Declarative and Procedural Learning in Alphabetic Retrieval

Todd R. Johnson (Todd.R.Johnson@uth.tmc.edu)
Hongbin Wang (Hongbin.Wang@uth.tmc.edu)
Jiajie Zhang (Jiajie.Zhang@uth.tmc.edu)
Department of Health Informatics[1]
University of Texas – Houston Health Science Center
7000 Fannin Suite 600
Houston, TX 77030

Abstract

This paper presents three experiments that study declarative and procedural learning in alphabetic retrieval. It is based on the view that speed-up during skill acquisition can result from acquiring either new procedural knowledge or new declarative knowledge, followed by speed up of both types of knowledge. In addition, both lead to different predictions of transfer due to the different retrieval characteristics of declarative and procedural knowledge. Specifically the paper uses three forms of alphabet arithmetic problems: 1) A + 3 = ?, 2) D – 3 = ?, and 3) ? + 3 = D, to further examine the acquisition and use of declarative and procedural knowledge. The first two forms replicate experiments conducted by Rabinowitz and Goldberg (1995), whereas the third experiment attempts to maximally discriminate between declarative and procedural skill acquisition. The results provide further support for the hypothesis that speed-up can result from either declarative or procedural acquisition and strengthening.

Introduction

The alphabet is a long and well-learned list. How such lists are structured in memory and how they are accessed is certainly one of the fascinating inquiries in cognitive science.

Different from the traditional serial memory tasks, where subjects are to recall a newly-learned list of items in a specified order, in alphabetic retrieval, the content retrieval is trivial. The central interest in alphabetic retrieval research is to study how the sequence information is maintained and accessed in human memory.

The techniques often used to study alphabetic retrieval can be roughly classified into two categories based on whether letters are to be actually retrieved or not. In the first category, no letters need to be retrieved. In Lovelace and Snodgrass (1971), for example, subjects were presented two letters in a pair, and had to judge if the two are in the correct alphabetic order. In the second category, one or more letters have to be retrieved. An example is the experiment by Lovelace, Powell, and Brooks (1973), in which subjects were presented a pair of letters and were instructed to retrieve (recite) the letters between the two (see also Browman and O'Connell, 1976).

A more elegant technique in the second category is the so-called *alphabet arithmetic* task. In an alphabet arithmetic task, subjects are presented a letter, *letter1*, and a *number*. The goal is to retrieve the letter, *letter2*, that is *number* letters after (or before) *letter1*. Reaction Time (RT) is usually measured. This technique can either take the form of questions like "what comes three letters after K" (e.g., Lovelace & Spence, 1972; Hovancik, 1975; Klahr, Chase, & Lovelace, 1983), or appear in the pure algebraic form (e.g., Rabinowitz & Goldberg, 1995; Johnson, Wang, & Zhang, 1998). An example of the latter is "K+3=?". Subjects have to provide *N* as the answer, because *N* is *3* letters after *K*.

Various alphabet arithmetic studies all produce a more or less consistent result pattern. When RTs are plotted as a function of the serial positions of letter1 (the stimulus), the curve ascends non-monotonically at the aggregated level, with local peaks and valleys. However, there is no general agreement upon the theoretical explanation.

Klahr, Chase, and Lovelace (1983) proposed a theory of the cognitive structure and process involved in alphabetic retrieval. According to this theory, the alphabet is represented hierarchically. At the top level, the whole list is represented as a set of groups. At the second level, each group is represented as a set of letters. Alphabetic retrieval is a search process that occurred sequentially on both levels. First, the correct group that the to-be-retrieved letter is in has to be found. Second, the letter then has to be found within that group. Both processes are conducted by self-terminating, serial searches, starting with the initial item at each level.

Obviously, based on this theory, the gradually ascending pattern of alphabetic retrieval results from this self-terminating, serial search process: it takes longer to retrieve later letters in the alphabet. In addition, since the search occurs on two levels, a sawtooth-shaped RT curve with local peaks and valleys is evident: valleys and peaks appear at the beginning (with minimum second-level search) and the end (with maximum second-level search) of each group, respectively.

[1] Portions of this research were conducted when the authors were at The Ohio State University, Departments of Pathology (Dr. Johnson) and Psychology (Drs. Wang and Zhang).

Although the idea that serial lists are represented hierarchically in memory is quite popular (e.g., Anderson & Bower, 1973; Estes, 1972; Johnson, 1991; Shiffrin & Cook, 1978; Slamecka, 1967), some researchers argue that such a structural speculation is unjustified and a simple associative model is at least equally plausible (e.g. Scharroo, Leeuwenberg, stalmeier, & Vos, 1994). According to this association idea, the alphabet is not represented hierarchically but as a single-level associative chain. In addition, alphabetic retrieval is often a direct access rather than a serial search from the very beginning. RTs in an alphabet arithmetic task are determined by the association strengths between the stimulus and the answer. The difference in association strengths is a function of past experience of how the alphabet is learned and practiced. In this view, therefore, the increasing RT curve across the alphabet is a result of the overall decreasing association strength across the alphabet. The concept of group (or chunk), which is critical in the hierarchical view and is assumed to be responsible for the fine structure of the RT curve, is nothing more a series of letters with strong associations.

Both views explain the data reasonably well (see Scharroo, Leeuwenberg, Stalmeier, & Vos, 1994). As a result, the debate continues. Fortunately, recent progress in cognitive architectures and distinction between declarative and procedural knowledge (e.g., Anderson, 1993; Anderson & Lebiere, 1998) shed new light on how alphabetic retrieval might work. Instead of treating alphabetic retrieval as either serial searching or direct associative access, it is now possible to incorporate the two views in a unified framework of declarative/procedural distinction. Specifically, the knowledge of alphabet arithmetic can be represented either declaratively (e.g., "N is 3 letters after K") or procedurally (e.g., "To find out the letter that is 3 letters after K, count from K three times, and output the result"). While procedural knowledge is universally applicable and supports more general problem solving, such as searching, it is time consuming. On the contrary, declarative knowledge supports direct memory retrieval thus is fast, but it is conditioned on the availability of the specific declarative knowledge. As a result, when a certain problem can be solved based on stored declarative knowledge, direct retrieval is applied – no serial counting is necessary. On the other hand, when the specific knowledge necessary to solve the problem is not readily retrievable, some problem solving methods based on generally-purposed procedural knowledge, such as serial counting, have to be used. This declarative/procedural approach incorporate the hierarchical searching view and the direct association view in the sense that strong association strengths are represented by retrievable declarative knowledge, and when retrievable declarative knowledge is not available, active searching with the aid of procedural knowledge starts.

Whether alphabetic knowledge is represented declaratively or procedurally is determined by, among other things, past experience. Repeatedly solving a problem procedurally may eventually result in declarative knowledge of that problem. Rabinowitz and Goldberg (1995) nicely illustrated this phenomenon. In one of their experiments, they asked subjects to solve 432 alphabet arithmetic problems and measured their RTs. For one group of subjects, the 432 problems include a set of 12 different problems, each repeated 36 times. For another group of subjects, the problem set consists of a set of 72 different problems, each repeated 6 times. They found that although the two groups had the same RTs at the beginning of training, the first group solved the problems much faster than the second group in the later stage of training. The reason, they argued, is that both groups solved problems procedurally (i.e., by counting) at the beginning. Since the first group solved the same set of problems over and over again, they acquired declarative knowledge about these problems and began direct retrieval in the later stage. The second group did not get enough practice for any problem, thus they kept procedurally searching in the entire session.

The idea that people solve problems by applying both declarative and procedural knowledge, whichever is appropriate, has received much support (e.g., Anderson & Lebiere, 1998; Reder & Ritter, 1992; Siegler, 1988). However, how well this framework can be applied to account for various alphabetic retrieval tasks remain unexplored. It is the purpose of this paper to report a study that empirically investigates the declarative/procedural distinctions in alphabet arithmetic.

Experiment

For any specific alphabet arithmetic fact (e.g., C is 2 letters after A), we distinguish three different evaluation forms (see Figure 1). The first one is the standard *addition* form, in which subjects are presented "A + 2 = ?" and required to produce "C". The second form is a *subtraction* form, in which subjects are asked to produce "A" with respect to the problem "C – 2 = ?". The third form is called *match*. In a match form, subjects are presented "? + 2 = C", and have to report "A" as the answer.

Addtion:	A + 2 = ?
Subtraction:	C – 2 = ?
Match:	? + 2 = C

Figure 1. Three forms of alphabet arithmetic

Solving the three problems in Figure 1 essentially requires the same alphabet arithmetic fact. However, due to the different evaluation forms, different declarative and/or procedural representations might be applied. More specifically, if it is available and retrievable, a single piece of declarative knowledge, "C is 2 letters after A", can be used to quickly solve both the addition and match problems. However, a different piece of declarative knowledge, "A is 2 letters before C", has to be available and retrievable to quickly solve the subtraction problem. On the other hand, when relevant declarative knowledge is not retrievable, these problems have to be solved procedurally. Specifically, while the addition problem requires a count forward procedure (i.e., “A, B, C..”), both the subtraction and match problems require a count backward procedure (i.e, “..C, B, A”). In addition, due to the addition format in the match problem, an extra step may be needed to convert it to a recognizable subtraction format so that the count backward procedure can be applied.

118 subjects from The Ohio State University participated in the experiment. They were divided into three groups, with at least 30 subjects in each group. A learning-transfer paradigm was adopted. In the learning phase, all three groups of subjects learned to solve alphabet arithmetic problems in the addition form. In the transfer phase, each group of subjects was instructed to solve only one type of problems, either addition, subtraction, or match.

One critical manipulation in the experiment is that each subject group was further divided into two subgroups, with each having different learning experience. Specifically, in the *consistent* subgroup, subjects solved a set of 12 problems over and over again, with each problem presented 36 times. In the *varied* subgroup, subjects solved a set of 72 problems, with each only presented 6 times. It is hypothesized that subjects in the consistent subgroups gradually developed declarative knowledge about the problems that they solved repeatedly, while subjects in the varied subgroups did not due to insufficient practice.

The experimental design is shown in Table 1. It is clear that in the transfer phase, subjects in the addition group were presented new problems that they had not seen in the learning phase, although they were in the same addition form. On the contrary, subjects in the subtraction and match groups were presented a subset of the problems they had seen in the learning phase, but in different forms. It is important to note that a portion of the experiment is essentially a replication of Rabinowitz and Goldberg (1995)’s experiment.

Table 1. Experimental Design*

<table>
<tr><td rowspan="2"></td><td rowspan="2">Learning</td><td colspan="3">Transfer</td></tr>
<tr><td>addition
“A+2=?”
[30]</td><td>subtraction
“C-2=?”
[36]</td><td>match
“?+2=C”
[39]</td></tr>
<tr><td>consistent</td><td>α
(36)</td><td rowspan="2">β2
(2)</td><td rowspan="2">α
(3)</td><td rowspan="2">α
(3)</td></tr>
<tr><td>varied</td><td>α+β1
(6)</td></tr>
</table>

*An alphabet arithmetic problem takes the form of letter1 +/- number = letter2. In the experiments, letter1, letter2 ∈ {A, B, …, Z}, number ∈ {1, 2, …, 6}. With the constraint that the problem must be valid, we have a total of 135 problems. Let α be a set of 12 such problems, where each possible number appears twice. Let β be a set of 96 such problems, and $\alpha \cap \beta = \phi$. In addition, β1 contains 60 problems of β, and β2 contains the other 36 problems. Each possible number appears in β1 10 times, and in β2 6 times. The numbers in parentheses are the number of times each problem was presentedThe numbers in brackets are the numbers of subjects.

The results show that subjects could solve these alphabet arithmetic problems quite accurately. The overall error rate is 8%. There were 13 subjects who either did not follow the instruction (e.g., writing down the alphabet on a piece of paper) or had more than 20% errors. They are excluded in further analysis. More detailed error rate information, conditioned on subject groups and experimental phases, is shown in Table 2. It is clear that subjects made significantly more errors in the transfer phase, especially when they tried to solve problems in different evaluation forms.

Table 2. Error Rate

%	Addition	Subtraction	Match	Overall
Learning	6.0	7.8	6.4	6.8
Transfer	7.5	30.2	25.9	18.9
Overall	6.3	9.5	7.9	8.0

The RT results are presented separately for the learning performance and the transfer performance.

Since all three groups of subjects were trained in the same evaluation forms, the learning data is combined across the three groups. Following the practice of Rabinowitz and Goldberg (1995), to show the trend, we divided the total number of trials (432) into 36 blocks, with 12 trials in each. For each subject, we calculated the median RT of each block. Then the mean of these medians was computed across the subjects. The results are shown in Figure 2, separately for the consistent and varied conditions. It indicates that although subjects showed the same level of performance at the beginning of learning, the subjects in the consistent condition solved problems much faster toward the later stage of learning than those in the varied condition. Statistics confirmed the result. A non-linear mixed-effect exponential model was nicely fitted to the data in each condition, and the fitting curve is also shown in Figure 2. Examining the parameter estimations, it is shown that the two conditions differ significantly in terms of both the decay rate ($z=-2.07$, $p<0.05$) and the asymptote ($z=14.15$, $p<0.01$). The effect size of the asymptote difference, about 1457ms, is a strong support for the argument that different problem solving strategies were adopted in the later stage of training.

The transfer performance, conditioned on three groups of subjects and two learning conditions, is shown in Figure 3. It is easy to observe that the transfer effect (i.e., the difference between the end-of-training performance and the transfer performance) is quite different across the experimental manipulations. An overall three-way MANOVA shows that the three-way interaction among the transfer condition (addition, subtraction, or match), the learning condition (consistent or varied), and the transfer itself (the end of training performance vs the transfer performance) is significant ($F(2,99)=6.23$, $p<0.01$). Further analyses show that, 1) while the interaction between the learning condition and the transfer effect is significant in both the addition and match groups ($F(1,28)=34.20$, $p<0.01$; $F(1,37)=5.15$, $p<0.05$, respectively), it is not significant in the subtraction group; 2) In the consistent learning condition, the transfer effect is significant in all three groups ($F(1,13)=82.63$, $p<0.01$; $F(1,19)=33.46$; $p<0.01$; $F(1,20)=64.58$, $p<0.01$; for the addition, subtraction, and match groups, respectively). In addition, both the end-of-training performance and the transfer performance are not significantly different across the three groups; 3) In the varied learning condition, the interaction between the transfer effect and the transfer condition is significant ($F(2,47)=7.39$, $p<0.01$): the transfer effects in the three transfer conditions are 345.3ms, 1637.7ms, and 2759.1ms, respectively; and 4) while in both the addition and subtraction transfer conditions, the performance in the transfer stage is not different between the consistent and varied learning conditions (3150ms vs 3343ms, and 4162ms vs 3922ms, respectively), in the match transfer condition, variedly-trained subjects performed much worse (i.e., longer RTs) than those consistently-trained (5396ms vs 3109ms, $F(1,37)=23.12$, $p<0.01$).

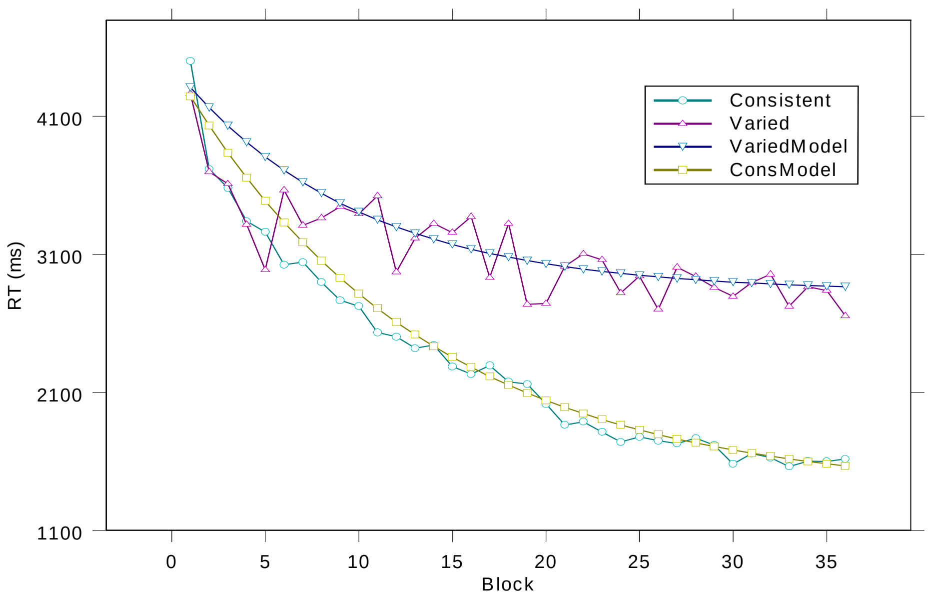

Figure 2. The Learning Performance.

Figure 3. The Transfer Performance.

Conclusions and General Discussions

Overall the results are consistent with our hypothesis about declarative/procedural distinction and interaction. First, in the addition-transfer condition, it is assumed that consistent training leads to declarative knowledge about the 12 over-learned problems, and varied training results in the counting procedure being well practiced though no declarative knowledge has been acquired. In the transfer stage, since different and new addition problems were presented, no relevant declarative knowledge was available, which left the counting procedure the only appropriate means. As a result, subjects with varied training benefited and showed a transfer effect because they had practiced and speeded up their counting procedure during their extensive training. On the contrary, subjects with consistent training showed no transfer to the new addition problems, presumably because the declarative knowledge they gained was specific to the training problems thus not useful and meanwhile they did not practice enough their counting procedure in the training.

Second, in the subtraction-transfer condition, both consistently-trained and variedly-trained subjects basically faced the same new challenge – counting down the alphabet. For those with consistent training, although the transfer problems were essentially equivalent to the training problems, due to different evaluation forms, declarative knowledge about these problems was not be applicable in the transfer stag e. In other words, it seems likely that most subjects did not realize that they could use their memory of addition results to solve subtraction problems. As a result, they had to join those variedly-trained subjects to try to adopt the brand new counting-back procedure to solve those transfer problems. Both groups showed no transfer.

Finally, the match problem serves an excellent condition to maximally discriminate between procedural and declarative learning. According to the model of alphabet arithmetic described above, to solve a match problem, subjects could use either declarative knowledge or procedural knowledge. To solve a match problem declaratively, one need only match the problem with their declarative knowledge in order to find an answer, which would suggest a perfect transfer when the corresponding declarative knowledge is available. To solve a match problem procedurally, normally one would need to first recognize that the problem is actually a subtraction problem by doing an algebraic transformation, then adopt a procedure to count back through the alphabet, which would suggest very little transfer from the previous training.

In the current match condition, subjects with varied training in fact had no choice: they had to solve the match problems procedurally, simply because they had not acquired the relevant declarative knowledge. This explains the worst transfer performance in the varied match condition. On the contrary, for those subjects with consistent training, they actually had a choice. They could solve the match problems by either using their newly acquired declarative knowledge, thus producing perfect transfer, or, they could adopt the similar converting-and-counting-back procedure, thus producing much worse transfer. The current data seems to indicate that subjects adopted neither one. In the consistent match condition, the transfer is neither perfect nor as bad as that in the varied match condition. It seems that subjects somehow combined the two approaches to solve the transfer problems. One possible scenario is that subjects used declarative knowledge to fetch the possible answer and then used a counting procedure to verify the solution. Another possibility is that some subjects used a declarative strategy and some subjects used a procedural strategy, which, when aggregated, produces the resultint pattern.

Johnson, Wang, and Zhang (1998) described an Act-R model of alphabet arithmetic that accounted for the results of the two experiments conducted by Rabinowitz and Goldberg. How this model can be applied to the match condition is of great importance to further clarify the declarative and procedural learning and application issues in alphabetic retrieval. Such a model is currently under development.

Acknowledgements

This work is funded in part by Office of Naval Research Grant No. N00014-95-1-0241.

References

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

Anderson, J. R. & Bower, G. H. (1973). *Human associative memory*. New York: Winston.

Anderson, J. R. & Lebiere, C. (1998). *The Atomic Components of Thought*. Hillsdale, NJ: Lawrence Erlbaum.

Browman, C. P. & O'Connell, D. C. (1976). Sequential phonological effects in recitation times. *Bulletin of the Psychonomic Society, 8,* 37-39.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding Processes in Human Memory (pp.161-190).* New York: Winston.

Hovancik, J. R. (1975). Reaction times for naming the first next and second next letters of the alphabet. *American Journal of Psychology, 88,* 643-647.

Johnson, G. L. (1991). A distinctiveness model of serial learning. *Psychological Review, 98,* 204-217.

Johnson, T. R., Wang, H., & Zhang, J. (1998). Modeling speed-up and transfer of declarative and procedural knowledge. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum

Klahr, D., Chase, W. G., & Lovelace E. A. (1983). Structure and process in alphabetic retrieval. *Journal of Experimental Pschology: Learning, Memory, and Cognition, 9(3),* 462-477.

Lovelace, E. A., Powell, C. M., & Brooks, R. L. (1973). Alphabetic position effects in covert and overt alphabetic recitation times. *Journal of Experimental Psychology, 99,* 405-408.

Lovelace, E. A. & Snodgrass, R. D. (1971). Decision times for alphabetic order of letter pairs. *Journal of Experimental Psychology, 88,* 258-264.

Lovelace, E. A. & Spence, W. A. (1972). Reaction times for naming successive letters of the alphabet. *Journal of Experimental Psychology, 94,* 231-233.

Rabinowitz, M. & Goldberg, N. (1995). Evaluating the structure-process hypothesis. In F. E. Weinert & W. Schneider (Eds.), *Memory Performance and Competencies: Issues in Growth and Development (pp. 225-242).* Hillsdale, NJ: Lawrence Erlbaum.

Reder, L. M. & Ritter, F. E. (1992). What determines initial feeling of knowing> Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 435-451.

Scharroo, J., Leeuwenberg, E., Stalmeier, P. F. M., & Vos, P. G. (1994). Alphabetic search: Comment on Klahr, Chase, and Lovelace (1983). *Journal of Experimental Pschology: Learning, Memory, and Cognition, 20(1),* 236-244.

Shiffrin, R. M. & Cook, J. R. (1978). A model for short-term item and order retention. *Journal of Verbal Learning and Verbal Behavior, 17,* 189-218.

Siegler, R. S. (1988). Strategy choice procedures and the development of multilication skill. *Journal of Experimental Psychology: General, 117,* 258-275.

Slamecka, N. (1967). Serial learning and order information. *Journal of Experimental Psychology, 74,* 62-66.