, (1994). On the interaction of data and theory in concept learning. *Cognitive Science.* Vol.18, 221-281.