COGNITIVE REPRESENTATIONS OF DOS COMMANDS AS A FUNCTION OF EXPERTISE

Kathleen M. Snyder

IBM Corporation
Dept F3B, Div 20
Room 12B174
360 Hamilton Avenue
White Plains, NY

James R. Lewis

IBM Corporation
T. J. Watson Research Center
Dept 564, Room H1-E22A
Route 9A
Hawthorne, NY

ABSTRACT

The purpose of this study was to examine the cognitive networks derived from the similarity rating of IBM Personal Computer Disk Operating System (IBM PC DOS) (TM) commands by computer users with varying levels of expertise. Naive, novice, and intermediate networks were examined to determine which links in their networks were also present in an expert network. Groups with a greater level of DOS expertise had more links in common with the experts. A core set of commands were identified which were linked in every network. As the level of expertise changed, it was possible to show the order in which links in the experts' network became present in the novice and intermediate groups.

> In some uninteresting sense, all computer-based representations are equivalent. This is so because computer-based representations are embedded ultimately in the symbolic structures available in a computer language like LISP and thence down into arrangements of bits in memory. Consequently, any representation that can be used to represent arrangements of bits can indirectly bear the information in any other representation.
>
> In a practical sense, however, some representations emphasize things that are important to solving a class of problems. One scheme, therefore, is more powerful than another because it offers more convenience to the user even though, theoretically, both can do the same work. Convenience, however, is perhaps too weak a word. In general, the much greater perspicuity and the inherent thinking advantages of powerful representations enable progress that would be impossibly difficult with anything less adequate. [1, p. 22]

Although Winston's [1] comments were given in the context of the advantage of good representation in artificial intelligence, the argument is also applicable to the use of multivariate statistical methods in psychology. Different methods allow different representations of the structure hidden in a complex dataset. Like any other craftsman, we must choose the appropriate tools to achieve our analytical goals.

The goal of this study was to demonstrate the use of network analysis for modeling cognitive representations of operating system commands as a function of expertise.

Expert-Novice Studies

A number of studies have been conducted which have examined the cognitive differences between experts and novices (e.g., [2,3]). The findings of these studies have been remarkably consistent.

Chase and Simon [2] studied the ability of novice and expert chess players to recall the board positions of chess pieces. They found that experts outperformed novice players. The experts' advantage was attributed to their ability to form larger chunks of information in memory rather than above-average memory capacity. They also found that if the arrangement of chess pieces was random rather than a midgame (and presumably meaningful) arrangement, then experts recalled no more piece positions than novices.

In an attempt to replicate these findings in a more applied setting, Egan and Schwartz [3] had both novice and expert electronics technicians attempt

0073-1129/89/0000/0447$01.00 © 1989 IEEE

to recall briefly presented (5 - 10 sec) circuit diagrams. Some diagrams were meaningful and others were random reconstructions of the meaningful diagrams. As with the chess masters, the expert technicians recalled more of the meaningful diagrams, but had no performance advantage for the random arrangements. Additional analyses showed that neither more skilled guessing by expert technicians nor spatial proximity could account for the experts' performance advantage.

Alwood [4] has recently reviewed the body of work examining the behavior of novice programmers with computers, contrasting their behavior with that of experts to illustrate that which is typical of novices. In his review, he points out that the expert advantage for the recall of meaningful information has been replicated several times with novice and expert programmers recalling lines of code (e.g., [5]).

These results have been explained by hypothesizing that experts can process global display properties or code a single chunk whose relation to a more general category is known. This conceptual knowledge would allow the expert to search visually and recall from memory more systematically than novices. Therefore, a critical difference between novices and experts would be the organization of concepts in long-term memory.

According to Bateson, Alexander, and Murphy [6], another explanation is that experts may use higher-level knowledge to understand problems while novices focus on specific details. Experts may also use high-level plan knowledge to direct their activities. Bateson et al. divided 50 computer science majors into novice and expert groups, and administered tests to assess syntactic memory, semantic memory, tactical skill, and strategic skill. The best predictors of programmer expertise were semantic memory, tactical skill, and strategic skill. In particular, they concluded that semantic memory is an important cognitive factor in the development of programming skill. This research also supports the hypothesis that the organization of concepts in memory should differ due to expertise.

Multivariate Statistical Methods and Psychological Research

Dillon and Goldstein [7] have described two primary categories of multivariate statistics: dependence and interdependence methods. Multiple regression and discriminant analysis are examples of dependence methods. Principal components analysis, factor analysis, multidimensional scaling (MDS), cluster analysis, and network analysis are types of interdependence methods. Dependence methods are used to explain or predict one or more dependent variables as a function of a set of independent variables. "Interdependence methods, on the other hand, are less predictive in nature and attempt to provide insights into the underlying structure of the data by simplifying the complexities, primarily through data reduction" [7, p. 19]. For this reason, interdependence methods have been used to assess the relationship among concepts in memory.

A recent addition to the collection of interdependence methods is a link-weighted network analysis procedure called Pathfinder [8, 9]. It can be applied to the same types of data typically analyzed by MDS or cluster analysis, i.e, estimates of pairwise distance between entities. With Pathfinder, distances between entities are represented as links in a network if the resulting links form the shortest possible path between the entities. Pathfinder solutions are affected by two variables: the value of the path length function (defined by the Minkowski r-metric) and the number of links examined in constructing the network. With ordinal level measurement, r should be set to infinity to ensure a unique network structure. To avoid triangle inequality violations, N-1 links should be examined when developing the network.

Since their development and availability on computer systems powerful enough to implement them, multivariate interdependence methods have been used as tools for psychological research. Rips, Shoben, and Smith [10] demonstrated that MDS could be used to model semantic distance in memory, and that these semantic distances could be used to predict reaction times in a categorization task. Schvaneveldt, Durso, Goldsmith, Breen, and Cooke [11] used both MDS and Pathfinder to study the organization of flight-related concepts by fighter pilots of varying degrees of expertise. They concluded that both MDS and Pathfinder revealed the underlying structure of the data, but highlighted different aspects. Pathfinder focused more on local relationships among concepts, while MDS provided more global information regarding the dimensions of the concept space. In research comparing the serial recall of concept lists organized by ALSCAL-S MDS and Pathfinder representations, Cooke, Durso, and Schvaneveldt [12] found that the network organization led to faster learning than the MDS organization. This implies that the Pathfinder networks better captured the relations important for recall.

These methods have also been applied in two areas of computer development: menu organization and expert systems. McDonald, Stone, Liebelt, and Karat [13] used MDS and cluster analysis to model experts' similarity ratings of word-processing functions. They compared the performance of a different set of word-processing experts using a menu based on experts' organization of the commands and one based on a random organization. Using a paired-associates learning paradigm, they found that fewer errors were made using the organized menu. Snyder, Paap, Lewis, Rotella, Happ, Malcus, and Dyck [14] used Pathfinder to develop IBM Personal Computer Disk Operating System (IBM PC DOS) (TM) menus based on the similarity ratings of novice and expert IBM PC DOS users. A third menu had an alphabetical organization. Three groups of naive participants used the menus with a paired-associates learning task. Participants using the Pathfinder menus learned at a significantly faster rate.

Knowledge acquisition has been frequently cited as a major bottleneck in the development of expert systems, and it is possible that the multivariate interdependence methods can be used to obtain expert knowledge more efficiently [15, 16]. Olson and Rueter [16] have recently reviewed methods for knowledge acquisition, classifying them into direct (interview, questionnaires, task observation, protocol analysis, interruption analysis, and drawing closed curves) and indirect (MDS, cluster analysis, network analysis, ordered trees from recall, and repertory grid analysis) methods. They point out that both classes have their limitations, since the direct techniques generally assume that experts can verbalize what they know or how they solve problems, and the indirect methods are less rich and must make some assumptions about the underlying structure of the representation of objects and their relations. People cannot always accurately verbalize what they know [17] and the assumptions required for indirect methods may not be met.

Cooke and McDonald [18] have reviewed the research issues associated with the use of multivariate techniques as knowledge-elicitation tools, and have argued for a formal methodology for acquiring and representing expert knowledge [15]. Their formal method requires three stages: initial elicitation of a set of concepts, application of a psychological scaling technique (i.e., multivariate interdependence method), and the interpretation of the resulting representation. For example, when using Pathfinder, the final stage is the interpretation of the nature of the relationship between concepts, achieved through either direct link-labeling or cluster analysis of similarity ratings of the links (applying a second interdependence method to the analysis of the output of the first).

Interdependence Analyses of Expert and Novice Computer Users

To draw together these two lines of research (expert-novice studies and the application of multivariate interdependence methods) in the context of computer use, we will review three applicable studies. Since these statistical methods have been used to model the relationship among concepts in long-term memory, they provide the means by which one may model the change in the relationships as a function of learning, i.e., the evolution from novice to expert.

Using the method of ordered trees from cued recall, McKeithen, Reitman, Rueter, and Hirtle [5] developed hierarchical representations of ALGOL W reserved words for beginners (just starting their first ALGOL W course), intermediates (just completed their first ALGOL W course), and experts (teachers of ALGOL W courses). The method of ordered trees from cued recall requires the participants to memorize the set of concepts. After learning the concept names, the participants are given a cue (a concept) and asked to recall the words that go with it, followed by the rest of the words. Each concept serves as a cue once. Then the participant's recall strings are searched for all groups of items that always appear contiguously, regardless of order. The final result is a hierarchical representation of the concepts.

Using this method, McKeithen et al. [5] attacked three questions: Are experts more organized than beginners? Does skill affect the depth of organization? Are organizations within a skill group similar? They determined that there was no strong evidence for a difference in the amount of organization among the groups. They also concluded that depth of organization did not seem to increase with skill level. After determining the amount of similarity between pairs of representations, they performed an MDS analysis and found that experts were more alike as a group than either beginners or intermediates. Beginners appeared to use general mnemonic techniques to memorize the words, such as alphabetical order or story sequences. The different organizations created by the application of different strategies probably accounts for the variability of position of the beginners' representations in the multidimensional space. Since the experts fully understood the meaning of the concept words in the context of ALGOL W, their representations were highly similar. The intermediates' representations showed some common understanding of the ALGOL W words, but also showed the lingering effects of some common-language sequences.

Cooke [19] divided her participants into four groups: naive (no programming experience), novice (up to one year of programming experience), intermediate (one to three years of programming experience), and expert (over three years of programming experience). Her concepts were words that had meaning in a programming context, but were not restricted to a specific language. She had the participants both provide similarity ratings for pairs of concepts and to practice free recall of the concepts until they were recalled without error twice. (Note that this recall task differs from the cued recall of the Reitman-Rueter technique used by McKeithen et al. [5].) Distance estimates were obtained from the recalled lists by counting the number of intervening concepts between each pair in a recalled list. Both the similarity and the recall data were analyzed using MDS. Within-group correlations based on recalled lists were uniformly poor, while within-group correlations based on similarity ratings showed an increase from novice to intermediate to expert, replicating the findings of McKeithen et al. [5]. However, the naive within-group correlation was as strong as the expert within-group correlation. Cooke [19] concluded that "the U-shaped function suggests that naive programmers initially agree with each other on the little that they know. As learning takes place, this agreement declines at first and then increases as the programmer develops expertise" (p. 28). Based on the lack of correlation between the recall and rating data, and on the consistency in results between the rating within-group correlations and those of McKeithen et al. [5], she concluded that ratings were preferable to free recall as a method of determining semantic distance.

Kay and Black [20] examined the changes in the knowledge representations of text editing commands with increasing experience. The novices in this experiment were psychology students who used the computer for text editing and data analysis, while the experts were computer science students. Participants rated the pairwise similarity of fifteen text-editing commands. These similarity ratings were analyzed using MDS, hierarchical cluster analysis, and additive clustering analysis (similar to hierarchical cluster analysis except that an item could belong to more than one group). As with the studies by McKeithen et al. [5] and Cooke [19], a three dimensional MDS solution was found to be best. Two dimensions were found to be the same between the groups: formatting vs. non-formatting commands and specific vs. general functions. For the expert users, the third dimension seemed to be based on the dimension of destructive vs. non-destructive commands. The novices' third dimension was defined by commands that began or ended editing sessions. The interpretation of the cluster analyses was consistent with that of the MDS. Kay and Black [20] further concluded that the knowledge representation for experts is more complex (contrary to the findings by McKeithen et al., [5]) and that, with experience, users combine individual commands to form sequences or plans to accomplish the various user goals.

Goals of the Present Study

The studies by McKeithen et al. [5], Cooke [19], and Kay and Black [20] have all demonstrated the utility of multivariate interdependence methods in examining structural changes in memory as a function of expertise. The specific methods employed were ordered trees from cued recall, MDS, and cluster analysis. In the present study, we attempted to demonstrate the development of computer expertise by examining network representations of IBM PC DOS commands by users at various levels of IBM PC DOS expertise. MDS representations show the arrangement of concepts in multidimensional space. Ordered trees and cluster analysis show the way in which users have organized concepts into groups. A network representation shows concepts as nodes connected by links. Given this representation, it should be possible to illustrate the development of expertise by determining which links are common to groups at various levels of expertise with the links present in an expert network; an examination that would not be possible using the MDS or cluster analysis modes of representation. One would predict that as expertise increases, the number of links in common (and, correspondingly, the statistical agreement) with the most expert group should also increase. It should be possible to identify a "core" set of commands which are common to all groups, and to show specifically which commands are added to this core set as a function of expertise. By studying a wide range of expertise, the analysis of intragroup agreement in this study may shed light on the discrepant results reported by McKeithen et al. [5] and Cooke [19] regarding the pattern of intragroup agreement as a function of expertise.

Method

Participants

Thirty-seven IBM employees took part in the study. Participants were assigned to one of six experience categories based on their performance on a recall test of the 43 commands in IBM PC DOS. The six categories, the criterion for inclusion in the category, and the number of participants assigned to each category are shown in Table 1.

Table 1. Categories of Expertise

Experience Category	Number of Commands Recalled	Number of Participants
Naive	0	5
Novice	1 - 4	5
Intermediate 1	5 - 10	9
Intermediate 2	11 - 15	7
Intermediate 3	16 - 24	5
Expert	Over 30	6

Materials and Procedures

All participants completed a questionnaire before beginning the experimental task. The goal of the questionnaire was to identify the participants' knowledge and use of computers, programming languages, operating systems and applications. One question asked participants to list as many IBM PC DOS operating system commands as they could. The participants' responses to this question were used to classify them into one of the six experience categories.

Descriptions of the commands were printed on 3X5 cards. Participants were asked to read the descriptions and to sort the cards into groups based on the similarity of the commands. Once the commands were grouped, participants assigned labels to the groups.

If participants had been asked to generate a similarity rating for each possible pair of commands, they would have been required to make over 900 paired comparisons. We had the participants group the commands since "clustering prior to pairing is a way of reducing the number of objects so that using paired comparisons becomes a reasonable technique" [21, p. 23].

A 9-point scale was used to collect similarity ratings between a participant's command groups. A "1" meant the command groups were perceived as highly similar and a "9" meant highly dissimilar. A program written for the IBM Personal Computer AT (TM) was used to capture this paired comparisons data.

Cognitive networks were derived for each participant and for each group. This resulted in 32 individual networks and 6 group networks. To generate the networks:

1. A program was written to build a 43 by 43 similarity matrix for each participant's data. The program looked at every possible pair of commands and generated similarity ratings for each pair. Pairs from the same group of commands were assigned a value of 0. Pairs of commands from different groups received the rating assigned to the pair of groups.

2. In deriving the group (various levels of expertise) networks a second program was written to sum all of the individual matrices from the same group.

3. The matrices were then used as input to the Pathfinder algorithm. Pathfinder generates link-weighted networks from a set of distance data. In this study the distance data consisted of the matrices constructed in steps 1 and 2. The Pathfinder network consists of a set of nodes and links that connect pairs of nodes that are highly related. In the present context the nodes are IBM PC DOS commands and the links represent the perceived similarity between the commands.

Results

The group networks were compared to uncover structural changes as a function of expertise, shown in Table 2. First, the links common to all groups were identified. These constituted the "core" set of commands. Next, links that were common to all groups except the Naive group were identified. The third comparison identified links common to all Intermediate and the Expert groups only. Next, we identified the links common only to the Intermediate 2, Intermediate 3, and Expert Groups. Finally, the additional links common to the Intermediate 3 and Expert groups were listed. (See Figure 1, the Expert network, for comparison with the information in Table 2.)

Using kappa [22], the intergroup agreement with the Expert group was measured for the Naive, Novice, and Intermediate groups (see Figure 2). Kappa is a measure of agreement or similarity which can be legitimately applied to a 0-1 level of measurement. It is a more accurate measurement of similarity than the number of links in common since it takes into account both the way in which two 0-1 matrices match (links in common) and the ways in which they fail to match (links not in common). All values of kappa showed significant agreement with the Expert group ($p<.05$), and also showed a monotonically increasing trend with the increasing expertise of the group. Over this range of experience, kappa changed from .16 to .30, almost doubling, with the greatest change occurring between the naive and novice groups.

The intragroup agreement was assessed by calculating kappa for each pair of participants within a group. As shown in Figure 3, the average kappa for intragroup agreement was an S-shaped curve with the increase in experience. A one-way analysis of variance conducted on the kappas calculated for each pair of participants within a group was significant ($\underline{F}(5,101)=3.4$, $\underline{p}=.007$). A post-hoc comparison using Bonferroni $\underline{t}$-tests [23] showed significant differences ($\underline{p}<.05$) between the Naive and Expert groups, and the Intermediate 2 and Expert groups.

Figure 4 shows the relationship of intragroup agreement and expertise with the three intermediate groups collapsed. In this case, the averaged group kappas increased monotonically with increasing expertise. A one-way analysis of variance was significant ($\underline{F}(3,101)=3.75$, $\underline{p}=.01$). Bonferroni $\underline{t}$-tests showed only the Naive and Expert groups to be significantly different ($\underline{p}<.05$).

Table 2. Development of Links Common to All Higher Networks

Level of Expertise	Common Links	
Naive	Backup-Restore Find-Sort	Date-Time Mkdir-Rmdir
Novice	Recover-Restore	
Int. 1	Break-Prompt Copy-Erase	Ctty-Keybxx Label-Vol
Int. 2	Backup-Recover Chdir-Mkdir	Chdir-Rmdir
Int. 3	Attrib-Comp Break-Verify Chdir-Dir Comp-Copy Comp-Erase Comp-Ren	Copy-Ren Ctty-Prompt Erase-Ren Find-More Mode-Prompt
Expert	47 additional links were only found in the Expert network (see Figure 1).	

Discussion

As predicted, the relationship between the expert group and the other groups increased as a function of expertise. While this is not surprising, it supports the use of Pathfinder for this type of cognitive modeling and also validates the use of free recall of commands as a measure of expertise. However, the use of free recall of concepts is restricted to domains in which the set of concepts to be recalled has already been identified, such as commands in an operating system or a programming language's reserved words. The evolution of expertise revealed in Table 2 shows the way in which an initial nucleus of properly linked commands (the core commands) grew with the increasing expertise of the groups. While no studies have been conducted demonstrating the utility of such a representation in the development of training materials, it is reasonable to suggest that modeling the change in network structure as a function of expertise could provide valuable guidance to such an effort. This representation would allow a developer to identify which concepts a naive user has properly related, and suggests the order in which new concepts should be introduced and related to old concepts. However, one should be cautious when

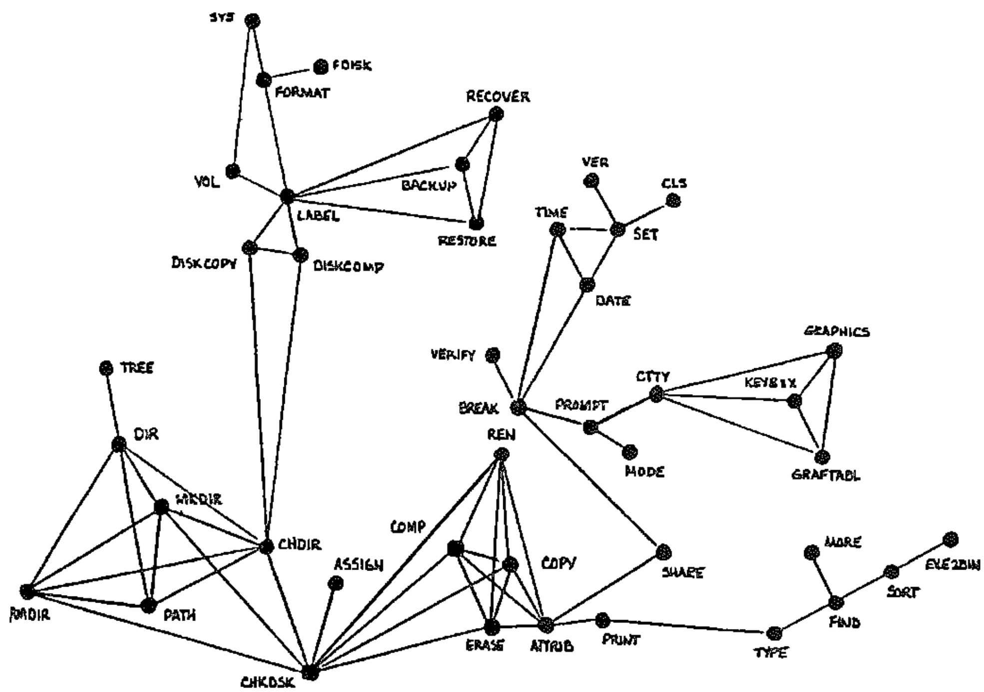

Figure 1. The expert network. (Note that the length of a line does not necessarily reflect its link-weight.)

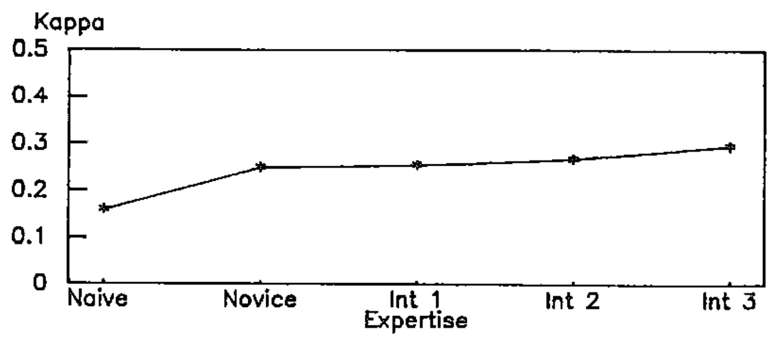

Figure 2. Relationship of Naive, Novice, and Intermediate groups to the Expert group.

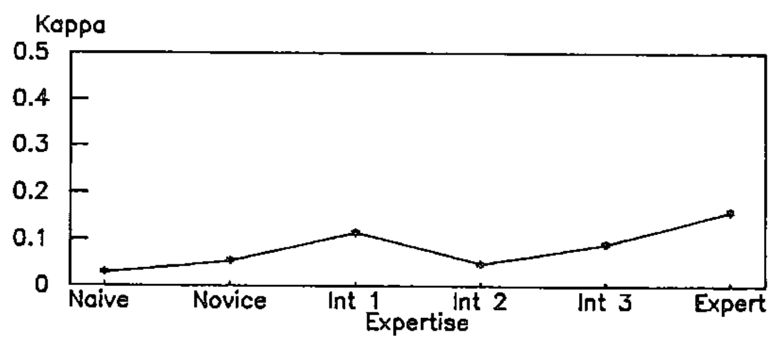

Figure 3. Intragroup relationship as a function of expertise.

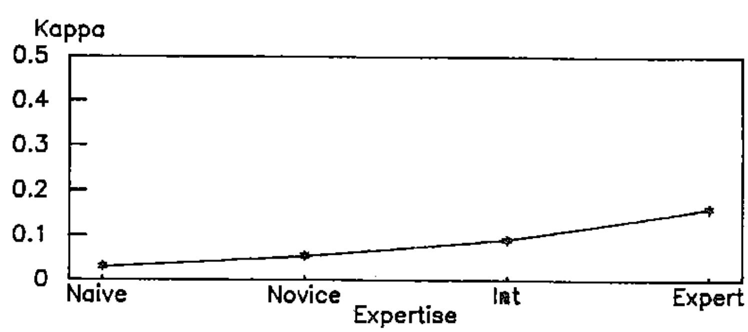

Figure 4. Intragroup relationship as a function of expertise. (Intermediate groups combined.)

interpreting links which are present in each network until determining if the nature of the link remains unchanged between levels of expertise [15].

The effectiveness of this type of representation for computer menu design has been suggested [14]. Its potential utility in the development of expert systems has also been documented [15]. With the advent of more intelligent computer interfaces, it may also be possible to use network representations of various levels of expertise to control adaptive interfaces, interfaces which change as a function of user behavior. Some interface elements which could be controlled in this way are menu organization, the content of on-line help and error messages, and hypertext.

The analysis of intragroup agreement in this study support the results of McKeithen et al. [5] rather than Cooke [19] since the data could be represented as a monotonically increasing curve (see Figure 4), but not a U-shaped one. Both McKeithen et al. and we used similar conceptual materials (ALGOL W reserved words and IBM PC DOS commands) while Cooke used words that were more general in nature (i.e., could be found in a standard English dictionary). Since Cooke's stimuli were more general, naive participants could have better intra-group agreement since they would have a common basis for judgement of similarity, unaffected by knowledge of the use of the words in the domain of computer programming. Naive participants rating words such as IBM PC DOS commands or ALGOL W reserved words would have little common basis for judging similarity, and would show poor intragroup agreement.

It is difficult to explain the S-shaped function (Figure 3) with the same simple mechanisms used to explain the monotonic or U-shaped functions. Despite the statistical significance of the drop in the curve, it may simply be the result of random variation. On the other hand, it is possible that the relationship between intragroup agreement and expertise is more complex than previously indicated. For example, if participants begin learning the relationships among concepts, but have no common basis for similarity judgement, the initial intragroup agreement will be poor. As they gain a common basis through experience, the intragroup agreement would increase. If, at some intermediate point in learning, new organizing principles must be learned, intragroup reliability might decrease since participants may not learn these principles at the same rate. As the slower learners catch up, the intragroup agreement would increase again. In general, a decrease in the intragroup agreement may reflect the introduction of major new organizing principles, while an increase may represent the gradual assimilation of the new principles. Only more studies examining a sufficient number of levels of expertise and exploring other operational definitions of expertise will resolve this issue. Also, future research should focus to some extent on longitudinal rather than cross-sectional experimental designs.

In conclusion, we have demonstrated that the analysis of similarity ratings of computer-related concepts using Pathfinder is consistent with previous findings in the relevant literature. We have also shown a way in which a network representation of concepts with nodes and links can reveal changes in memory structure which would be more difficult to study using other types of multivariate interdependence methods.

References

[1] P. H. Winston, Artificial intelligence, Reading, MA: Addison-Wesley, 1984.

[2] W. G. Chase and H. A. Simon, "Perception in chess," Cognitive Psychology, vol. 4, p. 55, 1973.

[3] D. E. Egan and B. J. Schwartz, "Chunking in recall of symbolic drawings," Memory and Cognition, vol. 7, p. 149, 1979.

[4] C. M. Allwood, "Novices on the computer: A review of the literature," International Journal of Man-Machine Studies, vol. 25, p. 633, 1986.

[5] K. B. McKeithen, J. S. Reitman, H. H. Rueter, and S. C. Hirtle, "Knowledge organization and skill differences in computer programmers," Cognitive Psychology, vol. 13, p. 307, 1981.

[6] A. G. Bateson, R. A. Alexander, and M. D. Murphy, "Cognitive processing differences between novice and expert computer programmers," International Journal of Man-Machine Studies, vol. 26, p. 649, 1987.

[7] W. R. Dillon and M. Goldstein, Multivariate analysis: Methods and applications. New York: John Wiley, 1984.

[8] D. W. Dearholt, R. W. Schvaneveldt, and F. T. Durso, "Properties of networks derived from proximities," New Mexico State University, Computing Research Laboratory, Las Cruces, NM, Tech. Memorandum MCCS-85-14, 1985.

[9] R. W. Schvaneveldt, F. T. Durso, and D. W. Dearholt, "Pathfinder: Scaling with network structures," New Mexico State Univeristy, Computing Research Laboratory, Las Cruces, NM, Tech. Memorandum MCCS-85-9, 1985.

[10] L. J. Rips, E. J. Shoben, and E. E. Smith, "Semantic distance and the verification of semantic relations," Journal of Verbal Learning and Verbal Behavior, vol. 12, p. 1, 1973.

[11] R. W. Schvaneveldt, F. T. Durso, T. E. Goldsmith, T. J. Breen, N. M. Cooke, R. G. Tucker, and J. C. De Maio, "Measuring the structure of expertise," International Journal of Man-Machine Studies, vol. 23, p. 699, 1985.

[12] N. M. Cooke, F. T. Durso, and R. W. Schvaneveldt, "Measures of memory organization and recall," New Mexico State University, Computing Research Laboratory, Las Cruces, NM, Tech. Memorandum MCCS-85-11, 1985.

[13] J. E. McDonald, J. D. Stone, L. S. Liebelt, and J. Karat, "Evaluating a method for structuring the user-system interface," in Proceedings of the Human Factors Society 26th Annual Meeting, p. 551, Seattle, WA: Human Factors Society, 1982.

[14] K. M. Snyder, K. Paap, J. R. Lewis, J. Rotella, A. Happ, L. Malcus, and J. Dyck, "Using cognitive networks to create menus," International Business Machines, Inc., Boca Raton, FL., Tech. Report 54.405, 1986.

[15] N. M. Cooke and J. E. McDonald, "A formal methodology for acquiring and representing expert knowledge," Proceedings of the IEEE, vol. 74, p. 1422, 1986.

[16] J. R. Olson and H. H. Rueter, "Extracting expertise from experts: Methods for knowledge acquisition," Expert Systems, vol. 4, p. 152, 1987.

[17] R. E. Nisbett and T. D. Wilson, "Telling more than we can know: Verbal reports on mental processes," Psychological Review, vol. 84, p. 231, 1977.

[18] N. M. Cooke and J. E. McDonald, "The application of psychological scaling techniques to knowledge elicitation for knowledge-based systems," International Journal of Man-Machine Studies, vol. 26, p. 533, 1987.

[19] N. M. Cooke, Memory structures of expert and novice computer programmers: Recall order vs. similarity ratings. Unpublished master's thesis, New Mexico State University, Las Cruces, NM, 1983.

[20] D. S. Kay and J. B. Black, "Changes in knowledge representation of computer systems with experience," in Proceedings of the Human Factors Society 28th Annual Meeting, p. 963, San Antonio, TX: Human Factors Society, 1984.

[21] P. Dunn-Rankin, Scaling methods, Hillsdale, NJ: Erlbaum, 1983.

[22] J. L. Fleiss, Statistical methods for rates and proportions. New York, NY: John Wiley, 1981.

[23] J. L. Myers, Fundamentals of experimental design. Boston: Allyn and Bacon, 1979.