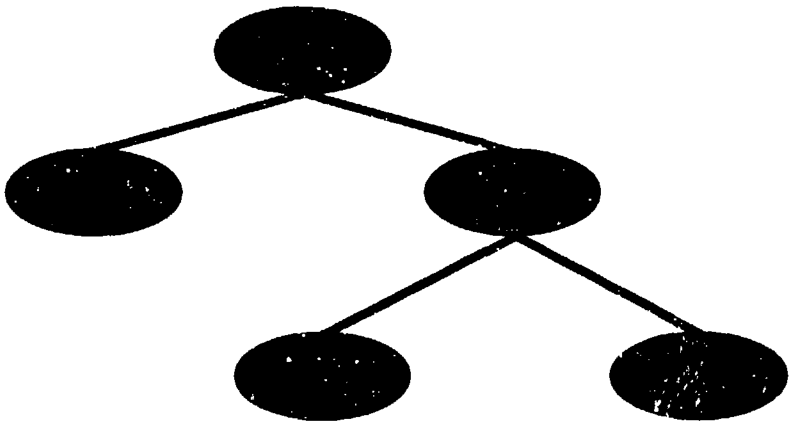
Toward a Model of Expert Knowledge Structure and Their Role in Cognitive Task Performance



Richard Koubek

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-92-J-1153, Contract Authority Identification No. NR4421558-04. The opinions expressed here do not necessarily reflect the position of ONR.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

[illegible]

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE November 1, 1993	3. REPORT TYPE AND DATES COVERED Final Report - Nov 89-June 93

4. TITLE AND SUBTITLE
Toward a Model of Expert Knowledge Structures and Their Role in Cognitive Task Performance

5. FUNDING NUMBERS
N00014-92-J-1153
R&T: 4421558-04

6. AUTHOR(S)

Richard Koubek

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

School of Industrial Engineering
Purdue University
West Lafayette, IN 47907-1287

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Cognitive Science Program
Office of Naval Research (Code 1142CS)
800 North Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

This project addressed the role of knowledge structure in skilled cognitive task performance. In particular, efforts focused on four areas: identification of attributes affecting the development of knowledge structures, derivation of theoretical dimensions for the knowledge structure construct, mapping the impact of knowledge structure on task performance and derivation of automated knowledge structure measures. This project has provided information useful for the selection and training of personnel, suggesting a focus on selecting and training individuals to develop the appropriate knowledge structure. Both the type of knowledge structures and examples of effective training methods are provided. Key unanswered questions include (1) the development of a comprehensive matching algorithm between particular task features and the type of knowledge structure which is best suited for these task attributes, and (2) identification of individual difference characteristics which influence the type of knowledge structure developed.

14. SUBJECT TERMS
Knowledge structure, skilled task, performance, expertise

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	

Toward a Model of Expert Knowledge Structure and Their Role in Cognitive Task Performance

Accesion For		
NTIS CRA&I	X	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

Richard Koubek

Purdue University
West Lafayette, Indiana 47907

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

This project addressed the role of knowledge organization in skilled cognitive task performance. In particular, this work focused on three areas for furthering the understanding and implementation of knowledge structures. These phases can be represented in figure 1. From the figure, the three main aspects are identified as follows: attributes effecting the development of knowledge structures, the knowledge structure construct and the impact of knowledge structure on task performance. In addition, measures of knowledge structure are explored Activities in this grant address these key aspects.

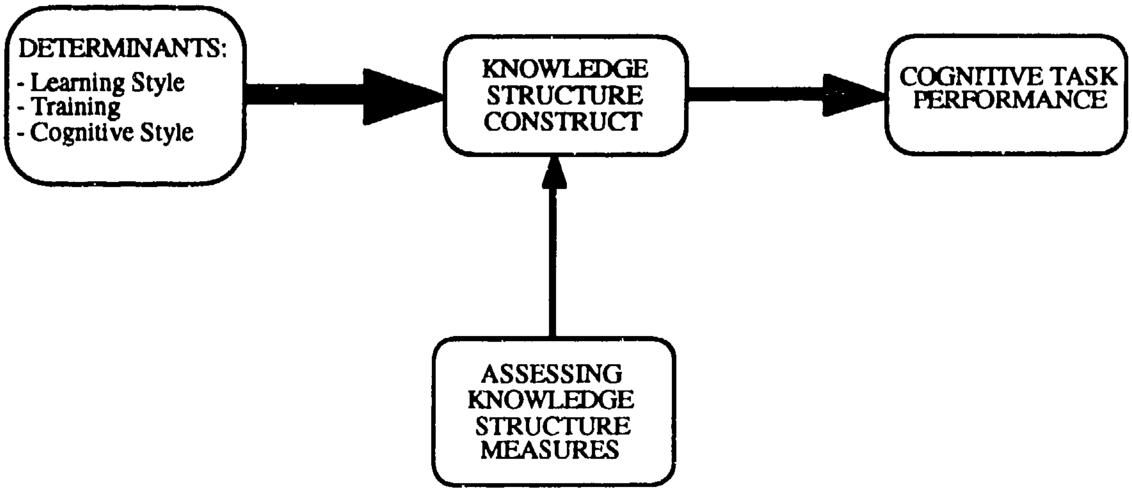


Figure 1. Knowledge Structure Domain.

The work performed under this contract is divided into three phases. The first phase can be considered exploratory, in which all three aspects of the model presented in the figure are examined simultaneously. Activities in the second phase resulted from deficiencies identified in the first phase regarding the current understanding of knowledge structure (KS) which constrained progress toward the desired objectives of this contract. Efforts in this phase were directed at developing a model of KS as well as suitable measurement techniques for the model dimensions. Finally, the third phase focused on two issues. First, determinants of an individual's KS were explored, and a new procedural-oriented model of KS was developed to account for more fine-grained behaviors.

Phase I: Initial Explorations

The first study performed under this grant set out to explore how knowledge structures can impact performance on controlled and automatized tasks. (Koubek and Mountjoy, 1991, Attachment 1). While previous research has shown an effect of KS on performance for controlled tasks, little research exists for automatized tasks. In particular, training paradigms for automaticity would suggest repeated practice on consistent task components, whereas training for knowledge structure development would lead to a more global training approach. As such, this effort attempted to define the scope of further work, more specifically, whether it is fruitful to consider knowledge structures as a performance determinant for tasks comprised of primarily consistent components.

As part of this study, two training scenarios were employed to produce groups with different knowledge structures: abstract and concrete. Also, field independence/field dependence was included as part of the design to account for variability in knowledge structure development resulting from this type of individual difference.

Results of the study support the view that KS is a high level determinant of operator performance on cognitive-oriented tasks. While it has been previously suggested that training to develop automatization should focus on consistent trials, this research has shown that initial performance and the rate of learning on these straightforward tasks are a function of knowledge organization. In addition, the KS was found to influence final performance on complex tasks and the strategies used to complete these tasks. However, this study was unable to identify training or cognitive style as factors which would lead to the different knowledge structures.

Phase II: Models and Measures of Knowledge Structure

One of the primary difficulties encountered in the first study was measurement of the knowledge structure. When searching for a suitable measure, it became immediately obvious that there was a lack of an explicitly defined model of the knowledge structure construct. Ideally, the underlying dimensions, or attributes, of knowledge structures would be clearly identified in order that measurement tools can be compared on the degree to which they can detect changes in these attributes. Therefore, in the second study (Koubek and Mountjoy, Attachment 2), we set out to define a set of human knowledge structure dimensions and provide some initial evidence for the existence of these dimensions.

To begin, a working definition of knowledge structure was stated as follows: Human knowledge structures are defined as the structure of interrelationships between concepts and procedures (elements) in a particular domain, organized into a unified body of knowledge. In addition, a preliminary effort to identify the parameters, or features, which would define the KS was undertaken. The above definition suggests two main components, of KS: elements and their interrelationships. These elements can be further described through two additional dimensions, or parameters. First, the knowledge structure elements may be either declarative or procedural concepts, and second, these elements can exist at various levels of abstraction. A cursory examination of concept lists used in previous KS research indicates both declarative and procedural concepts are used for relationship elicitation. Also, existing work has shown elements at varying degrees of abstraction.

In describing the second major component, relationships, two additional parameters can be identified as well. First, a relationship between elements can vary in degree, and second, more than one type of relationship can simultaneously exist between two or more elements. Obviously, the primary input to many KS measurement techniques is the proximity matrix for which varying degrees of relatedness is the basis. Also, techniques such as overlapping closed curves, multidimensional scaling and repertory grid have indicated the existence of multiple relationships between elements. The results of this study provided evidence for the existence of these dimensions.

With these dimensions explicitly listed, the second part of the research in this phase provided a comparative review of the existing KS measurement techniques in order to develop an intelligent basis for selection of assessment tools for later research. Three general classes of techniques were identified: verbal reports, clustering and scaling. Each group was reviewed according to: (1) the manner in which concepts and relationships are elicited, (2) the method used to derive the knowledge structure and (3) procedures used to analyze the knowledge structure (Benysh, Koubek and Calvez, 1993; Attachment 3). In addition, empirical data was collected on six popular techniques to examine how well they could measure the particular attributes put forth in the KS model. When comparing novice and expert representations, no techniques identified differences between the number of procedural and declarative concepts, and none identified differences in the number of multiple relations between concepts. Of particular interest, the techniques which transform the data to reveal latent structures, such as Hierarchical Clustering and Pathfinder, were the most effective.

At the conclusion of this second phase of the research, a model of KS attributes had been developed and existing techniques were examined to determine the degree to which they could assess these dimensions.

Phase III: Knowledge Structure Determinants and Procedural Structure Measures

With the model and measures in place, activity could now move forward and begin to explore how various KS attributes are developed within an individual. In phase I of this grant, this issue was explored in a global manner and we were unsuccessful at identifying attributes which determined the development of KS. This was attributed primarily to the weak measurement methods employed and lack of an understanding of the underlying dimensions being assessed. Work in Phase II provided the dimensions. A significant effort in this final phase was the development of KSAT (Knowledge Structure Assessment Technique) which combined many of the key attributes of existing KS assessment techniques and added additional assessment dimensions to measure all aspects of the KS model. With this automated tool in hand, the first study in phase III attempted to determine whether the development of an individual's knowledge structure in a particular domain can be manipulated through training. The experiment (Koubek, Clarkston and Calvez, in press; Attachment 4) utilized the manufacturing domain of plastic extrusion machine operation. Sixteen subjects, having no previous knowledge of the domain, were randomly assigned to one of two experimental groups. Each of the experimental groups corresponded to a distinct training condition. Over a three day period, both training groups received the same instructional content, however the sequence in which the training material was presented differed. One group initially received the abstract, conceptual relationships between domain concepts, followed by more detailed relationships associated with the lower level aspects of the domain. The other group received the training material in the reverse order; i.e. the lower level information followed by the abstract. Prior to and concluding the training sessions, each individual's knowledge structure was assessed along two dimensions, Hierarchical Levels and Multiple Relations, through the computer-based measurement technique entitled KSAT. The group which received the abstract relationships first showed significant improvement following training along both dimensions of knowledge structure. No significant changes in the knowledge structure dimensions were found for the group which received the lower level relationships first. This study suggests that an individual's knowledge structure can be manipulated through training, with a significant effect being attributed to the training sequence of abstract material followed by the more detailed material.

Following this, a second study (Clarkston and Koubek, Attachment 5) was performed with the cognitive style dimension Speed of Closure to examine its impact on the development of knowledge structures, interaction with training condition and its impact on final task performance. This study provided evidence that cognitive style and knowledge structure are fairly independent constructs. It appears from regression analyses that the effects of KS and Speed of Closure on task performance are actually additive. This would explain the lack of an effect for cognitive style found in Phase I of this research.

Toward the later stages of this research project, it became evident that a significant missing piece in KS research was the incorporation of procedural aspects of human knowledge. While procedural concepts may be used, capturing the essence of procedures in the traditional static KS approaches is highly inefficient. Of concern is the ability to predict fine-grained performance. Current KS research allows one to predict to groups, such as level of expertise, or particular task strategies.

The final study examined relevant literature in the various knowledge representation fields related to cognitive skilled task performance, including knowledge organization, exploratory cognitive models and functional cognitive models. The results of this review provided a compilation of a list of attributes common in many of the models, which were then specified as desirable in a new multi-factor model. This multi-dimensional approach was adopted due to the assumption that the factors associated with the various models are complimentary and that this integration would provide a more complete understanding of skill and expertise.

Following this, the study proposed a hybrid model of procedural knowledge structure representation, designed to unite these attributes into a common framework. This model embodied a number of structural dimensions which were hypothesized as being indicative of aspects of cognitive performance. A set of procedural knowledge structure measures were then defined based on these structural dimensions. Finally, this study validated that model with respect to performance on skilled cognitive task performance and traditional knowledge structure measures.

This validation indicated that the procedural model of KS is capable of predicting aspects of performance as well as expertise. There is also suggestive evidence that it can account for more of the variance than traditional KS measures when predicting errors and error rates during performance. Further, results also indicate that the model contributes to the predictive capability of the traditional KS model with respect to performance and expertise level. The procedural model independently did not perform better than the traditional model of KS when making predictions of expertise level.

In conclusion, work in this contract set out to explore the role of knowledge structures in skilled cognitive task performance. Probably the most significant accomplishments are those which were not initially planned: development of a KS model and a new assessment technique. In addition, this work has provided information directly useful for the selection and training of personnel, suggesting a focus on selecting people with the appropriate knowledge structures and/or training to develop these knowledge structures. Both the type of KS and examples of effective training methods have been provided. A key unanswered question in transferring this research to an applied setting is the relationship between various KS dimensions and different task attributes. It is likely that a match exists between particular task features and the type of KS which is best suited for these task attributes. This matching is yet to be determined. In addition, this work was unable to identify the individual characteristics which influence the type and ease of KS developed. These two questions remain open for future research

References

Koubek, R.J. and Mountjoy, D.N. (1991). "The role of training, individual differences and knowledge representation in cognitive-oriented task performance," *International Journal of Human-Computer Interaction* , 3(1), 31-41.

Benysh, D.V., Koubek, R.J. and Calvez, V.C. (1993). "A comparative review of knowledge structure measurement techniques for interface design," *International Journal of Human-Computer Interaction*, 5(3), 211-237.

Koubek, R.J., Clarkston, T.P. and Calvez, V. (in press). "The training of knowledge structures for manufacturing tasks: An empirical study," *Ergonomics*, 28 manuscript pages. Accepted 9 September 1992.

ATTACHMENT 1

The Impact of Knowledge Representation on Cognitive-Oriented Task Performance

Richard J. Koubek and Daniel N. Mountjoy

In: International Journal of Human-Computer Interaction 3(1) 31-47 (1991)

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1265 and # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

The Impact of Knowledge Representation on Cognitive-Oriented Task Performance

Richard J. Koubek
School of Industrial Engineering
Purdue University
West Lafayette, Indiana 47907-1287 USA

Daniel N. Mountjoy
Department of Biomedical and Human Factors Engineering
Wright State University
Dayton, Ohio 45435 USA

Running Title: Knowledge Representation

Acknowledgements: This research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-90-J-1256, # N00014-92-J-1153, and the Wright State University Research Challenge program. The opinions expressed in this paper do not necessarily reflect the position of ONR. The authors are indebted to Robin Stern and Kristen Barthelemy for their assistance in data collection.

ABSTRACT

This research examines the impact of training style and operator individual differences on the task representation developed, automatized task performance and controlled task performance. Results indicate that performance on relatively straightforward repetitive tasks usually associated with automatization is influenced by training style and the mental task representation held by operators. Also, domain representation is a significant determinant of performance on complex cognitive-oriented tasks requiring controlled processes. Therefore, the task representation is identified as a high level performance determinant for both simple and complex task performance. No effect for training style or individual differences was found. It is concluded that training programs for systems requiring human-computer interaction must account for this factor in order to facilitate the learning process and enhance task performance.

INTRODUCTION

As the United States begins to experience the effects of the "baby bust" and an increased emphasis on college training, rather than vocational, the labor pool from which to draw workers decreases. In addition, computer skill levels required for these jobs are increasing, thus further reducing the number of potentially qualified applicants. As more businesses become automated, persons with little computer experience are suddenly thrust into a virtual "computer world". Given this scenario, an effort must be made to determine those elements of training which will aide in the acquisition of skill on computer-based tasks.

In general, at least three primary factors exist, related particularly to the individual, which can impact productivity in cognitive-oriented work: previous learning (including training), present knowledge of the task and individual differences. While externally determined factors, such as work schedule and compensation are important, the present paper focuses on the internal factors.

It appears certain that the mental representation of the system one holds is a determining factor in the ability to solve complex problems. For example, Kieras and Bovair (1984) performed a study concerned with the importance of mental models in learning to operate an unfamiliar piece of equipment (a basic control panel). It was shown that the group trained with the mental model learned and executed the procedures more quickly, had superior retention and simplified inefficient procedures more often than the group trained without the model. In the realm of computer programming, Mayer (1989) suggests that the presentation of a concrete model early in a novice programmer's training program may have beneficial effects on his or her encoding and use of new technical information. Adelson (1981) demonstrated that expert programmers reorganize randomized computer code in a hierarchial structure, while novices group code according to its syntactic similarity. Likewise, McKeithen, Reitman, Reuter and Hirtle (1981) studied the memory strategies of novices and experts. In the reproduction of a computer program, novices used general mnemonic strategies, such as an alphabetic strategy, while experts used a more specific strategy of grouping the words according to their functions. While the results of Mayer's study would tend to favor the incorporation of a concrete model in training a novice programmer, the results of the Adelson and McKeithen et al. studies would suggest the use of more hierarchial representations for attaining higher levels of skill.

Rather than focus on a single training style, recent research has suggested the need to tailor training programs to individual differences. Sein and Bostrom (1989), for example, have found that people with an "abstract" learning style will perform significantly better when provided a training program which emphasizes the abstract features of the domain, while those with a "concrete" learning style perform almost twice as well when given an analogical (concrete) training program compared to the abstract training program. Some researchers have attempted to train expert and novice programmers to form a particular mental representation. For example, Adelson (1984) showed that novices could be forced into a semantic representation (as opposed to their preferred syntactic representation), and experts could be forced into a syntactic representation (as opposed to their preferred semantic representation). However, these representations proved unstable, and both groups eventually switched back to their more "natural" representation. It would appear that, as implied by the results of Kolodner (1983), Murphy and Wright (1984), Novick (1988), and Koubek and Salvendy (1989), this "change-over" from a concrete mental representation of the novice to an abstract representation of the expert will occur over time, only as the novice gains additional experience in a particular task domain.

Others have shown the importance of a breadth-first, compared to a depth-first training style for final performance on troubleshooting tasks. Zeit and Spoehr (1989) concluded that the degree of hierarchial structure within a learning tool is reflected in the structure of the learner's knowledge representation. In addition, a hierarchically organized knowledge base, along with applied practice, will lead to procedural representations, while subjects who lack a hierarchial knowledge base will not develop procedural representations. In support of the interplay of one's knowledge structure and the performance level on a given task, Lambert and Newsome (1989) studied the impact of question format and organization presented by an intelligent system on the problem-solving performance of experts (high-skill employees) and novices (low-skill employees). The results provide further evidence that experts and novices organize conceptual knowledge of a problem in different manners. When questions were posed by the system requiring concrete information organization, low-skill employees performed significantly faster than when the questions required abstract information organization. Additionally, high-skill employees performed faster in response to questions which required abstract information organization as compared to concrete information organization. These findings may have far reaching implications in the development of expert systems, as well as in training novices to program and debug efficiently.

Another emphasis in the literature suggests training to develop automatic processes (Fisk and Gallini, 1989). This is supported by Wiedenbeck (1985) who found that, even in simple, automatized tasks, experts are significantly faster and more accurate than novices. The Dual Processing Code Theory (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977) suggests that two qualitatively different cognitive processes exist: controlled and automatic. The key distinguishing features between these two processes include the use of cognitive resources and control. Automatic processes have been found to require little or no cognitive resources and are subject to lack of control over their processing. Controlled processing is resource intensive and monitored, or controlled, by the cognitive system. Therefore, training individuals on consistent task components is suggested to rapidly attain high levels of automatization.

Another key distinction between the two processes lies in their development. Initially, all newly acquired processes are controlled. However, as practice accrues on consistent task components, then the processes associated with executing these consistent components become automatized, and automatized processes require no cognitive resources. Non-consistent task components, however, must be executed with controlled processes, which are resource intensive. Therefore, if one were able to be trained to automatically process certain cognitive information (i.e. computer code), thereby requiring less cognitive resources for the performance of that particular cognitive task, the speed and efficiency of task performance would increase.

Finally, cognitive style may also be a determining factor of one's asymptotic skill level for computer-oriented tasks. For example, the cognitive style of field independence is "definable in terms of degree of dependence on the structure of the prevailing visual field, ranging from great dependence, at one extreme, to great ability to deal with the presented field analytically, or to separate an item from the configuration in which it occurs, at the other" (Witkin, Lewis, Hertzman, Machover, Meissner and Wapner, 1954). In a study which examined student and professional programmers' cognitive representations of software, Holt, Boehm-Davis and Schultz (1987) found that the mental models formed (which were examined while subjects performed either simple or complex modifications to a program) were affected by problem structure, problem type, and ease of program modification. Specifically, the mental models of the professionals were most affected by modification difficulty, while the mental models of the students were most affected by the structure and content of the programs. This suggests that the professionals may act in a "field independent" manner, since they were less influenced by the surface structure of the program. Conversely, it is possible that the students, who were primarily affected by the surface structure and content of the programs, may be classified as "field dependent". While evidence supports each of the above stated factors as performance determinants, it is beginning to appear that a complex interaction exists between individual differences, training and the current knowledge representation of the task.

Derivation of Hypotheses

The primary focus of the present research is to provide exploratory evidence examining the effect of knowledge representation on performance of two task types, where controlled and automatized processes are required respectively. In an effort to examine these effects on each task type separately, a diad of primary hypotheses are proposed:

Hypothesis One: The task representation affects controlled process task performance.

Hypothesis Two: The task representation affects the development and performance of automatized processes.

From the above review, several other factors, particularly training and cognitive style, have been identified as potential determinants of task performance. To ascertain whether performance effects tested in the first two hypotheses above are confounded with these two variables, the following hypotheses are proposed:

Hypothesis Three: Training and cognitive style affect controlled process task performance.

Hypothesis Four: Training and cognitive style affect the development and performance of automatized processing.

Finally, it is possible that the representation, training and cognitive style are in fact, interrelated. In particular, the final hypothesis of the present study tests whether the training received and individual differences (cognitive style) are determinants of the type of task representation developed by an individual.

Hypothesis Five: Training and cognitive style affect the representation developed.

METHOD

In order to test the above hypotheses, subjects of varying cognitive styles were trained in either an Alphabetical or Hierarchial manner to use a word processor. Following training, their task representation was assessed and they were required to perform both controlled and automatic process tasks.

Task

Subjects were required to perform four tasks: cognitive style assessment, domain training, representation evaluation and stimulus task execution. The cognitive style of Field Independence (FI) - Field Dependence (FD) was used to categorize subjects with respect to their individual differences. Based on the cognitive style theory mentioned previously, one would predict that FI and FD subjects would tend to form conceptually different knowledge representations depending on the structure of their training.

Following the administration of the Hidden Figures Test (Ekstrom, French, Harman and Dermen, 1976) to assess cognitive style, subjects were trained to use a computer word processor (Microsoft Word version 4). Subjects received training on the word processor in one of two ways. One group received the commands arranged alphabetically while the other group received training in which the commands were arranged in a hierarchial manner, based on their functional interrelationships. Each group was given the same commands and examples from which to learn. The only difference was presentation order.

The third task required subjects to complete a representation evaluation form. This form presented 17 learned word processing commands, paired with one another, to yield a total of 136 items. Subjects were asked to rate the degree of similarity on a 5-point Likert-type scale for each pair. This data was evaluated through clustering techniques to identify the subject's representation of the word processing domain.

In order to test hypotheses one through four listed above, two distinct tasks were required which sampled automatic and controlled processes respectively. Therefore, the fourth experimental component required subjects to perform two text editing tasks using the word processing skills learned in the second phase.

In the first text editing task (designated the AP, or automatic processing task), subjects were presented a document and asked to perform a centering task 30 times, once each on evenly spaced lines. This task required repeated execution of a fixed sequence of commands. Since execution of this perceptual-motor based task remained consistent during all trials, it is assumed that automatized processes would develop based on the Dual Processing Code Theory. Each subject repeated the centering task 30 times in one sitting and the time to complete each trial was recorded. Traditionally, research in the automatic processing literature requires subjects to repeat trials well over several hundred times. However, due to the broad-based, exploratory nature of the present study, such a number was inappropriate due to the extensive other testing conditions required. Rather, this data was extrapolated using mathematical modelling techniques (to be discussed later) to simulate a significantly larger number of trials. (From a post-hoc analysis based on the modelling technique, subjects, on average, reached 82 percent of their maximum performance on the AP task within the 30 trials.)

The second editing task required subjects to place two paragraphs side-by-side in the document. The side-by-side procedure required a combination of several separate procedures and was relatively complex. Subjects were allowed 15 minutes to complete this task. Since subjects had not practiced the side-by-side procedure, by definition, this can be classified as the CP (controlled process) task.

Subjects

While 20 subjects volunteered for the experiment, one was eliminated due to her experience with the stimulus task. The remaining 19 (9 male and 10 female) were undergraduate university students from a variety of academic majors with little or no general word processing experience and no prior experience with the present system. Based on their Hidden Figures Test score, subjects were classified as either FI or FD. The national average score on this test, 16, is used as the criterion for placement into groups. Ten subjects had scores below 16 (FD) while nine had scores of 16 or above (FI). These subjects were randomly divided into the training conditions, yielding nine trained alphabetically and ten hierarchically.

Variable Definition and Experimental Design

The independent variables for this study were cognitive style and training method. As described above, the representation was elicited through cluster analysis of similarity ratings. This analysis provided insight into the manner in which subjects group various commands and can suggest evidence regarding the accuracy and completeness of their mental representation. From the cluster analysis, the representation was characterized by the variables listed in Table 1.

Three variables were derived to characterize automatization: alpha, T_0 and T_{1000}. For the AP task, performance was described by the log-linear function $T_n=T_0 n^{-alpha}$. In this equation, alpha represents the rate of learning and T_0 is the time for completion of the first trial. These parameters were derived directly from the data. Using these values, the time for the 1000th trial, T_{1000}, was calculated as an estimate of asymptotic performance. CP task performance was characterized by whether the subject completed the task in the allotted time. Therefore, the dependent variables are T0, T1000, alpha (for the AP task) and whether the CP task was completed. In addition, the representation oriented variables serve as either dependents or independents as a function of the analysis.

The first hypothesis was tested with Chi-Square and discriminant analysis techniques while the second hypothesis, designed to examine the relationship between task representation and automatization, was performed with canonical correlation and multiple regression procedures. Due to sample size restrictions, the third hypothesis was tested with a Chi-Square procedure. A 2x2 ANOVA design was used to test the hypotheses four and five. The independent variables were cognitive style (FI versus FD) and training (Alphabetical versus Hierarchial). The dependent variables for each analysis were those derived from the representation and automatized process tasks respectively.

Procedure

Prior to the training phase, subjects were administered the Hidden Figures Test to determine their cognitive style and assigned to the appropriate training group. An attempt was made to evenly distribute FI and FD subjects into the training conditions. In the second phase, subjects received their respective training modules. During training, subjects read hard-copy descriptions of each command and were required to practice each command with the actual system before proceeding. Subjects were allowed as much time as necessary.

Upon completion of training, subjects were given the representation evaluation form and asked to bring the completed form back the next day, when they would perform the stimulus tasks. On the day following training, subjects were allowed to re-familiarize themselves with the system and then perform the AP and CP tasks in that order. The subjects were provided with a keyboard and a mouse as their computer interface for the two tasks. Their training manuals were also furnished for assistance. A concurrent verbal report was required of the subjects throughout the CP task. Each subject was allowed 15 minutes to complete the CP task. The testing session was video taped for later analysis. Following testing, a second representation evaluation form, identical to the first, was completed to identify any possible changes in knowledge representation.

RESULTS

Presentation of hypotheses in the previous section was structured to represent the logical derivation and rationalization for each hypothesis in a top-down fashion. When presenting the results, it is more efficacious to provide a bottom-up description since the findings build upon one another.

Training and Individual Difference Effects

Representation Development. *Hypothesis: Training and cognitive style affects the representation developed.* The potentially complex interactions of various representation variables warrant multivariate analysis procedures, therefore, a 2X2 MANOVA was performed with cognitive style and training as the independent variables. (Vertical commands, horizontal commands and font purity were not included in this analysis due to their non-normality). The dependent variables were derived from the cluster analysis as described previously. No significant effects were found for training, cognitive style or the interaction. Therefore, this hypothesis is not supported.

Automatized Task Performance. *Hypothesis: Training and cognitive style affect the development and performance of automatized processes.* As described above, the variables used to characterize automatized task performance are alpha, T_0 and T_{1000}. As might be expected, these variables are all significantly correlated with each other at the $p<.02$ level. For the purposes of this experiment, each variable is examined independently through a 2x2 ANOVA procedure with training and cognitive style serving as the independent variables. Regarding initial performance, T_0, while no main effects occur, a significant interaction between training and cognitive style is evident ($F(1,15)=7.04$; $p<.018$). See Table 2 for these results. With regard to learning rate (Table 3), once again, a significant interaction occurs ($F(1,15)=6.45$; $p<.023$). The highest values (or fastest learning rate) are found for FD subjects with Alphabetic training (significantly different from all other means at $p<.05$) while the lowest learning rate occurs for FD subjects with Hierarchial training. There appears only a slight trend for FI subjects to acquire automatized processes more quickly with a hierarchial representation (see Figure 1).

From Figure 1, it appears that FD subjects perform better initially when presented with hierarchial training ($p<.05$ level; Newman-Keuls test). After practice, however, the computed T_{1000} value indicates that final performance for FD subjects is best served with the Alphabetic training rather than Hierarchial training ($p<.05$). Neither main effects nor the interaction were statistically significant for T_{1000}.

Apparently, the hierarchial representation training allows FD subjects to more quickly orient to the problem. As expected by the definition of Field Independence, performance of subjects classified into this group appears unaffected by the training style. Further research is needed to examine this issue more clearly.

Controlled Task Performance. *Hypothesis: Training and cognitive style affect controlled process task performance.* On the CP task, six of the 19 subjects found the correct solution within the allotted time of 15 minutes. Due to the limited number of those completing the task, the effect of training and cognitive style on CP task completion were analyzed separately using Chi-Square procedures. The statistic was identical for both variables: $^2=.693$; $p<.405$. Therefore, this hypothesis cannot be confirmed.

Representation Effects

Automatized Process Task Performance. *Hypothesis: The task representation affects automatized processes.* To evaluate hypothesis four, a canonical correlation was first performed on the representation variables (excluding overall cluster purity since it is a linear combination of existing variables) and the automatization variables listed previously. This multivariate procedure determines the relationship between two sets of variables (SAS Institute, 1988). Results indicate a statistically significant correlation between the two groups. The Squared Canonical Correlation is 0.82, which is significant at the $p<.02$ level. It is therefore suggested that a relationship exists between the representation subjects possess and their performance on automatized tasks, supporting hypothesis two.

In an effort to examine this relationship in more detail, three stepwise multiple regression analyses were performed on alpha, T_0 and T_{1000} respectively, using the representation variables as independent variables. From this analysis, the rate of learning (alpha) can be predicted by the Maximum Distance Between Clusters and the Purity of Font Cluster variables (using 0.15 as the entry and removal criterion). With these two independent variables, 31.55 percent of the variance in alpha can be predicted ($F(2,16)=3.69$; $p<.048$).

In addition to the above two independent variables, the regression equation to predict initial time to perform the AP task, T_0, includes the Number of Horizontal Commands Misclassified. This is logical since the AP task dealt primarily with horizontal page layout. With these three variables, the computed statistics are as follows: $R^2=.593$, $F(3,15)=7.28$; $p<.003$. No variables met the 0.15 significance level for entry into the model for predicting estimated final performance, T_{1000}. From the above results, hypothesis two is supported and it can be concluded that the knowledge representation impacts AP task performance, at least in the initial stages of developing automatization.

Controlled Task Performance. *Hypothesis: The task representation affects controlled process performance.* For this analysis, subjects were divided into two groups based on whether they successfully completed the task in the allotted time. From this, a Chi-Square analysis was performed using the variables Overall Cluster Purity (grouped as 0-1 and 2-3) and successful or non-successful task completion (see Figure 2). This analysis reveals that there is a significant dependency between these variables ($^2=6.094$; $p<.025$). In order to determine the utility of this finding for predicting performance on CP tasks from knowing the knowledge representation features, a discriminant analysis was performed on the 19 subjects. Using the computed discriminant function with Overall Cluster Purity, 89.13 percent of the subjects were correctly classified as successful or unsuccessful. More specifically, one successful subject and three unsuccessful subjects were misclassified. With only one variable, the accuracy of this discriminant function supports hypothesis one, that representation significantly influences controlled task performance.

In order to determine the CP solution strategies of the subjects, a GOMS analysis was performed on the verbal protocol data (selection rules were not obtained in this analysis). A "master" GOMS solution containing a set of three goals and their coinciding methods and operators to accomplish those goals was developed upon which to compare subject solution strategies. From the analysis of the subjects' solutions, three strategies became evident: Direct, Single Branch and Multiple Branch.

Those subjects whose strategies contained no incorrect methods (that is, all methods utilized led the subject closer to the task goal) were classified as Direct. The only deviations of these subjects' solutions from the master GOMS solution were individual "operators" within the chosen methods. Four subjects were placed in this category, and of these, three successfully completed the task.

A second strategy classification is Single Branch. These subjects tended to follow a single solution path, even when that particular path was not leading them closer to the task goal. The classification criterion for this category required the subject to have performed three successive methods (different by no more than one operator) that did not advance the subject closer to the ultimate goal. This pattern may have occurred at any point within the solution set. Five subjects were determined to be Single Branch, and none reached the solution of the CP task.

The final strategy is Multiple Branch. The remaining ten subjects moved from method to method in search of the correct solution pattern (which three subjects located). To be placed in this category, subjects must have implemented 2 or less incorrect methods consecutively, while not following the Direct pattern. It should be noted that, in order to ensure the correct placement of subjects into their respective solution strategy groups, the classification process was performed independently by two raters.

Following the placement of subjects into their respective solution strategy groups, a Wilcoxon's Rank-Sum Test was performed, using Overall Cluster Purity as the ranked variable, in order to determine if representation differences existed between the groups. Purity scores were then tested for each solution strategy group against the purity scores of the other groups independently. The overall purity scores in the Direct group were found to be higher than those in the Single Branch group ($W_s(n_1=4,n_2=5)=11.5$; $p<.05$), and those in the Multiple Branch group ($W_s(n_1=4,n_2=10)=17$; $p<.05$). The overall purity scores appear to be slightly higher in the Single Branch group than in the Multiple Branch group, but the result was not significant. Figure 3 shows the relationship among the group means. From the previous results, it can be seen that subjects with high purity scores tend to utilize a Direct solution strategy when performing a controlled process task, while subjects with low overall purity scores follow a Multiple or Single Branch approach. It is noteworthy to mention that 75% of those subjects with Direct strategies completed the task compared to 30% and 0% of the Multiple Branch and Single Branch Groups respectively. In summary of this section, it appears that task representation affects the solution strategy employed in a complex cognitive task, which in turn is a determining factor of successful task completion. A complete summary of the statistical analyses performed in this study is given in Table 4.

CONCLUSION

While the above results should be viewed in the context of an exploratory analysis, several tentative conclusions can be drawn. Hypothesis five of this study (training and cognitive style affect the representation developed) was not supported. While this may be due to a lack of statistical power in the tests, these results are similar to that of Adelson (1984). The particular training (Hierarchial versus Alphabetical) administered to the subjects, regardless of their cognitive style, did not affect their developed representation. Again, it appears that subjects will maintain their most natural representation.

Significant results were obtained for the effect of training and cognitive style on the development and performance of automatized processes. The fastest learning rate was found for Field Dependent subjects with Alphabetic training, while the slowest learning rate occurred in Field Dependent subjects trained hierarchically. Initial AP task performance of FD subjects was actually aided by hierarchial training, but the final performance was best aided by alphabetic training. This result is related to previous findings which suggest a switch from concrete to an abstract representation as one becomes experienced in a task domain. In particular, an individual is not necessarily an expert simply because a particular task has been automatized.

No evidence was found to support the effect of training and cognitive style on controlled process task performance (that is, whether the subjects finished the CP task in the allotted time). Further research is needed in this area.

Automatized processes were found to be affected by the task representation. More precisely, the rate of learning, alpha, was predicted by two independent variables: Maximum Distance Between Clusters and the Purity of Font Cluster. In addition, the initial time to perform the AP task, T_0, could be predicted with the inclusion of a third independent variable, the Number of Horizontal Commands Misclassified.

Controlled process task performance was also found to be affected by the task representation. Simply by knowing the subjects' task representations (based upon Overall Cluster Purity), approximately 89% of the subjects were correctly classified as to whether they finished the CP task. In addition, the particular strategy utilized to perform the CP task was affected by the task representation. Overall Purity scores were highest for those subjects who approached the task with a Direct strategy, and 75% of those subjects completed the CP task successfully.

The above results support the view that a high level determinant of operator performance on cognitive-oriented tasks exists: domain representation. Previously, it has been suggested that training to develop automatization and high level performance simply requires repetitive practice. However, the present results appear to indicate that, depending on individual operator characteristics, a higher level factor can significantly influence initial performance and the rate of learning on mundane and straightforward tasks which are well suited for automatization. In addition, the task representation influences performance on more complex tasks, including the strategy used to complete them. Since computer-oriented tasks may require both types of performance from operators (automatic and controlled processes), emphasis should be placed on selecting and reinforcing the correct representation for the particular task requirements and individual operator characteristics. However, further research is necessary to determine mechanisms for teaching and reinforcing these representations. The present study did not identify factors which lead to the particular representation developed. With this knowledge, training programs could be targeted to develop representations most suited to the task and operator, thereby decreasing training time and increasing task performance.

REFERENCES

Adelson, B. (1981). The development of abstract categories in programming languages. *Memory and Cognition*, 9, 422-433.

Adelson, B. (1984). When novices surpass experts: the difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 483-495.

Card, S.K., Moran, T.P. and Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ekstrom, R.B., French, J.W., Harman, H.H. and Dermen, D. (1976). *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton: Educational Testing Service.

Fisk, A.D. and Gallini, J.K. (1989). Training consistent components of tasks: Developing an instructional system based on automatic/controlled processing principles. *Human Factors, 31*, 453-464.

Holt, R.W., Boehm-Davis, D.A. and Schultz, A.C. (1987). Mental representations of programs for student and professional programmers. In G.M. Olson, S. Sheppard and E. Soloway (Eds.), *Empirical Studies of Programmers: Second Workshop*. Norwood, NJ: Ablex Publishing Co., 33-46.

Kieras, D.E. and Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, 8, 255-273.

Kolodner, J.L. (1983). Towards an understanding of the role of expertise in the evolution from novice to expert. *International Journal of Man-Machine Studies*, 19, 497-518.

Koubek, R.J. and Salvendy, G. (1989). Cognitive performance of super-experts on computer program modification tasks. Working Paper, School of Industrial Engineering, Purdue University, West Lafayette IN. 47907.

Lamberti, D.M. and Newsome, S.L. (1989). Presenting abstract versus concrete information in expert systems: what is the impact on user performance. *International Journal of Man-Machine Studies*, 31, 27-45.

Mayer, R.E. (1989). The psychology of how novices learn computer programming. In E. Soloway and J.C. Spohrer (Eds.), *Studying the Novice Programmer*. Hillsdale, NJ: Lawrence Erlbaum Associates, 129-159.

McKeithen, K., Reitman, J.S., Reuter, H. and Hirtle, S.C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.

Murphy, G.L. and Wright, J.C. (1984). Changes in conceptual structure with expertise: differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 144-155.

Novick, L.R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 510-520.

SAS Institute, (1988). *SAS/STAT User's Guide*. Cary, NC: SAS Institute.

Schneider, W. and Shiffrin, R.M. (1977). Controlled and automatic human information processing:I. Detection, search, and attention. *Psychological Review, 84*, 1-66.

Sein, M.K. and Bostrom, R.P. (1989). Individual differences and conceptual models in training novice users. *Human Computer Interaction, 4*, 197-229.

Shiffrin, R.M. and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attention, and a general theory. *Psychological Review, 84*, 127-190.

Wiedenbeck, S. (1985). Novice/expert differences in programming skills. *International Journal of Man-Machine Studies*, 23, 383-390.

Witkin, H.A., Lewis, H.B., Hertzman, M., Machover, K., Meissner, P.B., and Wapner, S. (1954). *Personality Through Perception*. New York: Harper.

Zeit, C.M. and Spoehr, K.T. (1989). Knowledge organization and the acquisition of procedural expertise. *Applied Cognitive Psychology*, 3, 313-336.

TABLE 1. Description of Representation Variables.

VARIABLE	DESCRIPTION
Maximum Distance Between Clusters	Provides overall rating of the differentiation between clusters. Low values indicate little distinction between commands.
Total Number of Clusters	Calculated by counting the number of separate clusters that exceed one-half the maximum distance between clusters. Clusters which exceed this value can be considered prominent and significant.
Number of Horizontal Layout Commands Misclassified	Three conceptual clusters exist in the task: horizontal page layout commands, vertical page layout commands and font commands. This variable indicates the number of horizontal layout commands which were incorrectly classified into vertical layout or font clusters.
Number of Vertical Layout Commands Misclassified	See above.
Number of Font Commands Misclassified	See above.
Purity of Horizontal Layout Cluster	Binomial variable which indicates whether subjects had a single cluster which included all the horizontal layout commands and no others. Impure=0 and pure=1.
Purity of Vertical Layout Cluster	See above.
Purity of Font Cluster	See above.
Overall Cluster Purity	Composite value which is computed by summing the individual purity values.
Number of Commands Not Clustered	Provides an indication of domain representation completeness.

TABLE 2. ANOVA results for the effect of training and cognitive style on the development and initial performance (T_0) of automatized processes.

Source	df	Squares	Sum of Square	Mean F	p-value
Training	1	0.3709	0.3709	3.20	0.094
Cognitive Style	1	0.1929	0.1929	1.66	0.217
Interaction	1	0.8162	0.8162	7.04	0.018
Error	15	1.7393	0.1159		

TABLE 3. ANOVA results for the effect of training and cognitive style on the development of automatized processes.

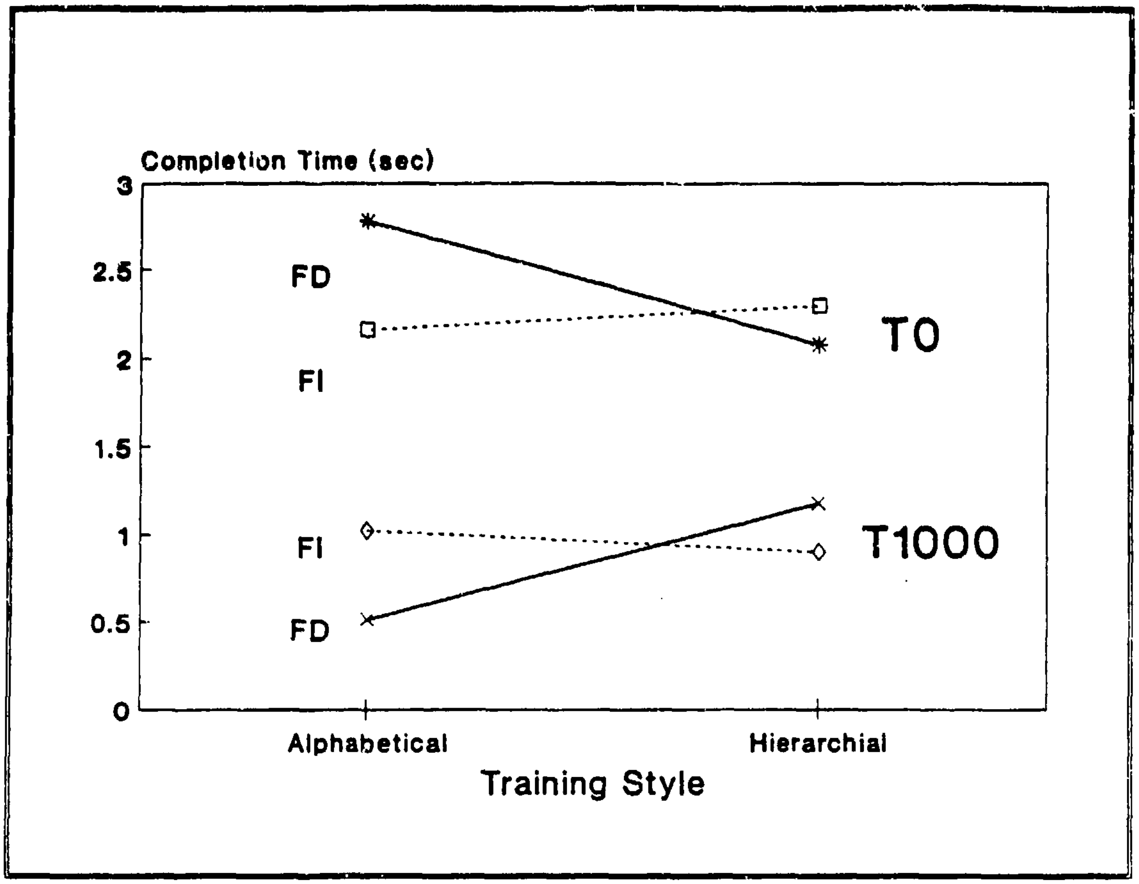
			Sum of	Mean	
Source	df	Squares	Square	F	p-value
Training	1	0.0381	0.0381	3.94	0.066
Cognitive Style	1	0.0091	0.0091	0.94	0.348
Interaction	1	0.0624	0.0624	6.45	0.023
Error	15	0.1453	0.0097		

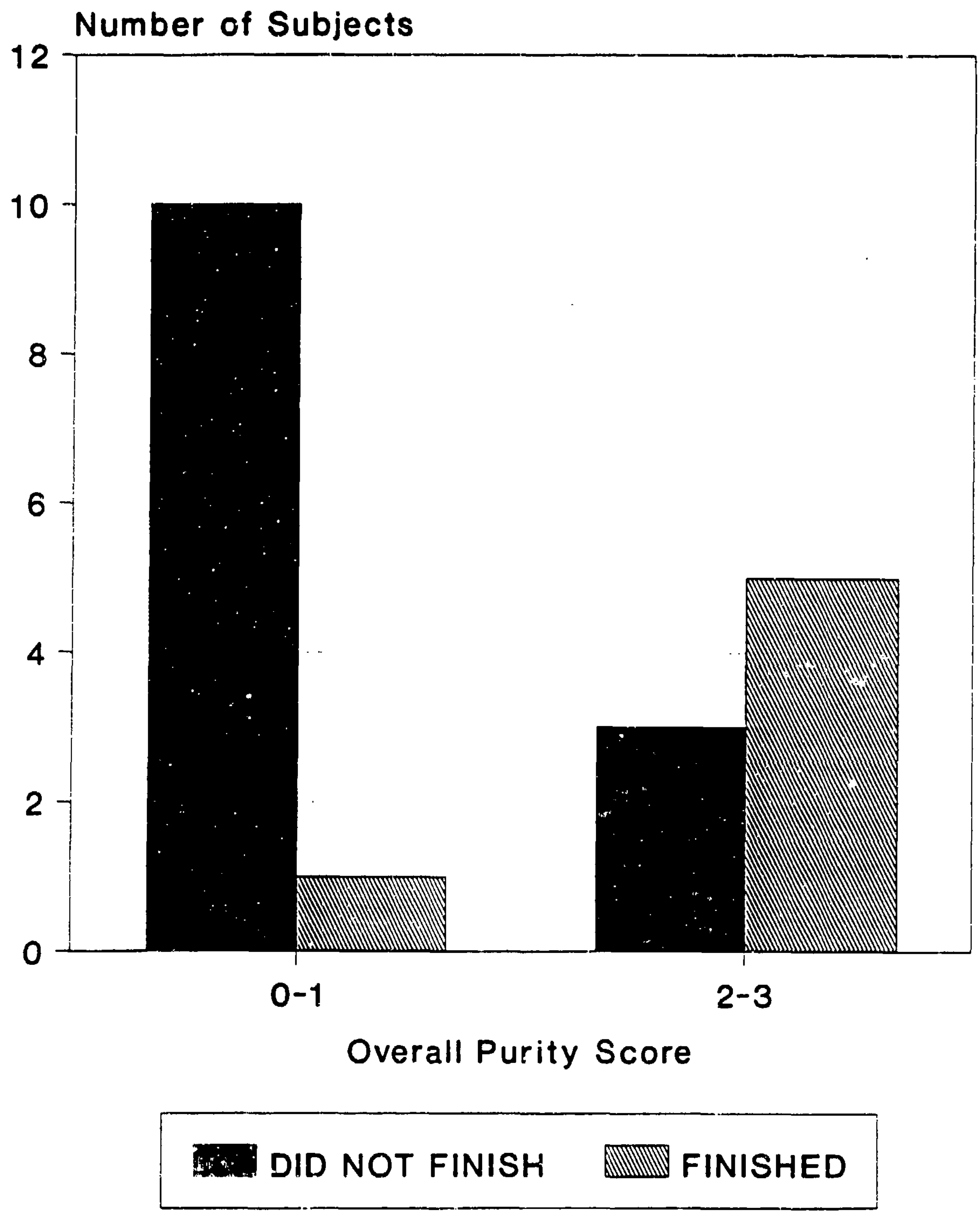
TABLE 4. Summary of Results

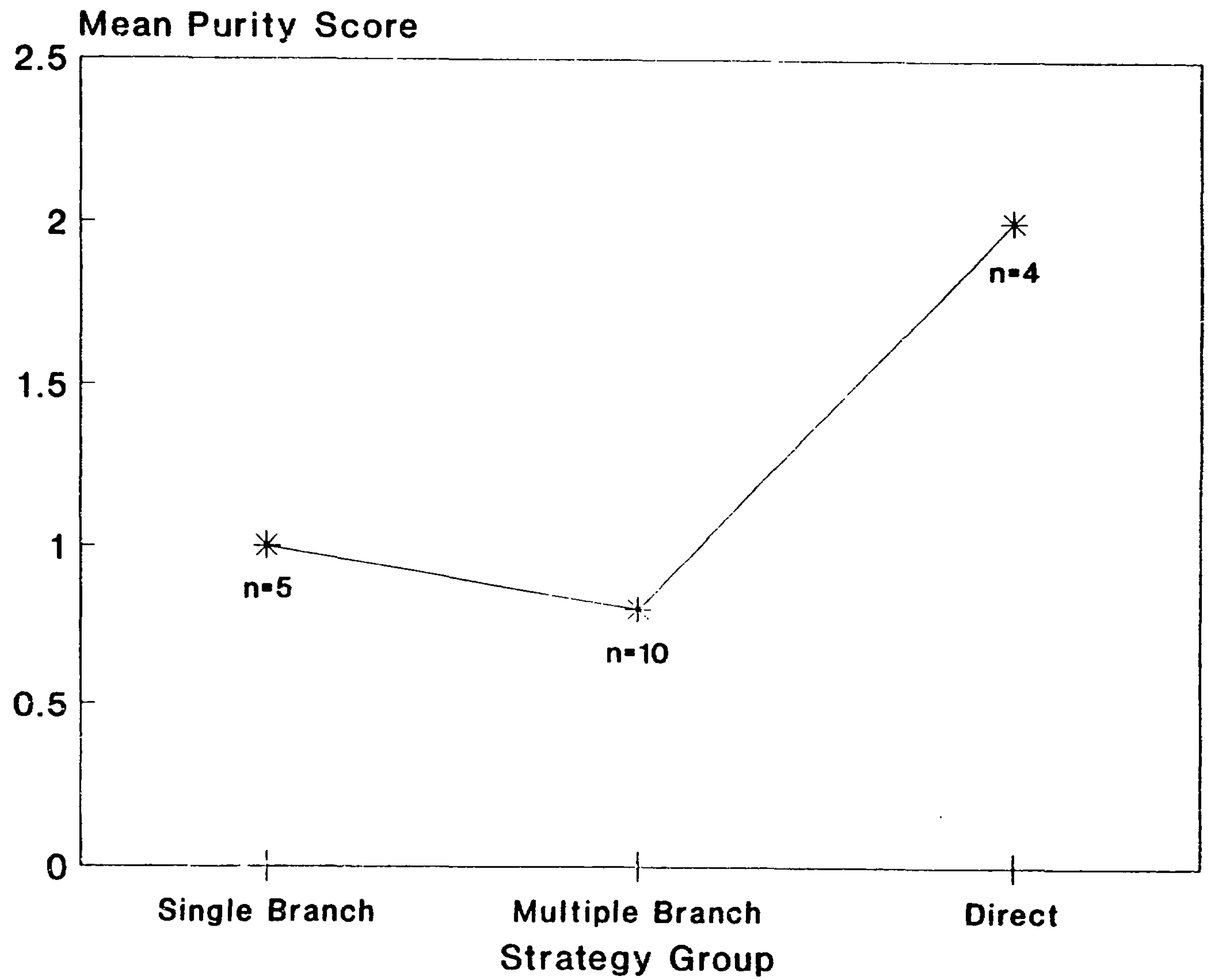
HYPOTHESES	STATISTIC	SIGNIFICANCE
1. Task representation affects controlled process performance		
(a) Representation affects probability of success	Chi-Square	$p<0.025$
(b) Representation affects solution strategy		
• Higher overall purity scores in Direct group than in Single Branch	Wilcoxon's Rank-Sum	$p<0.05$
• Higher overall purity scores in Direct group than in Multiple Branch	Wilcoxon's Rank-Sum	$p<0.05$
2. Task representation affects automatized processes	Squared Canonical Correlation	$p<0.02$
(a) Initial performance (T_0)	Multiple Regression	$p<0.003$
(b) Final performance (T_{1000})	Multiple Regression	N.S.[1]
(c) Rate of learning (Alpha)	Multiple Regression	$p<0.048$
3. Training & Cognitive Style affect controlled process task performance	Chi-Square	N.S.
4. Training & Cognitive Style affect development and performance of automatized processes		
(a) Initial Performance (T_0)		
· Main Effects	2x2 ANOVA	N.S.
· Interaction	2x2 ANOVA	$p<0.018$
· Field Dependent subjects better with Hierarchial training than Alphabetical training	Newman-Keuls	$p<0.05$
(b) Final Performance (T_{1000})		
· Main Effects	2x2 ANOVA	N.S.
· Interaction	2x2 ANOVA	N.S.
· Field Dependent subjects better with Alphabetical training than Hierarchial training	Newman-Keuls	$p<0.05$
(c) Learning Rate (Alpha)		
· Main Effects	2x2 ANOVA	N.S.
· Interaction	2x2 ANOVA	$p<0.023$
· Highest for Field Dependent subjects with Alphabetical training	Newman-Keuls	$p<0.05$
5. Training & Cognitive Style affect representation developed	2x2 MANOVA	N.S.

[1]N.S. = Not significant at $p<0.05$ level.

LIST OF FIGURES


Completion Time (sec)
3
2.5
2
1.5
1
0.5
0
FD
FI
TO
FI
FD
T1000
Alphabetical
Hierarchial
Training Style


Number of Subjects
12
10
8
6
4
2
0
0-1
2-3
Overall Purity Score
DID NOT FINISH
FINISHED


Mean Purity Score
2.5
2
1.5
1
0.5
0
n=5
n=10
n=4
Single Branch
Multiple Branch
Direct
Strategy Group

ATTACHMENT 2

Towards the Development and Validation of a Model for Knowledge Structures

Richard J. Koubek and Daniel N. Mountjoy

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1256 and # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

Toward The Development and Validation of a Model for Knowledge Structures

Richard J. Koubek[1] and Daniel N. Mountjoy[2]

[1]School of Industrial Engineering
Purdue University
West Lafayette, IN 47907-1287

[2]Department of BHE
Wright State University
Dayton, OH 45435

Acknowledgments
This work was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1256 and N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

ABSTRACT

The impetus for this research was the historical lack of an operational definition of knowledge structure, and, likewise, the non-existence of a formally defined model of knowledge structure. Such information is necessary in order to further the advancement of work in understanding skilled task performance by bringing parsimony to the field. A model of human knowledge structure is proposed as a first attempt to identify the parameters of the human knowledge structure. Furthermore, it is shown that the parameters of "Multiple Levels of Abstraction" and "Varying Degree of Relatedness" were affected by representation differences between experience levels. Therefore, since experience is often equated with skill, it may be possible to predict cognitive-oriented task performance based upon the characteristics of these two model dimensions. This has implications for personnel selection, training and job design. Finally, differences between the measurement capabilities of the various knowledge structure measurement techniques were revealed. None of the techniques utilized were able to determine representation differences for each proposed dimension of knowledge structure. Those techniques which transform the data to reveal latent structures, such as Hierarchical Clustering and Pathfinder, were most effective. This points to the need for measurement methodologies capable of eliciting information regarding all knowledge structure parameters. A research agenda in knowledges structures is suggested.

1.0 INTRODUCTION

1.1 Background

The necessity for proper selection, training and job design to insure workers are prepared for the decision-making activities and additional responsibilities brought upon by advancing technology becomes an absolutely critical link in the process of improving global competitiveness for both private corporations and government agencies. Traditional physical testing for manual work and simple cognitive ability tests have become insufficient for such tasks. As a result, work in cognitive science is now making inroads to analyzing and understanding performance on complex cognitive tasks.

Glaser and Bassok (1989) have identified three directions, or research thrusts, categorizing ongoing efforts to understand complex cognitive skills. One focus addresses the acquisition of rapid, proceduralized cognitive processes found in well-practiced consistent tasks. Early efforts in this area (Shiffrin and Dumais, 1981) identified automatization as the process by which declarative knowledge is transformed into compiled procedures. Later work (Anderson, 1982) advanced this effort by not only presenting mechanisms (proceduralization and composition) for initial automation, but also suggesting processes, such as tuning and strengthening, which can account for performance improvement beyond the initial acquisition phase.

A second line of work identified by Glaser and Bassok (1989) is that of metacognition. Here, researchers attempt to identify general control strategies to account for learning and skilled performance. Such effort has been paralleled by work in Artificial Intelligence and expert systems, where the knowledge base and control strategies are separated, but goes beyond that to suggest control mechanisms for determining resource allocation, monitoring performance and predicting outcomes (Simon and Simon, 1978).

A third research focus in skill acquisition is knowledge structure, or knowledge representation. Recent research (Chi et al. 1981, Adelson, 1981, 1984; Koubek and Salvendy, 1989, in press) has shown that the manner in which knowledge is structured is a significant determinant of skilled task performance.

This empirical finding for individuals directly parallels the Knowledge Representation Hypothesis found in Artificial Intelligence literature:

> Any mechanically embodied intelligent process will be comprised of structural ingredients that (a) we as external observers naturally take to represent a propositional account of the knowledge that the overall process exhibits, and (b) independent of such external semantical attribution, play a formal but causal and essential role in engendering the behavior that manifests that knowledge. (Smith, 1982).

In order to move forward in understanding and improving high level skill acquisition, parsimony dictates efforts toward developing theoretical foundations. Glaser (1976) has identified three key elements necessary for such theory development. First, the initial state of knowledge and skills possessed by the individual must be assessed. Second, a description of the desired knowledge state must be determined and finally, the mechanisms and processes which move the individual from the initial state to goal state must be identified. Each of the above research approaches would suggest a different underlying structure to serve as the basis for analyzing the initial and goal states. For example, automatization could employ a production rule system, metacognition may be represented as a set of heuristics and knowledge structure as a set of interconnected elements.

To be consistent with Glaser's (1976) theory development components, it is critical that assessment of initial and desired goal knowledge states are performed in a manner which allows comparative evaluations. Fundamental to comparing knowledge states is the consistency of dimensions describing the states. It is important, at this point, to distinguish between dimensions and the values which are attained for these dimensions. For example, while a dimension titled "number of relationships between two concepts" should be evident in each state when measuring knowledge structures, the actual value assigned to this dimension would be expected to vary. Therefore, foundational to any effort in understanding complex cognitive performance is the availability of a suitable set of dimensions upon which to compare knowledge states. These dimensions, or variables, are often based on empirical or practical issues, such as ease of computation, and determined after the data has been collected (Koubek and Mountjoy, 1991). However, to proceed in an efficient and logical manner, the elements which are to be assessed should be determined a priori on the basis of a model for skilled cognitive behavior such that data collection can then be guided by the dimensions, i.e., model-based measurement.

With regard to knowledge structures, research evidence has begun to show the importance of this construct for understanding performance on complex cognitive tasks. For example, Koubek and Mountjoy (1991) have provided empirical quantitative evidence regarding the impact of human knowledge structure (KS) on performance of multiple task types. This work supports the contention that KS cannot be ignored in repetitive task performance, but rather appears to provide the organizational framework in which subjects can identify task consistencies and eventually develop automatization. One particular hinderance in executing that study was the difficulty in obtaining a measure of KS which yielded assessment variables which could be compared across individuals and groups.

From this, an exhaustive literature review regarding KS measurement techniques was performed (Calvez, Koubek and Mountjoy, 1991). To summarize this review, it appears that a wide variety of KR measurement techniques have been used in the literature, ranging from verbal reports to discriminant functions.

Each KR measurement methodology implicitly makes assumptions regarding the contents, or components, of the human knowledge structure. Rarely, however, are these assumptions made explicit. In fact, few publications go so far as to provide an operational definition of knowledge structure: the object of measurement. According to Goldsmith and Johnson (1990), "Palmer (1978) noted that the field is 'obtuse, poorly defined, and embarrassingly disorganized'(p.259). Although more than a decade has passed since Palmer made these observations, there is ample evidence to suggest that his observations remain valid (p.243)." A potential explanation of this disarray may rest in the fact that specific techniques are designed to measure particular components of the KR, and conflicting outcomes are a result of measuring different elements of the overall construct labeled knowledge structure.

Present understanding of KR has been driven primarily by the individual measurement techniques available and associated mathematical assumptions. While this work has provided a rich empirical base, a comprehensive model of KS is now necessary for further systematic advancement in understanding human knowledge structure. Such a model can yield parsimony, serve as a foundation and impetus for model-based KR measurement techniques and identify KR parameters for predicting operator performance in cognitive-oriented tasks.

1.2 Model of Knowledge Structure

To begin, a working definition of knowledge structure was defined as follows: "Human Knowledge structure is defined as the structure of interrelationships between concepts and procedures (elements) in a particular domain, organized into a unified body of knowledge". In addition, a preliminary effort at identifying parameters, or features, which define the KR model is undertaken. The above definition suggests two main components of knowledge structure: elements and their interrelationships. These elements can be further described through two additional dimensions, or parameters. First, the knowledge structure elements may either be declarative or procedural concepts, and second, these elements can exist at various levels of abstraction. A cursory examination of concept lists used in the previously described research indicates both declarative and procedural concepts are used for relationship elicitation. Also, existing work, (Adelson, 1984; Koubek and Salvendy, 1989) has shown elements at varying degrees of abstraction. For example, critical elements in the domain of computer programming may be "FOR-NEXT" loop and "CONTROL STRUCTURE". The first element is procedural in nature, while CONTROL STRUCTURE is considered more declarative. Also, FOR-NEXT LOOP is a component of CONTROL STRUCTURE and therefore is at a lower level of abstraction.

In describing the second major component, relationships, two additional parameters can also be identified. First, the degree of relationship between elements can vary in degree, and second, more than one type of relationship can simultaneously exist between two or more elements. Obviously, the primary input to many knowledge structure measurement techniques is the proximity matrix, for which varying degrees of relatedness is the basis. Also, techniques such as overlapping closed curves, multidimensional scaling and repertory grid have indicated the existence of multiple relationships between elements. For example, three potential elements in a process control task, FLOW RATE, TEMPERATURE, and VOLUME, may be related to one another in varying degrees and in more than one way. For example, FLOW RATE and VOLUME may be related based on the physical proximity of their displays and/or their interaction with PRESSURE. The proposed model is outlined in table 1.

1.3 Derivation of Hypotheses

The purpose of this study is three-fold. The first issue is to show that the KR dimensions in the model proposed do indeed exist. Next, it is determined whether an individual's experience level in a given cognitive-oriented domain affects the characteristics of the proposed model dimensions, and finally, the capability of existing KR measurement techniques to identify these differences in knowledge structure are compared. Therefore, the following primary hypotheses are proposed:

Hypothesis One: Each proposed dimension of knowledge structure exists.

Hypothesis Two: The characteristics of the proposed KR model dimensions are affected by domain experience.

Hypothesis Three: Differences exist between the measurement capabilities of current KR measurement methodologies across the proposed model dimensions.

The purpose of the first hypothesis is to provide support for the proposed KR model, and in effect, give cause to conduct the remainder of the study. The second hypothesis attempts to reveal which of the proposed model dimensions are a function of experience level in a cognitive domain. Finally, the third hypothesis attempts to expose the effectiveness of available KR measurement techniques, which will point towards areas of those techniques that are in need of future improvement.

2.0 METHOD

This experiment required experienced and naive individuals in a given domain to perform a series of tasks associated with selected KR measurement techniques. Representational differences between the two subject groups were quantitatively assessed based on the output from the knowledge structure measurement techniques.

2.1 Task

In order to test the above hypotheses, the domain used must be cognitive-oriented, and must be broad enough to encompass each of the proposed KR model dimensions. The domain chosen which meets these criteria is that of clerical work. The particular contents of this domain are outlined in the Dictionary of Occupational Titles (DOT 201.362-030).

The particular tasks performed were those necessary to carry out the procedures dictated by the individual KR measurement techniques. Because, at this point in time, our understanding of knowledge structure is dependent upon the measurement techniques currently available, a battery of techniques was used in an attempt to gain a more complete perception of the presence of the proposed model dimensions. The KR measurement techniques selected to be tested are representative of the previously discussed prominent categories. In particular, card sorting, hierarchical clustering analysis (Proc Cluster, SAS Institute, 1988), multidimensional scaling (ALSCAL) and Pathfinder were incorporated. Furthermore, an analysis of the repertory grid technique and pairwise similarity ratings was performed. Verbal report techniques, such as GOMS analysis and problem behavior graphs, were not selected for inclusion in the study because of the subjectivity involved in the derivation of the knowledge structures and their lack of well defined analysis methods.

2.2 Stimulus

The experimental stimulus used in each of the following tasks consisted of 30 clerical domain-relevant concepts which were elicited in pilot study interviews. Six subjects participated in the pilot study, of which three individuals were experienced secretaries and three were naive. Each subject was asked to list as many concepts as possible which they considered to be important to the general secretarial job. These concepts, they were told, may be anything from equipment used to procedures followed throughout the day.

Upon completion of the interviews, one master list of concepts was generated for each subject group. To insure that the test stimulus would consist of concepts representative of both subject groups, the intersection of the master lists was recorded. This process yielded 24 concepts. Since the purpose of the study was to draw out differences between experience levels, an additional six concepts were extracted from the master list of the experienced secretaries. These six were concepts that at least two of the three secretaries held in common. Therefore, the stimulus list was comprised of 30 domain-relevant concepts, 24 in common to both subject groups, and six in common only to the experienced group.

2.3 Subjects

Thirty subjects participated in the experiment. Fifteen were secretaries with at least five years of experience, and the remaining fifteen were considered naive in the field (having a maximum of one year of secretarial experience). The mean experience of the secretaries was 15.20 years with a standard deviation of 7.92 years, while the mean experience of the naive group was 0.20 years with a standard deviation of 0.41 years. Each of the subjects were informed as to the purpose of the experiment and were paid for their participation.

2.4 Experimental Design

The experimental design is a straightforward comparison of experienced and naive group means on each of the dependent variables tested throughout the experiment. However, because of the nature of two of the dependent variables, tests between proportions were performed.

The independent variable in this study is experience level: Experienced and Naive. Because of the nature of the specific KR measurement methodologies used, the dependent variables differ from technique to technique across the model dimensions. Each are quantitative in order to reduce the subjectivity in the analysis, and have been derived from the assumptions of the various techniques or previous literature. The dependent variables are presented in table 2, and are described here for each technique.

2.5 Dependent Variables

2.5.1 Declarative and Procedural Concepts

This particular dimension is independent of the measurement techniques, and is determined by the percentage of procedural concepts listed during a concept elicitation process.

2.5.2 Multiple Levels of Abstraction

Card Sorting: The average number of cards in a pile. A large number of concepts in a pile represents a higher order of abstraction in one's knowledge structure. This is based on the assumption that more concepts present in a pile require a more abstract classification label.

HCS: The average number of concepts in a cluster at the median joining distance. A large number of concepts in a cluster represents a higher order of abstraction in one's knowledge structure. This is based on the same assumption as above.

Pathfinder: The number of stars present in a representation (where a star is defined as a concept with at least five incident links). A larger number of stars represents a higher order of abstraction in one's knowledge structure. This is based on the same assumption as above.

2.5.3 Multiple Relations

Card Sorting: The number of repeated cards. A larger number of repeated cards provides evidence for more types of simultaneous relations existing between concepts.

Repertory Grid: The number of dimensions elicited. A greater number of dimensions elicited provides a basis for more types of simultaneous relations between concepts.

MDS: The number of distinct dimensions labeled by the subject. Subjects are shown a plot of one MDS dimension at a time, and are asked to label the dimensions on which the concepts are arranged. The more distinct number of dimensions labeled, the more types of simultaneous relations exist between concepts.

2.5.4 Varying Degree of Relatedness

Card Sorting: The co-occurrence of a randomly selected concept pair in a pile. Two concepts occurring together in the same pile are more strongly related than two in separate piles.

HCS: The co-occurrence of a randomly selected concept pair in a cluster. Two concepts occurring together in the same cluster are more strongly related than two in separate clusters.

Repertory Grid: The distance between a randomly selected concept pair. All dimensions are collapsed, and the distance is taken between the average rating for each selected concept. The more related two concepts are, the smaller the distance between them.

MDS: The distance between a randomly selected concept pair. The euclidean distance is calculated between the concepts in the two-dimensional solution. The closer the two concepts are in space, the more related they are.

Pathfinder: The number of links between a randomly selected concept pair. The number of links is determined by the shortest possible path between the two concepts. A smaller number of links between concepts denotes a higher relation between those concepts

Pairwise Similarity Ratings: The similarity of a randomly selected concept pair. Therefore, a higher similarity rating denotes more similarity between given concepts.

2.6 Procedure

Card Sorting: Each subject was presented with a pile of 30 index cards. Each card contained one of the domain concepts elicited during the pilot study interviews. The subjects were asked to sort the cards into piles based upon which concepts they felt "go together". They were told that, when they finished, there may be as few as one pile (if all concepts were seen as going together) or as many as 30 piles (if all concepts were seen as being separate). Additional cards could be requested if the subject determined that a particular concept belonged in more than one pile. It was stressed that there were no right or wrong answers and there was no time constraint. When the sorting was complete, the subjects were asked to provide a label for each pile (that is, to provide a reasoning for their placement of the concepts). The total number of piles and their respective labels were recorded.

Repertory Grid: Following the card sorting procedure, subjects were presented with three randomly selected concept cards. They were asked to determine a dimension that would separate two of the concepts from the other one. Random triads were presented until the experimenter determined that a representative list of dimensions had been elicited. To construct the rating grid, the domain concepts were listed down the left-hand side of the grid while the elicited dimensions were listed across the top. Because of the dependency of the dimensions on an individual subject, each grid was tailored specifically for that subject. Subjects were presented with the grid the following day, and were allowed to complete it at home. They were asked to rate each concept on an ordinal scale from one to seven on each dimension, and to return the rating sheet as soon as they had finished.

Concept Elicitation: On a second day, subjects were asked to list as many concepts as possible which they felt to be important in the secretarial field of work. They were told that these concepts could be anything from equipment used to procedures followed throughout the day (identical to the pilot study interviews mentioned earlier). The percentage of procedural and declarative concepts were determined for each subject (i.e., "files" is considered to be a declarative concept, whereas "filing" is considered procedural).

Pairwise Similarity Ratings: A computer program was developed which presented all possible pairs of the domain concepts to the subjects (resulting in 435 concept pairs). The subjects were shown one pair at a time, and were asked to rate their perceived similarity of the two concepts on a Likert-type scale from one to seven (one meaning extremely dissimilar and seven meaning extremely similar). The subjects were asked to use the entire range of values as necessary. The subjects were also told that a number of factors may influence the similarity of the concepts, and that all factors should be considered when a rating was assigned. This data was later submitted to Proc Cluster, ALSCAL and Pathfinder for further analysis. The default settings were used for each of these techniques.

3.0 RESULTS

3.1 Hypothesis One

As stated earlier, the purpose of this hypothesis is to show that the proposed dimensions within the model of knowledge structure do indeed exist. To do this, it was necessary to determine if the values collected within each dimension for at least one measurement technique were either different from zero (as in the case of Static and Procedural Concepts and Multiple Relations) or different between concepts. The specific processes involved for each dimension are discussed below.

Static and Procedural Concepts

To test for the existence of this dimension, a t-test was performed on the average percentage of procedural concepts elicited from the Experienced group in order to determine that this percentage was greater than zero. The results of this test were significant, $t(14) = 12.07$; $p < 0.005$.

Multiple Levels of Abstraction

The starness values of two randomly selected concepts (Communication and Liaison) were compared to determine if differences in levels of abstraction exist within the Experienced group. The t-test was significant, $t(28) = 3.90$; $p < 0.01$, indicating a difference in the level of abstraction between the two concepts.

Multiple Relations

The number of repeated cards in the card sorting procedure (Experienced group) was tested to determine if this number was greater than zero. The t-test was significant, $t(14) = 2.43$; $p < 0.025$, providing evidence for the simultaneous existence of multiple relations between concepts.

Varying Degree of Relatedness

To determine the existence of this dimension, three concepts were selected by an independent observer: Accounting, Payroll and Shorthand. A t-test was then performed on the similarity ratings of the two concept pairs Accounting-Shorthand and Accounting-Payroll for the Experienced group. The test was significant, $t(28) = 2.82$; $p < 0.01$, indicating that varying degrees of relatedness exist between concepts within the knowledge structure.

3.2 Hypothesis Two

Static and Procedural Concepts

This variable was not significantly different between experience level groups, $t(23) = 0.90$; $p < 0.38$ (see table 3).

Multiple Levels of Abstraction

Card Sorting: The test between experience group means did not yield significant results for this variable, $t(28) = 0.57$; $p < 0.57$.

HCS: The results of this test were statistically significant, $t(28) = 2.53$; $p < 0.02$. The clusters of the Experienced group contained more concepts than the clusters of the Naive group.

Pathfinder: While not statistically significant, $t(28) = 1.91$; $p < 0.07$, there appears to be a trend for the Experienced group's representations to contain more stars than the Naive group's representations.

Multiple Relations

Card Sorting: This test did not yield statistically significant results, $t(28) = 0.51$; $p < 0.62$.

Repertory Grid: The number of dimensions elicited were not significantly different for the two experience level groups, $t(28) = 0.74$; $p < 0.46$.

MDS: This test did not yield statistically significant results, $t(27) = 0.47$; $p < 0.64$.

Varying Degree of Relatedness

The measurement of this dependent variable required that a pair of concepts be randomly selected by an outside party. The restriction placed on this selection was that the pair must have come from the six domain concepts that were only in common to the experienced secretaries in the pilot study. This would allow an observation of the differences in the degree of relatedness between the two experience groups for a presumably higher level pair of concepts. The pair chosen was "Accounting" and "Supervising Employees".

Card Sorting: The test between group proportions did not yield statistically significant results, $z = 0.46$;

$p < 0.32$.

HCS: Statistically significant results were found with this technique. A test between group proportions yielded $z = 2.46$; $p < 0.007$.

Repertory Grid: The result of this test was not statistically significant, $t(28) = 1.88$; $p < 0.08$.

MDS: This technique did not yield statistically significant results, $t(28) = 0.56$; $p < 0.58$.

Pathfinder: The result of this test was not statistically significant, $t(28) = 0.49$; $p < 0.63$.

Pairwise Similarity Ratings: The result of this test was not statistically significant, $t(28) = 1.7$; $p < 0.10$.

3.3 Hypothesis Three

Based on the results of the individual techniques used within Hypothesis Two, it is readily seen that differences do exist in the KR measurement capabilities between techniques on the model dimensions. In particular, HCS was the only technique that extracted significant differences ($P < .05$) in the knowledge structures of the two experience level groups.

4.0 DISCUSSION

4.1 Existence of Model Dimensions

The first hypothesis, that each proposed dimension of knowledge structure exists, was supported. This was shown by testing that the values obtained during experimentation for each of the dimensions were either different from zero (in the case of Procedural and Declarative Concepts and Multiple Relations) or different between concept pairs (in Multiple Levels of Abstraction and Varying Degree of Relatedness). Therefore, it has been shown that the proposed dimensions are a good *starting point* in the development of a comprehensive model of knowledge structure. Further, the support of this hypothesis provided the necessary foundation on which to perform the testing of Hypotheses Two and Three.

4.2 Affect of Experience Level on Model Dimensions

In regard to the second hypothesis, that the characteristics of the proposed knowledge structure model dimensions are affected by domain experience, three outcomes were possible for each of the proposed model dimensions: (1) None of the techniques administered show differences between experience levels, (2) a subset of the techniques administered reveal differences between experience levels, and (3) each of the techniques administered find differences between experience levels. As long as at least one of the measurement techniques administered revealed significant differences between experience level groups for a given model dimension, there is evidence to suggest that the characteristic of the particular dimension is affected by differences in experience level. As such, it could be suggested that each dimension in the model differs with experience level and may influence an individual's performance on a cognitive task. Such evidence was found for the dimensions of Multiple Levels of Abstraction and Varying Degree of Relatedness. However, the dimensions that did not yield significant differences between subject groups should not be discarded as possible determinants of performance. Perhaps differences would have been found had the techniques used been more sensitive to those dimensions of knowledge structure. Furthermore, one cannot rule out the possibility that the domain chosen for this study was not one in which experience level necessarily produces differences on all knowledge structure dimensions.

In addition to the proposed model dimensions, a post-hoc analysis indicated that another dimension measured with Pathfinder, Representation Complexity (the total number of links present in the representation), is an indicator of representation differences between experience levels, $t(28) = 2.43$; $p < 0.03$. Therefore, any future model of knowledge structure should include a complexity parameter.

4.3 Differences in Measurement Techniques

The third hypothesis, that differences exist in knowledge structure measurement capability among techniques on the model dimensions, was supported. This was shown by the different results obtained by the various techniques across the model dimensions. The results of this study indicate that hierarchical clustering analysis was the only technique to extract significant differences ($p < .05$) between the experience groups on any of the model dimensions. It appears then, that the remaining techniques were not sensitive enough in this scenario to yield significant differences between experience groups. It is interesting to note that the one technique that extracted representation differences between skill levels used statistical methods to transform the data prior to producing the representation. This provides support for the latent trait techniques as discussed in Calvez, et al. and suggests that, indeed, these transformation techniques are useful in eliciting the contents of knowledge structure. Although they did not yield significant results at the 0.05 alpha level, in comparison to the other techniques, Pathfinder (Multiple Levels of Abstraction) and the repertory grid technique (Varying Degree of Relatedness) appear to have revealed differences between experience levels and should, therefore, be considered as useful techniques for measuring these respective dimensions. Given these results, it appears that Pathfinder and hierarchical clustering analysis are useful techniques in revealing differences in the levels of abstraction between different experience groups, while the repertory grid technique provides an effective indication of the varying degree of relatedness between pairs of domain-related concepts. As mentioned earlier, none of the measurement techniques employed in this study were able to detect differences between experience levels for the dimension of Multiple Relations. Although differences were not found for this dimension, the author believes this may be due to a weakness of the available measurement techniques. More research is needed to determine the useability of this dimension to determine knowledge structure differences.

4.4 Research Agenda

Three specific areas of future work are discussed below: Refinement of the model of knowledge structure, the development of new measurement methodologies, and the application of this research in the work place.

From the results of this study, it may be seen that there is a need to refine the proposed model of human knowledge structure. In particular, the inclusion of an additional dimension that represents the complexity of the knowledge structure is necessary. Furthermore, it is imperative that work continue in the development of knowledge structure measurement techniques. For, none of the techniques used here have the ability to measure each of the parameters of the proposed model of knowledge structure.

The focus of future work should be on the development of a comprehensive measurement methodology capable of eliciting reliable information from each of the parameters of knowledge structure. The results of this experiment necessitate that any new technique be able to determine differences in the dimensions of Multiple Levels of Abstraction, Varying Degree of Relatedness, and Representational Complexity. In addition, the technique should elicit quantitative data to reduce subjectivity in the interpretation of the knowledge structure dimensions, and should be applicable for use in a variety of work place domains.

When these goals have been accomplished, it may be possible to predict an individual's performance on a given cognitive-oriented task based upon the characteristics of that individual's knowledge structure. Such information would be applicable in personnel selection, training and job design. For example, certain jobs may be better suited for individuals with a higher level of abstraction in their knowledge structure. Training could be geared toward the development of certain knowledge structure attributes needed for an operator to efficiently perform a given job. Finally, jobs themselves could be designed in such a way as to be "matched" with the knowledge structure of the operator, for there has not been much success with training individuals to adopt a particular knowledge structure (Adelson, 1984).

In summary, this experiment has validated the existence of four proposed dimensions of knowledge structure, identified two as being affected by differences in domain experience levels and identified an additional dimension to be included in a future model of knowledge structure. Furthermore, differences were found in the measurement capabilities of available knowledge structure measurement techniques. Therefore, the primary objectives of this effort were satisfied. These results support the importance of validating a formally defined model of knowledge structure, and the need for the development of new measurement methodologies in order to advance our understanding of the field.

REFERENCES

ADELSON, B., 1981, Problem solving and the development of abstract categories in programming languages. *Memory & Cognition*, 9, 422-433.

ADELSON, B., 1984, When novices surpass experts: The difficulty of a task may increase with expertise. Journal of Experimental Psychology, 10,(3), 483-495.

ANDERSON, J.R., 1982, Acquisition of cognitive skill. *Psychological Review*, 89,(4), 369-406.

CALVEZ, V., KOUBEK, R.J. AND MOUNTJOY, D.M., 1991, A comparative review of knowledge structure measurement techniques for interface design. Working Paper, School of Industrial Engineering, Purdue University, West Lafayette, IN 47907.

CHI, M.T.H., FELTOVICH, P.J., AND GLASER, R., 1981, Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.

GLASER, R., 1976, Components of a psychology of instruction: Toward a science of design. Review of Educational Research, 46,(1), 1-24.

GLASER, R. AND BASSOK, M., 1989, Learning theory and study of instruction. Annual Review of Psychology, 40, 631-666.

GOLDSMITH, T.E., AND DAVENPORT, D.M., 1990, Assessing structural similarity of graphs. In R.W. Schvaneveldt (ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations* (New Jersey: Ablex Publishing Corporation), 75-88.

KOUBEK, R.J., AND MOUNTJOY, D.N., 1991, The impact of knowledge representation on cognitive-oriented task performance. *International Journal of Human-Computer Interaction*, 3, 31-48.

KOUBEK, R. J., AND SALVENDY, G., In press, Cognitive performance of super-experts on computer program modification tasks. *Ergonomics*.

PALMER, E.S., 1978, Fundamental aspects of cognitive representation. In E. Rosch and B. B. Lloyd (eds.), *Cognition and Categorization* (Hillsdale, NJ: Earlbaum), 259-303.

SAS Institute (1988). *SAS/STAT User's Guide*. (Cary, NC: SAS Institute Inc.).

SHIFFRIN, R.M. AD DUMAIS, S.T., 1981, The development of automatism. In J.R. Anderson (ed.), *Cognitive Skills and their Acquisition* (Hillsdale, New Jersey: Erlbaum), 111-140.

SMITH, B.C., 1982, Reflection and semantics in a procedural language. Ph.D. thesis and Tech. Report MIT/LCS/TR-272, M.I.T., Cambridge, MA.

Table 1. Proposed model of knowledge structure.

COMPONENT	FEATURE
Element	Unique concepts, or knowledge units, of a domain exist as elements in the KR.
	Elements can exist as declarative or procedural concepts.
	Multiple levels of abstraction among elements can exist in a single KR.
Interrelationship	Relationships exist between the elements.
	The relationship between elements can vary in degree.
	More than one type of relationship can simultaneously exist between two or more elements.

Table 2. Dependent variables for Hypothesis Two.

KR MODEL DIMENSION	TECHNIQUE					
	Card Sorting	HCS	Repertory Grid	MDS	Pathfinder	Similarity Ratings
Declarative & Procedural Concepts	Percentage of Procedural Concepts Elicited					
Multiple Levels of Abstraction	Average Number of Cards Per Pile	Average Number of Concepts Per Cluster	NA	NA	Number of Stars	NA
Multiple Relations	Number of Repeated Concepts	NA	Number of Dimensions Elicited	Number of Dimensions Labeled	NA	NA
Varying Degree of Relatedness	Co-Occurrence of Pairs in Piles	Co-Occurrence of Pairs in Clusters	Distance Between Concept Pairs	Distance Between Concept Pairs (2-D)	Number of Links Between Concept Pairs	Similarity of Concept Pairs

NA = No quantitative measure available

Table 3. Summary of results for Hypotheses Two and Three.

KR MODEL DIMENSION	TECHNIQUE						SUMMARY
	Card Sorting	HCS	Repertory Grid	MDS	Pathfinder	Similarity Ratings	Significant Differences
Declarative & Procedural Concepts			$X_E = 0.82$ $SD_E = 0.13$ $X_N = 0.74$ $SD_N = 0.24$ $p < 0.39$				NO
Multiple Levels of Abstraction	$X_E = 6.28$ $SD_E = 2.54$ $X_N = 5.77$ $SD_N = 2.34$ $p < 0.57$	$X_E = 3.19$ $SD_E = 0.64$ $X_N = 2.71$ $SD_N = 0.39$ **$p < 0.02$**	NA	NA	$X_E = 13.60$ $SD_E = 5.64$ $X_N = 10.33$ $SD_N = 3.46$ $p < 0.07$	NA	YES
Multiple Relations	$X_E = 1.93$ $SD_E = 3.08$ $X_N = 1.47$ $SD_N = 1.77$ $p < 0.62$	NA	$X_E = 9.47$ $SD_E = 2.17$ $X_N = 8.87$ $SD_N = 2.26$ $p < 0.46$	$X_E = 3.2$ $SD_E = 0.94$ $X_N = 3.36$ $SD_N = 0.84$ $p < 0.64$	NA	NA	NO
Varying Degree of Relatedness	$p_E = 0.13$ $p_N = 0.07$ $p < 0.32$	$p_E = 0.33$ $p_N = 0.00$ **$p < 0.007$**	$X_E = 1.01$ $SD_E = 0.54$ $X_N = 0.67$ $SD_N = 0.44$ $p < 0.07$	$X_E = 2.19$ $SD_E = 0.51$ $X_N = 2.06$ $SD_N = 0.75$ $p < 0.58$	$X_E = 3.00$ $SD_E = 1.60$ $X_N = 3.27$ $SD_N = 1.39$ $p < 0.63$	$X_E = 1.73$ $SD_E = 1.28$ $X_N = 2.60$ $SD_N = 1.50$ $p < 0.10$	YES
Significant Differences	NO	YES	NO	NO	NO	NO	

NA = No quantitative measure available

ATTACHMENT 3

A Comparative Review of Knowledge Structure Measurement Techniques for Interface Design

Darel V. Benysh, Richard J. Koubek, and Vance Calvez

In : International Journal of Human-Computer Interaction 5(3) 211-237 (1993)

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1256 and # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

A Comparative Review of Knowledge Structure Measurement Techniques for Interface Design

Darel V. Benysh[1], Richard J. Koubek[1], and Vance Calvez[2]

[1]School of Industrial Engineering
Purdue University
West Lafayette, IN 47907

[2]Department of Biomedical and Human Factors Engineering
Wright State University
Dayton, OH 45435

Running Title: Measuring Knowledge Structures

Please address all correspondence to the Richard Koubek, Purdue University. This work was supported by the Cognitive and Neural Sciences Division, Office of Naval Research, under grants # N00014-90-J-1256 and # N00014-92-J-1153. The opinions expressed in this paper do not necessarily reflect the position of ONR. The authors would like to acknowledge Dan Mountjoy and Shaun Daugherty for their help in preparation of this document.

ABSTRACT

Recent research suggests that the structure of user knowledge, in addition to the content, is a significant determinant of user behavior in computer-oriented tasks. The purpose of this paper is to provide the researcher in human-computer interaction with a comparative review of available knowledge structure measurement techniques and to summarize their potential application for aiding the designer at various stages of interface design. Three general classes of techniques are identified: verbal reports, clustering and scaling. Each group is reviewed according to (1) the manner in which concepts and relationships are elicited, (2) the method used to derive the knowledge structure and (3) procedures used to analyze the knowledge structure. With this information, the human-computer interface designer can more effectively use these techniques for their particular application.

1.0 INTRODUCTION

1.1 Overview of User Centered Design

The expansion of information technology in the workplace has provided engineers and psychologists with new possibilities for designing jobs and tasks to best match operator characteristics. Given the power and flexibility of a computational environment, there are new opportunities for following the often-stated principle of matching the task and user. Such user-centered design overcomes the costs of requiring the worker to conform to a pre-determined work environment. For example, if a job is fixed such that only a narrow class of personnel is qualified, labor costs will increase as the qualified applicant pool decreases. Interfaces created through user centered design principles can consist of either an intelligent system which is capable of evaluating and adapting to the user, or a system which is designed to conform to the characteristics of the expected user population. In either case, the capability to identify the user characteristics which impact performance becomes critical to the design process.

1.2 User Characteristics

Work in the area of identifying user characteristics has traditionally focused on measuring general cognitive abilities, as found in the personnel selection literature, or identifying task specific knowledge possessed by the operator. However, recent research suggests that the structure of knowledge, in addition to the content or general global abilities, is a significant determinant of user behavior in computer-oriented tasks.

For example, Koubek and Mountjoy (1991) demonstrated that the manner in which an individual structured knowledge of word processing commands affected the performance of a complex text-editing task. In particular, those individuals with an abstract knowledge structure were more likely to complete the task than those with a concrete structure. Therefore, it appears that in order to maximize performance on any given cognitive task, it would be beneficial to "match" an individual possessing a particular knowledge structure with a task or interface most suited for that structure type.

Knowledge structure differs from knowledge content in that the structure is basically the way in which a person's body of knowledge is organized. They are also different than mental models in that knowledge structures represent static relationships between concepts rather than an operational model of systems. The structure of the organization inherently affects the way the person utilizes their knowledge. An inappropriate structure can make even the most complete body of knowledge ineffective.

Work in understanding this phenomenon of knowledge structure may be traced back to the research of Tulving (1962) who determined that, given a list recall task, humans tend to order that list according to their structure of the relationships between concepts. Early work examining individual differences in the performance of master and novice chess players (de Groot, 1965; Chase and Simon, 1973) found that master chess players were able to interpret patterns or "clusters" in the meaningful arrangements of chess pieces which allowed them to streamline their memory processes.

These initial experiments spawned a flurry of research which attempted to discover the properties of knowledge structures which impact cognitive task performance. This work has encompassed a variety of domains such as physics and mathematics problem solving, computer programming, electronics, medical diagnosis, software design, and psychological diagnosis (e.g. Egan and Schwartz, 1979; Adelson, 1981; Shoenfeld and

Herrmann, 1982; Weiser and Shertz, 1983; Murphy and Wright, 1984; Adelson, 1984; Barfield, Koubek, and Hwang, 1985; Barfield, 1986; Hobus, Schmidt, Boshuizen, and Patel, 1987; Hardiman, Dufresne, and Mestre, 1989;). Researchers attempting to obtain information about the characteristics of knowledge structures have developed a diverse group of techniques which, to some extent, present an individual's knowledge structure. These techniques vary greatly in the information obtained, method of elicitation, and the format with which the structure is presented.

1.3 The Design Process

When designing human computer interaction, different stages of the design process require different types of information regarding the potential user base. Many attempts have been made to classify the stages of software development. Some researchers have proposed detailed processes while others stress a more situationally dependent general rule–based method (Booth, 1989; Henderson, 1991; Eason and Harker, 1991; Gardner, 1991). It is not the intent of this paper to introduce more speculation this topic. Instead, we will adopt a general framework of interface software development involving four stages in which to discuss the various contributions of knowledge structure assessment techniques.

The first stage of the software development process addresses evaluation of the existing system. This opens the possibility for a variety of interpretations. For instance, this includes the evaluation of an existing software package for which a successor or competing product is being designed. This also includes evaluation of the methods used in a non-computerized task situation which is being computerized. The premise at this stage is that operators familiar with the task will have developed certain procedures and relationships between concepts. Many times it is preferable to design within these predefined knowledge structures. Othertimes incongruities between how the operator conceptualizes and actually performs a task will indicate need for a new design approach which would bring the two closer together.

The second stage to be considered is the initial design stage. This includes such tasks as defining design objectives and specifing interface characteristics. Issues at this stage include interface consistency and information structure problems. The third stage, formative design, is the actual production of the interface. Besides the actual creation of the interface software, one of the key issues here is the comparison and selection between alternative interface designs. The final stages, summative evaluation and installation, will be lumped together and will not be included in our discussion of knowledge structure assessment technique contributions, since the contribution of knowledge assessment to the development process should be included in the early stages of development to be of any use.

Typically, at each of these stages, the designer uses an explicit model of the user or an implicit understanding of the users' needs based on the designer's understanding of these needs. Therefore, providing user characteristic information, such as knowledge structure, to the designer can benefit the entire design process.

1.4 Purpose

The purpose of this paper is to provide the researcher in human-computer interaction with a comparative review of available knowledge structure measurement techniques and to summarize their potential application for aiding the designer at various stages of interface design. With this information, the investigator can select the appropriate technique for their particular application. Three general classes of techniques are identified: verbal reports, clustering and scaling. Each group is reviewed according to (1) the manner in which concepts and relationships are elicited, (2) the method used to derive the knowledge structure and (3) procedures used to analyze the knowledge structure.

1.5 Fundamental Issues

All of the techniques discussed here are fundamentally introspective in nature. The techniques differ in terms of how systematic and formal that introspection is, but all techniques begin with individuals making some type of metacognitive judgements (i.e., judgements about what they know). However, many of the existing techniques then perform statistical or multivariate analyses of these judgements. These analyses may give the investigator the impression that the source of the information is somehow more objective, such as with a physiological or performance-based measure. As a result, it is important for the investigator to keep a proper perspective regarding the origins of the resulting representations.

While it should be acknowledged that the existing techniques elicit information which is interpretive, one should not ignore the utility of that information source. The questioning of the introspectionist paradigm around the beginning of the twentieth century was in part due to the fact that individuals were unable to reliably introspect about their cognitive processes. Because of this, the use of introspection as a data source has developed a negative connotation. However, metacognitive judgments can contain information which is difficult or impossible to obtain from more objective sources (Ericsson and Simon, 1984). Therefore, one may argue that investigations into knowledge structure measurement techniques should not attempt to avoid introspective aspects of the measurement process.

There are several common features which characterize the type of knowledge elicited by these techniques. First, the knowledge elicited is relatively stable over time (Chi, Hutchinson, and Robin, 1989). Second, the type of knowledge elicited tends to be somewhat more declarative than procedural in nature. While some of the techniques can elicit more procedural types of knowledge, such as in terms of "if-then" rules (Anderson, 1983), the majority of research has produced techniques which are more effective at eliciting declarative knowledge (Rips, Shoben, and Smith, 1973). Third, the existing knowledge measurement techniques provide representations of complex, cognitive domains. As a result of this, each technique may be more applicable to some domains than others. Further, it is important to note that underlying user structures provide only one of the many user characteristics which can be considered in HCI design.

In general, the measurement of knowledge structures, has three component parts. First, the important concepts and relationships in the domain of interest are elicited. For example, if an expert's knowledge structure of programming were being measured, a list of concepts such as "LOOPING", "RECURSION" and "ARRAY" would be developed. In addition, the relationships, or degree of relationship, between the various concepts would be determined. In the programming example, subjects may be requested to rate the degree of similarity between "LOOPING" and "ARRAY" using a 7-point scale. With the raw data collected, the second phase is the derivation of the underlying structure. In this process, the raw data is transformed to reveal a structure for later analysis. The third and final stage is analysis of the derived knowledge structure. It is in this stage that dependent variables can be derived which characterize an individual's knowledge structure. The validity and reliability of the final dependent variables are a function of all proceeding steps described above.

From review of the literature, three general classes of knowledge structure measurement techniques emerge: verbal reports, clustering methodologies and scaling methodologies. The output from these techniques range from rather qualitative (verbal reports) to highly quantitative (scaling methodologies). Techniques within each of these classes tend to have similar properties with regard to the three stages of measurement listed above. Therefore, the structure of this paper is as follows. Each of the general classes will be discussed in turn. The review of each measurement class will focus on four components: structure elicitation, structure derivation and structure analysis and application to one of the design stages. Tables 1, 2 and 3 provide a summary of the discussion to follow.

2.0 VERBAL REPORTS

Verbal reports are based on the intuitive notion that, if one wishes to obtain information concerning an individual's knowledge structure, then it is appropriate to have the individual verbally describe the characteristics of his or her domain knowledge. As will be discussed, there are several critical drawbacks associated with the use of verbal reports for knowledge structure measurement. Nevertheless, verbal reports are a popular means of eliciting cognitive information in a variety of domains. The common elements among the verbal report techniques are as follows. First, all data is generated from verbal descriptions of the domain of interest. Second, the construction of a representation of the knowledge structure is based on the investigator's interpretation of the data. Finally, the investigator often applies a particular cognitive model, such as GOMS, NGOMSL, Task Action Grammar, or Production Rules, to organize the data.

The following discussion of verbal reports is divided into two stages: information elicitation and representation development. At present, the analysis of structures in basically left to investigator interpretation. The information elicitation stage discusses verbal techniques which strictly involve the "information gathering" portion of the knowledge representation definition process. The representation development phase describes two popular models of cognitive task performance that investigators have used as a framework for the results of the verbal reports.

2.1. Information Elicitation

2.1.1 Interviewing. Interviewing has been the most widely used technique to elicit information from experts in the development of expert systems (Evans, 1988). Individuals are asked a series of questions concerned with the performance of the given problem or job, structured in such a way as to, first, elicit general concepts, and later, to determine relationships among those concepts (Meister, 1985; Olson and Rueter, 1987; McPherson and Thomas, 1989; Means and Voss, 1985; Leinhardt and Smith, 1985). Further, investigators use a funnel approach in which the line of questioning starts very broad and then becomes more specific as the interview proceeds. Interviewing is primarily a retrospective technique in that the individual being interviewed elicits information subsequent to performance of domain related tasks.

2.1.2 Questionnaires. Questionnaires may also be an effective manner in which to elicit variables and their rélationships important to performing a given task. The questionnaire is presented in a written format with open ended questions. Questions are posed in such a manner as to allow the individual to identify critical variables, and the relationship between these variables. Like interviewing, a questionnaire is a retrospective technique.

2.1.3 Interruption Analysis. Interruption analysis is performed by observing the subject while they perform the task. If at any point during task performance the subject's actions or thought processes become unclear to the observer, the latter interrupts the task in order to probe the individual for information regarding the reasoning behind those unclear actions or thought processes (e.g. "Why did you do that?" or "What did you gain from that?"). As noted by Olson and Rueter (1987), once the task has been interrupted, it is difficult to resume.

2.1.4 Protocol Analysis. Protocol analysis is a further attempt to eliminate inferences drawn by outside sources. In the generation of a protocol, an individual is asked to "think out loud" as he/she performs the task to be analyzed. The individual should identify any goals and methods currently being utilized to reach the task solution.

A number of researchers have used protocol analysis to elicit knowledge in a variety of domains. The areas of computer programming, software design, and word processing have been popular domains for the application of protocol analysis (Vessey, 1985; Soloway and Adelson, 1985; Koubek, Salvendy, Eberts and Dunsmore, 1987; Koubek and Salvendy, 1989; Zeitz and Spoehr, 1989; Koubek and Mountjoy, 1991). Its application toward the general domain of education and problem solving is also prevalent in the literature (Chi, Feltovich and Glaser, 1981; Sweller, Mawer, and Ward, 1983; Leinhardt and Smith, 1985; Sweller and Owen, 1985; Lunderberg, 1987; Ploger, 1988). The use of protocol analysis in these domains is reasonable because the thought processes may be verbalized quite naturally. In other domains, in which the learning process is less verbal (such as playing golf), protocol analysis is much less effective (Cooke and McDonald, 1986).

This technique, according to Ericsson and Simon (1984), is a valid means of obtaining an individual's momentary cognitive processes since the individual's immediately preceding thoughts are stored in short term memory, and can be accessed by the individual directly. However, some investigators are concerned that task performance is altered during the verbalization process and would therefore produce unnatural protocols (Ericsson and Simon, 1984; Musen, 1989; Nisbett and Wilson, 1977).

2.1.5 Discussion of Information Elicitation with Verbal Reports. In past years, verbal reports have generally been labeled as introspective in their generation of data (Nisbett and Wilson, 1977), and, therefore, have been discarded by many as being valuable tools of knowledge elicitation. However, Ericsson and Simon (1984) argue that some of these techniques may be unfairly categorized as highly introspective (e.g. protocol analysis). As is true with the application of any knowledge elicitation technique, strict guidelines must be followed in order for the output of verbal reports to be valid. Ericsson and Simon (1984) have identified three criteria that must be met to insure the proper use of verbal reports. The first of these is the "Relevance Criterion", which states that verbalization should be a normal part of the performance of the given task. The "Consistency Criterion" explains that consecutive verbalizations should be logically consistent. Finally, the "Memory Criterion" requires that the person remember a subset of the information that was attended to during task performance. Ericsson and Simon (1980, 1984) contend that if these guidelines are followed, verbal reports can produce useful reliable data for the study of cognitive processes.

2.2 Representation Generation

Once the verbal data has been collected, it must be transformed into an analyzable format. This process is neither easy nor well defined for most of the techniques previously described. Three popular techniques which do have guidelines to follow in the structuring of verbal data are GOMS models, grammars and production rule systems. For a complete description of these techniques, see Olson and Olson (1990).

In the representation development techniques for verbal report data, much subjectivity is left to the analyst. Even if an attempt is made to quantify certain aspects of the representations, the ultimate decision as to which aspects are to be quantified is made by the analyst. This may tend to bias the results toward revealing the kind of information that the analyst wishes to obtain. However, if care is taken in the selection of the quantifiable aspects, as well as in the experimental design, interesting relationships between individual characteristics may be found. The primary strength of verbal reports is the flexibility they give to an investigator.

The analyst may decide to use a cognitive model, such as GOMS, NGOMSL, Task Action Grammar, or Production Rules, to organize the resulting data. GOMS modeling (Card, Moran, and Newell, 1983) attempts to capture the goals of the user, how these goals are decomposed into sub-goals, and how observable behaviors are used to satisfy these goals. GOMS organizes these sub-goals and their associated behaviors as elements of a goal stack which represents the steps required to complete a goal. This model can be used to make estimates of relative execution time, learning times, and transferability. The NGOMSL model (Kieras, 1988), is similar to the GOMS model in the way it decomposes a task. However, it uses a much more formalized language and organizes the elements into a computer programming type subroutine structure. Additionally, NGOMSL modeling enables the analyst to make much more accurate predictions of execution time, learning time, consistency gains, and working memory load. Grammars, specifically Task Action Grammar (TAG) (Payne and Green, 1986), and production rules (Newell and Simon, 1972) attempt to organize the Human-Computer interaction language into a set of rules for specifying commands. Typically, grammars will be represented as a set generalized syntactic forms for specifying a whole host of possible commands. Production rules are presented as a series of IF-THEN statements which characterize an even broader range of interaction possibilities. Once in either of these formats, they may be used to make measurements of interface consistency, relative learnability, and generation and generation and testing of research hypotheses.

2.3 Application of Verbal Reports to HCI Design

The first design stage, system evaluation, is primarily concerned with gathering information about the tasks, procedures, and the user organization of task specific knowledge. Research has shown that verbal reporting techniques, which consist of concept elicitation and investigator's interpretive representation, are useful tools for extracting higher level and global task element information (Koubek and Salvendy, 1986). However, these techniques are less effective at eliciting task specific information. These techniques will aid the interface designer in identifying tasks and other design issues to be considered throughout the entire design cycle. Indeed, user experiences, revealed through interviews and protocol analysis, can uncover or emphasize design issues previously unsuspected or ignored by the designer.

3.0 CLUSTERING METHODOLOGIES

The clustering methodologies can viewed as a midpoint between the scaling methodologies and the verbal reports that were previously discussed. The clustering techniques are more formal and systematic than the verbal reports and less so than the scaling methodologies. The important difference between the clustering methodologies and verbal reports is that with verbal reports, the analyst produces the representation of the elicited knowledge based on their subjective interpretation of the data. With the clustering methodology, the representation is generated directly from the user inputs (e.g. by having them perform a series of tasks) or is analytically derived from the results of a task. Thus, the investigator has substantially less influence on the characteristics of the representation obtained.

The foundations of the clustering methodologies began with the early work in expert-novice differences (de Groot, 1965; Chase and Simon, 1973; Reitman, 1976; Egan and Schwartz, 1979). This research suggested that domain-relevant concepts are grouped together in clusters by those individuals proficient in the domain of interest. Further, these clusters were viewed as the building blocks from which more elaborate knowledge representations might be constructed. De Groot (1965) and his colleagues examined the performance of master and novice chess players. The major finding of their work was that master chess players showed a superior ability to recall chess board configurations only when the boards were arranged in actual game situations as opposed to random arrangements of the pieces. This result suggested that expert chess players perceived patterns (e.g. offensive and defensive groups) in the chess piece configurations. Indeed, the existence of a superior knowledge structure or problem structure appeared to increase the ability of a master chess player to efficiently recall large amounts of relevant information. Chase and Simon (1973) used a recall technique to establish boundaries of these cognitive clusters in chess board configurations. These techniques have also been used to elicit information in the field of computer programming (Barfield, Koubek, and Hwang, 1985).

The following techniques may counter some of the problems inherent in verbal report techniques, providing an alternative to these less formal and objective techniques. Clustering methodologies are divided into three main stages: concept elicitation, cluster elicitation, and the analysis of representations. This division of concept list elicitation from the elicitation of the clustering relationships is one of characteristics that distinguishes clustering methodologies from verbal reports.

3.1 Concept Elicitation

The elicitation of a list of domain relevant concepts is a significant, yet often underemphasized, aspect in the generation of a representation of human knowledge structures. To date, few formal techniques have been created to systematically extract the important domain-related concepts (Cooke and McDonald, 1987).

The techniques given below are, for the most part, borrowed from the verbal report techniques such as interviewing and protocol analysis. However, concept elicitation techniques differ from the verbal reports in two important ways. First, the format of the resulting information in concept elicitation is predetermined. Indeed, the output of these techniques is required to be a list of domain relevant concepts. This is in contrast to verbal reports, where the characteristics of the output are to a large degree uncertain. Also, verbal reports attempt to obtain all important knowledge concerning the domain simultaneously, whereas concept elicitation only generates domain relevant concepts.

3.1.1 Investigator Judgement. In the majority of studies examined in the literature, investigators obtained the concept list from researcher's intuition or written technical information on the subject rather than the individuals under study. This approach to concept elicitation would certainly seem to be counter-productive because the information does not originate from the individuals being measured. Nevertheless, an investigator may be compelled to use this technique because it may reduce costs substantially. Further, in domains such as computer programming or problem solving, a list of concepts may be self-evident.

Cooke and Schvaneveldt (1988) selected programming concepts from the chapter headings of an introductory text on the subject. McKeithen, Reitman, Rueter, and Hirtle (1981) used single syllable ALGOL reserved words as the concept list. Chi, Feltovich, Glaser (1981) selected physics problems from a physics reference for use in their experiment. Adelson (1981, 1984) subjectively selected computer programs and flowcharts to present to novice and expert computer programmers. Hollands and Merikle (1987) selected psychological terms from several psychology texts.

3.1.2 List Generation. This technique involves the manual listing of a set of domain relevant concepts by an individual or group. Cooke and McDonald (1986) describe three different ways in which domain critical concepts may be elicited in list form: critical concept listing, step listing and an outline generation method. In critical concept listing, the individual is simply asked to list as many domain relevant concepts as possible. In step listing, the individual is prompted to list concepts related to a particular task that they have observed or performed. In the outline generation method, the individual is asked to list the headings and subheadings of a hypothetical book on the domain of interest. Rips, Shoben, and Smith (1973) used a technique similar to critical concept listing to elicit familiar bird and mammal terms. The 12 terms most frequently mentioned by students in a five minute period were selected.

3.1.3 Interviewing and Protocol Analysis. Cooke and McDonald (1986) also discussed and investigated an interviewing method. In their method, an individual observes an interviewing scenario in which one person asks another person open ended questions about the pertinent domain. Similarly, in protocol analysis, an individual would observe a recorded protocol of a domain relevant task and would be asked to record all significant words mentioned. These techniques, of course, are similar to the corresponding techniques in the verbal reports section. Schvaneveldt, Goldsmith, Durso, Maxwell, Acosta and Tucker (1982) used an iterative series of literature searches, task analyses, and interviews with fighter pilots to obtain a concept list for two tactical flight maneuvers.

3.1.4 Discussion of Concept Elicitation Techniques. The trend in the literature is for the list of domain relevant concepts to be much smaller than the actual size of the concept list that an expert possesses. To some extent, this occurs because of practical limitations on budgets and time, but also occurs because researchers often believe that a partial list of concepts is sufficient for the applications of the analysis.

Studies have found that different methods of concept elicitation tend to obtain certain types of information. Cooke and McDonald (1986) compared several concept elicitation techniques (concept lists, interview, task-based lists, chapter lists). For the domain of "driving a car", they found that concept listing and task-based listing generated mostly general rules ("Wear seat belts", p. 1428) while interviewing and chapter listing revealed mostly concepts ("brakes", p. 1428). They argued that different techniques are differentially suitable to obtain various types of information. To produce an adequate concept list, investigators must make a decision regarding the desired depth of analysis and closely examine the domain of interest to produce an operational definition of a concept which can guide the decision of an appropriate technique.

3.2 Cluster Elicitation.

A cluster elicitation technique takes information in terms of a list of domain related concepts, establishes discrete or continuous relationships between the concepts, and outputs a spatial presentation of these relationships.

3.2.1 Card Sorting. The basic idea behind the card sorting technique is somewhat self explanatory. Individuals manually sort a set of concepts into several piles. Each concept in the concept list is labeled on a card. The analyst presents all of the concepts to the person. Typically, the individual is instructed to divide the cards into groups based on which concepts "go together" (Gobbo and Chi, 1986, p. 224). Most frequently, there is no restriction placed on the number of groups which can be created. Further, duplicate concept cards are often encouraged (McDonald, Paap, and McDonald, 1990) so that a concept may appear in more than one group. This increases the variety of types of relationships which can be elicited. Often in card sorts, individuals are encouraged not to sort concepts with which they are not familiar. In addition, the labels of the piles are often elicited. Weiser and Shertz (1983) had individuals sort programming problems and compared the labels individuals given to the clusters across groups.

3.2.2 Ordered Trees. The basis of this technique states that the way in which concepts are stored in an individual's long-term memory is reflected by the order in which they are recalled. Specifically, it is proposed that humans store and recall all items in a particular cluster before recalling items in another cluster (Olson and Rueter, 1987). Given the assumption that humans store information in chunks or "clusters" of concepts, attempts have been made to examine the recall process to identify the boundaries of these clusters.

Chase and Simon (1973) had chess players reconstruct chess board configurations from memory to determine clusters of chess pieces. They argued that the response latencies between clusters of chess pieces should be longer than latencies within a cluster. Tulving (1962), McLean and Gregg (1967) and Bower and Springsteen (1970) also used inter-response latency information to derive cluster-type relationships for a set of concepts. Reitman (1976) attempted to replicate Chase and Simon's study using another board game, Go. She found that the error variation in inter-response latencies was so large that she was unable to make useful conclusions about the boundaries of the clusters. Further, researchers had, at that time, begun to believe that clusters were often arranged hierarchically (Chase and Simon, 1973). This technique was not well suited for identifying nested clusters.

Reitman's dissatisfaction with the use of inter-response latencies led to a new technique which focused on the sequence of concepts in the recall as opposed to temporal aspects (Reitman and Rueter, 1980). In this technique, individuals were asked to repeatedly recall the concept list. Prior to recall trials, the individual completely memorized the concept list. On each trial, a different one of the concepts was presented as a cue word to initiate the recall. This encouraged the individual to base recall on their internal organization of the concepts rather than on the memorization of one sequence. An algorithm, developed by Reitman and Rueter, identified strings of concepts which were always adjacent to each other, regardless of order, and clustered them together. This technique is capable of identifying clusters of concepts but it forces a hierarchical format on the data. The result of the algorithm is a spatial, hierarchical representation of the clusters. McKeithen et al. (1981) used this technique to examine the knowledge structures of computer programmers. A distinctive aspect of this procedure is that it can identify patterns of sequences in a particular cluster. This technique is often called ordered trees due to the sequence patterns developed in the hierarchical clusters.

3.2.3 Closed Curve Analysis. The method of closed curves allows one to discover the relationships between objects which are represented spatially (Olson and Rueter, 1987). Traditionally, closed curves have been used to determine differences in the chunking ability of novices and experts in various spatial domains such as board games and circuit analysis (Reitman, 1976; Egan and Schwartz, 1979). The application of the closed curves technique is straightforward. Subjects are instructed to circle everything that "goes together". Thus, this task may be intuitively easy for individuals to perform. The subject matter for a closed curve technique may be any graphical representation of domain information, such as a plant layout or interface display panel.

In the case of Reitman (1976), the master Go player was shown a game board configuration which had been reproduced from memory at an earlier time. The master was then asked to circle the partitioning of elements which they saw in that configuration. Reitman discovered that the master Go player tended to chunk game patterns in overlapping clusters, not only as separate chunks or nested hierarchies. In addition, she found that the master Go player produced very reliable closed curve representations. In another example, Egan and Schwartz (1979) asked a skilled electrician to circle functional units of a circuit diagram and provide a verbal label for each functional group. Skilled electricians grouped schematics in functional units (e.g. amplifiers, feedback networks, filters and rectifiers), whereas novice electricians grouped identical diagrams in haphazard manners.

This method is unique in that it allows the individual to indicate chunks directly. In addition, closed curves is the only cluster elicitation technique (with the possible exception of card sorting using duplicate cards) which produces overlapping clusters representations as well as nested cluster representations.

3.2.4 Spatial Reconstruction. In this technique, individuals are asked to reconstruct an existing spatial system, such as a board game or circuit board. As the individual performs the reconstruction, the investigator collects information regarding the elements that are placed in the work space between glances back to the reference system. Elements placed between two glances are claimed to be chunked together. Spatial reconstruction produces a single level clustering representation.

Chase and Simon (1973) asked chess players of varying levels of experience to reconstruct mid- and end-game chess configurations on an adjacent chess board with the target board in plain view. The investigators recorded the order of reconstruction and the chess pieces placed between glances to the target board. They also recorded the between and within glance latencies for placement of each chess piece. Chase and Simon argued that the master chess player structures the chess configuration into patterns or "chunks" (p. 57). They hypothesized that the pieces placed between glances reflect the individual's interpretation of a structural relationship between those chess pieces.

This technique is limited to spatial domains, such as board games, where the individual applies a structure to an existing configuration. There is no reasonable analogy of this technique to domains without a strong spatial component such as programing. Because of the spatial domain restrictions, there is no concept elicitation step for this technique.

3.2.5 Discussion of Cluster Elicitation Techniques. As a whole, the clustering methodologies give the individual being measured control in deciding what the representation will look like, i.e., what the relationships between the concepts may be. Clustering methodologies inherently have problems because of the discrete nature of clusters. There are often occasions when it is neither appropriate to totally group or separate a pair of concepts. Those concepts that are weakly related may cause some amount of uncertainty on the part of the domain expert regarding whether to cluster the pair or not. The clustering methodologies do not provide for this possibility. The scaling methodologies are better at representing varying degrees of relationships.

3.3 Analysis of Representations

Upon review of the literature, a formal quantitative analysis methodology for cluster representations was not readily apparent for most clustering techniques. However, card sorts using repeated concepts may be converted to distance data using co-occurrence measures (See section 4.1.3). Due to this, the investigator or domain expert usually subjectively interprets the resulting representations.

3.4 Application of Clustering Methodologies to HCI Design

Clustering techniques inherently find relationships between concepts in the user's knowledge structure. This will aid the designer in fulling the requirements of the second stage of the design process, namely specifying interface characteristics and addressing consistency and structure requirements for the interface. For instance, clustering may be used to identify commands that are grouped together in the user's knowledge structure and that should, therefore, be placed in close proximity in the interface design.

The strength of these techniques is that the individual being measured has a significant amount of flexibility in producing a representation that best matches their particular knowledge organization. While at the same time the investigator controls the scope of the concepts being examined, and can thus restrict the domain to concepts relevant to the issues at hand.

4.0 Scaling Methodologies

The scaling methodologies can be logically subdivided into four stages: concept elicitation, relationship elicitation, representation development, and representation analysis. Concept elicitation involves the elicitation of a list of domain relevant concepts, as discussed above. This is followed by relationship elicitation in which the interrelationships between the concepts are determined. This data is then submitted to a scaling algorithm in the representation development phase to produce a representation. Finally, this representation is examined in the representation analysis stage. Here, an investigator attempts to draw conclusions about the representation.

There are cases, such as the repertory grid technique, where the concept and relationship elicitation stages may be simultaneous. While one might argue that the separation of concept and relationship elicitation is an unnatural and constrictive way for humans to elicit structural information, separation of these two stages distinguishes the scaling and clustering methodologies from verbal reports. In addition, certain potential applications of this work (e.g. personnel selection and training) emphasize the consideration of individual differences in human knowledge structures. However, the most popular approach is to obtain a common set of concepts for the whole group and then apply a particular relationship elicitation technique. This method reduces individual difference variance in the data analysis.

4.1 Concept and Relationship Elicitation

The input for scaling techniques (with the possible exception of list generation) is a list of domain relevant terms obtained from the concept elicitation stage and their degree of relationship. The techniques for eliciting concepts were described in section 3.1 above. The techniques for eliciting the relationships between the concepts are described below. The output of the scaling techniques typically entails a set of numbers, in the form of a matrix, in which each number describes the psychological distance, or similarity, between a pair of concepts. Often this matrix is presented graphically.

4.1.1 Pairwise Similarity Ratings. The technique of pairwise similarity ratings is based on the hypothesis that individual's judgement or rating of the conceptual similarity of (or distance between) two concepts is directly related to the psychological distance between the two concepts in the person's memory. In this technique, individuals supply a similarity judgement for every possible pair of concepts. The pairwise similarity rating technique has been one of the most commonly used techniques for the scaling methodologies. Pairwise similarity ratings take the "microscopic" approach with regard to knowledge elicitation. That is, the individual is presented only two concepts at a time and asked to rate the degree of similarity between these concepts.

Schvaneveldt et. al. (1982) and Schvaneveldt, Durso, Goldsmith, Breen, Cooke and De Maio (1985) used pairwise similarity ratings in the context of flight maneuvers of pilots and pilot trainees. Cooke and Schvaneveldt (1988) implemented pairwise similarity ratings on abstract programming concepts. Koubek and Mountjoy (1991) used pairwise similarity ratings in the domain of word processing. Hopkins, Campbell, and Peterson (1987) obtained judgements of the relative predictability of values and properties in a heart vessel system. Many other instances of the use of pairwise similarity ratings are found in the literature (e.g. Enkawa and Salvendy, 1989; Esposito, 1990; Dayton, Durso, and Shepard, 1990).

4.1.2 Repertory Grid. A central contribution of the repertory grid approach is the idea that concepts may be related on a variety of dimensions. As an alternative to pairwise similarity ratings where all dimensions of comparison are assumed to be lumped into one judgement (i.e. similarity), repertory grid allows the researcher to obtain a clear idea of the basis by which similarity judgements are made.

The repertory grid technique is based on personal construct theory proposed by Kelly (1955) for clinical psychology applications. Boose (1985, 1986) provides a thorough explanation of the development of personal construct theory for applications in measuring knowledge representations. Colthart and Evans (1981) used the repertory grid technique to elicit knowledge structures concerning bird taxonomies. Details regarding the repertory grid methodology and applications can be found in Boose, 1985, Olson and Rueter, 1987, Shaw, 1980 and Keen and Bell, 1981.

4.1.3 Co-occurrence Analysis. Co-occurrence is a statistical term that is a measure of the likelihood that two concepts will appear together. The most commonly used measure of co-occurrence is conditional probability. The actual task involved here is identical to the card sorting task mentioned in the clustering methodology section. Here, however, several methods are available which convert single iteration card sorts into a proximity matrix. Using a group of participants, one can measure the conditional probability that two concepts will be clustered together across all of the individuals' card sorts. Alternatively, one may have a single person complete many card sorts, using the same list of concepts, and then compute conditional probabilities across trials. Other statistical measures of co-occurrence have also been investigated (McDonald, Plate, and Schvaneveldt, 1990).

4.1.4 Recall Techniques. There are recall techniques that can produce proximity data. One of these techniques is very similar to the card sorting methodology. In this procedure, the person recalls the concept list several times and the analyst computes the conditional probability that two items are adjacent. This procedure suffers from the fact that many of the concepts will never be adjacent (Cooke and McDonald, 1987). Another way that proximity information can be obtained is by the measurement of inter-item distance between concepts in the sequential recall list. Thus, if a pair of concepts have five items between them in a particular recall list, the psychological distance between the items is defined to be five. Friendly (1977) also discusses this method. Repeated measures can be used, along with the use of rotating "seeds" or cue concepts (Reitman and Rueter, 1980; Gammack, 1990), to obtain measures of the consistency of the inter-item distance.

Event Record Analysis is similar to a recall list in the sense that the relationships between concepts are obtained by processing sequential records of concepts. However, instead of the list being obtained from a standard set of concepts, the list is derived from a protocol or observation of a domain related task (Cooke and McDonald, 1987). The methods of obtaining similarity information from this data are identical to that of the inter-item distance measures for recall measures. Finally, the twenty question technique is based on the idea that if the individual tries to guess the identity of a hidden target item, they will tend to ask questions that will narrow the possible alternatives. By recording the questions asked and the concepts eliminated after each question, one obtains information about the interrelationships between the concepts. Gammack (1990) used this technique to classify locomotives.

4.2 Representation Development

Inputs to these representation generation techniques generally are similarity matrices as described in the previous section. The outputs of these techniques are spatial, symbolic representations of the knowledge structures.

4.2.1 General Weighted Networks (GWN). General weighted networks are structures based on graph theory (Harary, 1969). Concepts are represented as points called "nodes" in the pictorial representation. The relationship between concepts are represented as "links" which are lines that connect certain pairs of nodes. Each link has a weight which is equivalent to the distance measure obtained for the two concepts that the link connects. A series of links (in which all nodes are distinct) is called a path. General weighted networks reduce a set of inter-concept distance measurements by only including the most salient connections in the network. The basic idea behind the construction of a GWN is as follows: A link between any two nodes is included in the network if and only if the weight of that length is at least as small as the combined weight of any other possible path that connects the two nodes (Dearholt and Schvaneveldt, 1990).

There are several different applications in the literature where a GWN has been used as a technique to represent knowledge structures. Schvaneveldt, Durso, and Dearholt (1987) developed an algorithm called Pathfinder which produces a class of networks known as PFnets. Two parameters control the complexity and characteristics of the network. The r parameter determines how the paths between any two nodes are calculated. The selection of r also determines the measurement assumptions in the data (e.g. when r is set to infinity, ordinal data is assumed). The q parameter is the maximum number of links between two concepts that will be considered for inclusion in the network. PFnets have been widely used as a scaling methodology for knowledge structure measurement (Schvaneveldt et al., 1985; Cooke, Durso, and Schvaneveldt, 1986; Cooke and Schvaneveldt, 1988). Hopkins et al. (1987) use another type of GWN called a directed graph (Miyamoto and Nakayama, 1984). All GWN's function basically the same way. They only differ in the way that path lengths are computed and in the provisions for link inclusion.

4.2.2 Multidimensional Scaling (MDS). Multidimensional scaling is a mathematical procedure with its roots found as early as 1928 (Schiffman et al., 1981). MDS represents the similarity of concepts spatially on an n-dimensional map. In particular, the more similar two concepts are, the closer together they appear on the map.

According to Schiffman et al. (1981), MDS models are algebraic equations which summarize certain information in the data. These algebraic models have geometric complements which yield the characteristic resultant spatial map of concepts.

MDS has been used in a variety of applications over its lengthy history, and in addition, has been implemented into many software packages (e.g. ALSCAL, MINISSA, POLYCON, INDSCAL/SINDSCAL, KYST, and MULTISCALE). Polzella and Reid (1989) employed MDS techniques to discover differences in performance characteristics of expert and novice combat pilots. Polzella and Reid argue that MDS allows deeper insight than simpler metrics "...into the underlying representation associated with performance of a complex task" (p.141). Similarly, MDS has been used to evaluate cognitive representations of fighter pilots (Schvaneveldt et al., 1982; Schvaneveldt et al., 1985). Other areas of application have included memory organization (Cooke, Durso and Schvaneveldt, 1986), the analysis of spatial aptitude (Pellegrino and Kail, 1982), the representation of perceived relations among properties of the human heart/blood vessel system (Hopkins et al., 1987), the changes in structure of text-editing commands with experience (Kay and Black, 1984) and the determination of the interaction between learning processes and the cognitive representation of problem solving (Enkawa and Salvendy, 1989).

4.2.3 Hierarchical Clustering Schemes (HCS). The final technique to be discussed is hierarchical clustering schemes (HCS). This technique differs from those listed in the clustering section, which allow the individual to directly produce the knowledge representation, by generating the knowledge representation through the use of an outside algorithm. Johnson (1967) is generally credited with the development of a particular set of hierarchical clustering schemes that are popular in the cognitive measurement literature. It should be noted that the clustering process described by Johnson is only one of many clustering methods currently available. Romesburg (1984) describes many other clustering techniques.

HCS has been used extensively in the literature. Once again, computer-oriented knowledge has been a popular domain for the application of clustering schemes. Adelson (1981) used HCS techniques to discover the differences in the way that expert and novice programmers organize programming concepts. The same type of study was also performed by McKeithen et al. (1981). Kay and Black (1984) employed HCS as a method to determine the changes in knowledge structure as novices developed into experts in text-editing. Other areas of application have included the work of Hopkins et al. (1987) which looked at the understanding of relationships between features of the human cardiovascular system, Schoenfeld and Herrmann's (1982) study of novice/expert perception of mathematical problems, and Schvaneveldt et al. (1982) who were concerned with the organization of critical flight information in memory.

4.2.4 Discussion of Representation Development Techniques. Several studies have been completed which quantitatively compare the performance of various representation development techniques. Schvaneveldt et. al. (1985) compared the ability of MDS, GWN, and raw proximity ratings to predict membership of an individual into groups of instructor pilots, guard pilots and novice pilots. They found that both MDS and GWN outperformed the raw ratings in predictive success. Further, they found that the MDS representation marginally outperformed the GWN. Goldsmith and Johnson (1990) obtained somewhat different findings in the domain of classroom learning. They found that GWN generally outperformed both MDS representations and raw ratings in the prediction of students' final grades based on the match between the students' and the instructor's structures. As a whole, the representation generation methodologies provide a way for the researcher to obtain a quick, pictorial summary of the knowledge structure.

4.3 Analysis of Representations

One of the criticisms with the use of the scaling methodologies is that once the pictorial representation is obtained, it is often unclear how this information can be applied to solve problems. Certainly, the representation may be subjectively evaluated by the analyst or domain experts. To some degree, this path of analysis results in the same difficulties as the verbal report methodologies discussed earlier. Granted, these representations certainly can take the role of supporting the traditional knowledge elicitation process (Olson and Rueter, 1987). However, for a variety of computer system design applications, it is important to find ways to quantitatively evaluate and compare structures across individuals. While several existing techniques can analyze representations, only comparative analyses are provided.

4.3.1 Analysis of General Weighted Networks There are a variety of methods available that attempt to assess the similarity or relatedness of two networks (structural similarity should not be confused with concept similarity, such as in pairwise similarity ratings). these include Subentities (Graham, 1987), Path Lengths (Goldsmith and Johnson, 1990), Neighborhoods (Goldsmith and Davenport, 1990) and Feature Analysis (Schvaneveldt et al., 1985). See Goldsmith and Davenport (1990) for a formal mathematical treatment of these techniques.

4.3.2 Analysis of Multidimensional Scaling Solutions and HCS The similarity of two MDS solutions can be obtained by calculating the Pearson correlation coefficient of the corresponding euclidean distances between each concept pair for the two solutions. Although this method does not indicate how the individuals are similar, it at least provides enough information to determine which individuals' representations are most alike. Goldsmith and Johnson (1990) used this technique to compare MDS solutions of students with the instructor's representation. INDSCAL (individual differences MDS), developed by Carroll and Chang (1970), may be used in order to see where certain groups, or individuals, are located with respect to the dimensions derived from the original MDS solution (Schvaneveldt et al. 1985).

Schvaneveldt et al. (1985) also describe the use of linear discriminant functions in order to classify individuals into two or more groups based upon their cognitive structures. Basically, this method utilizes vectors in order to divide the solution space into regions which are comprised only of particular types of patterns (e.g. expert/novice). This technique has been used by Schvaneveldt et al. (1982,1985) to classify experienced and unexperienced pilots based upon the representations of their understanding of flight concepts. Schvaneveldt et al. (1982) also analyzed PFnets in this manner. At present, HCS output is left to experimenter judgement and ad hoc algorithms (Koubek and Mountjoy, 1991).

4.4 Application of Scaling Methodologies to HCI Design

Scaling techniques provide a much more quantitative representation of the user's knowledge structures, and therefore are capable of distinguishing differences between highly similar structures. This makes them highly useful in the third stage of the design process, summative evaluation and installation. These sensitive techniques can be used to measure knowledge structures resulting from alternative interface designs. As such, they are a tool for evaluating and selecting interfaces which provide or produce more accurate and usable knowledge structures.

5.0 Conclusion

In conclusion, knowledge structures are being identified as a significant determinant of operator performance and one for which the HCI specialist must consider when designing interfaces and interaction sequences. In review of the available literature three classifications of knowledge structure measurement techniques emerge. Furthermore, these three categories have the potential of providing a variety of information to the human computer interface designer with varying scope and quantitative value. The high level or global task information provided by the first classification of techniques, verbal reports, is an important consideration in the early stages of the design process. The clustering methodologies are a source of concept relationship information, which is useful in specifying consistency and structure requirements in the inital design stages of interface design. The quantitative nature of the scaling methods enable the interface designer to evaluate and compare the user structures resulting from alternative designs during the evaluation stage of the design process. Therefore, using these knowledge structure measurement techniques presented here in conjunction with nearly any formal design methodology, a designer will be able to identify the knowledge structure of their user population and incorporate this knowledge into an interface best suited for their user group.

REFERENCES

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. *Memory & Cognition*, 9, 422-433.

Adelson, B. (1984). When novices surpass experts. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 483-495.

Anderson, J. R. (1983). *The Architecture of Cognition*, London, England: Harvard University Press.

Barfield, W. (1986). Expert-novice differences for software: Implications for problem solving and knowledge acquisition. *Behaviour and Information Technology*, 5, 15-29.

Barfield, W., Koubek, R. J., Hwang, S. (1985). Eliciting expert behavior for software: A cognitive psychology perspective. In R. Eberts and C. Eberts (Eds.), *Trends in Ergonomics/Human Factors II*, North Holland Publishing: New York. 303-310.

Boose, J. H. (1985). A knowledge acquisition program for expert systems based on personal construct psychology. *International Journal of Man-Machine Studies*, 23, 495-525.

Boose, J. H. (1986). *Expert Transfer for Expert Systems Design*, New York: Elsevier.

Booth, P. A., (1989). *An Introduction to Human-Computer Interaction*, Earlbaum: London.

Bower, G. H., and Springsteen, F. (1970). Pauses as recording points in letter series. *Journal of Experimental Psychology*, 83, 421-430.

Card, S. K., Moran, T. P., and Newell, A. (1983). *The Psychology of Human-Computer Interaction*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Carroll, J. D., and Chang, J. (1970). Analysis of individual differences in multidimensional scaling via and n-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35, 283-319.

Chase, W. G., and Simon H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.

Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.

Chi, M. Hutchinson, J., and Robin, A. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly*, 35, 27-63.

Colthart, V., and Evans J. St. B. T. (1981). An investigation of semantic memory in individuals. *Memory & Cognition*, 9, 524-532.

Cooke, N. M., Durso, F. T. and Schvaneveldt, R. W.(1986). Recall and measures of memory organization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 538-549.

Cooke, N., and McDonald, J. (1986). A formal methodology for acquiring and representing expert knowledge. *Proceedings of the IEEE*, 74, 1423-1431.

Cooke, N., and McDonald, J. (1987). The application of psychological scaling techniques to knowledge elicitation for knowledge-based systems. *International Journal of Man-Machine Studies*, 26, 533-550.

Cooke, N. J., and Schvaneveldt, R. W. (1988). Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies*, 29, 407-427.

Dayton, T. Durso, F. T., and Shepard, J. D. (1990). A measure of the knowledge reorganization underlying insight. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 267-277.

De Groot A. D. (1965). *Thought and choice in chess*, The Hague: Mouton.

Dearholt, D. W., and Schvaneveldt R. W. (1990). Properties of Pathfinder networks. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organization*, New Jersey: Ablex Publishing Corporation. 1-30.

Durso, F. T. and Coggins, K. A. (1990). Graphs in the social and psychological sciences: Empirical contributions of Pathfinder. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 31-51.

Eason, K. and Harker, S. (1991). Human factors contributions to the design process. In B. Shackel and S. J. Richardson (Eds.), *Human Factors or Informatics Usability*, New York: Cambridge Press. 73-96.

Egan, D. E., and Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. *Memory & Cognition*, 7, 149-158.

Enkawa, T., and Salvendy, G. (1989). Underlying dimensions of human problem solving and learning: Implications for personnel selection, training, task design and expert systems. *International Journal of Man-Machine Studies*, 30, 235-254.

Ericsson, A. K., and Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.

Ericsson, K. A., and Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*, Cambridge Massachusetts: MIT Press.

Esposito, C. (1990). A graph-theoretic approach to concept clustering. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 89-100.

Evans, B. T. (1988). The knowledge elicitation problem: A psychological perspective. *Behaviour and Information Technology*, 7, 111-130.

Friendly, M. L. (1977). In search of the M-gram: The structure of organization in free recall. *Cognitive Psychology*, 9, 188-249.

Gammack, J. G. (1990). *Expert conceptual structure: The stability of Pathfinder representations.* In R. W. Schvaneveldt (Ed.), Pathfinder Associative Networks, New Jersey: Ablex Publishing Corporation. 213-226.

Gardner, A., (1991). An approach to formalised procedures for user-centered system deign. In B. Shackel and S. J. Richardson (Eds.), *Human Factors or Informatics Usability*, New York: Cambridge Press. 133-150.

Gobbo, C., and Chi, M. (1986). How knowledge is structured and used be expert and novice children. *Cognitive Development*, 1, 221-237.

Goldsmith, T. E., and Davenport, D. M. (1990). Assessing structural similarity of graphs. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 75-88.

Goldsmith, T. E., and Johnson P. J. (1990). A structural assessment of classroom learning. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 241-254.

Graham, R. L. (1987). A similarity measure for graphs–reflections on a theme of Ulam. *Los Alamos Science*, 114-212.

Harary, F. (1969). *Graph Theory*, Reading, Massachusetts: Addison-Wesley.

Hardiman, P. T., Dufresne, R., and Mestre, J. P. (1989). The relation between problem categorization and problem solving among experts and novice. *Memory & Cognition*, 17, 627-638.

Henderson, A., (1991). A development perspective on interface, design, and theory. In J. M. Carroll (Ed.), *Designing Interaction: Psychology at the Human-Computer Interface*, Cambridge Press: New Yourk. 254-268.

Hirtle S. C., and Mascolo, M. F. (1986). Effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 182-189.

Hobus, P. P. M., Schmidt, H. G., Boshuizen, H. P. A., and Patel, V. L. (1987). Contextual factors in the activation of first diagnostic hypotheses: Expert-novice differences. *Medical Education*, 21, 471-476.

Hollands J. G., and Merikle, P. M. (1987). Menu organization and user expertise in information search tasks. *Human Factors*, 29, 577-586.

Hopkins, R. H., Campbell, K. B., and Peterson, N. S. (1987). Representations of perceived relations among the properties and variables of a complex system. *IEEE Transactions on Systems, Man, and Cybernetics*, smc-17, 52-60.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32, 241-255.

Kay, D. S., and Black, J. B. (1984). Changes in knowledge representations of computer systems with experience. *Proceedings of the Human Factors Society*, 28, 963-967.

Keen, T. R. and Bell, R. C. (1981). One thing leads to another: a new approach to elicitation in the repertory grid technique. In M. Shaw (Ed.), *Recent Advances in Personal Construct Technology*. New York: Academic Press.

Kelly, G. A. (1955). *The Psychology of Personal Constructs*, Norton: New York.

Kieras, D. E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Heilander (Ed.), *Handbook of Human Computer Interaction*. North-Holland: Elsevier Science Publishers.

Koubek, R. J., Salvendy, G., Eberts, R. E., and Dunsmore, H. (1987). Eliciting knowledge for software development. *Behaviour and Information Technology*, 6, 427-440.

Koubek, R. J., and Mountjoy, D. N. (1991). The impact of knowledge representation on cognitive oriented task performance. *International Journal of Human-Computer Interaction*, 3, 31-48.

Koubek, R. J., and Salvendy, G. (1991). Cognitive performance of super-experts on computer program modification tasks. *Ergonomics*, 8, 1095-1112.

Leinhardt, G. and Smith, D.A. (1985). Expertise in mathematics instruction: Subject matter knowledge. *Journal of Educational Psychology*, 77, 247-271.

McDonald, J. E., Paap, K. R., and McDonald, D. R. (1990). Hypertext perspectives: Using Pathfinder to build Hypertext Systems. In Schvaneveldt, R. W. (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 197-212.

McDonald, J. E., Plate, T. A., and Schvaneveldt, R. W. (1990). Using Pathfinder to extract semantic information from text. In R. W. Schvaneveldt (Ed.), *Pathfinder Associative Networks: Studies in Knowledge Organizations*, New Jersey: Ablex Publishing Corporation. 149-164.

McDonald, J. E. and Schvaneveldt, R. W. (1988). The application of user knowledge to interface design. In R. Guindon (Ed.), *Cognitive Science and its Applications for Human-Computer Interaction*, Hillsdale, NJ: Erlbaum. 289-338.

McKeithen, K. H., Reitman, J. S., Rueter, H., and Hirtle, S. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.

McLean, R. S., and Gregg, L. W. (1967). Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology*, 74, 455-459.

McPherson, S. L., and Thomas, J. R. (1989). Relation of knowledge and performance in boys' tennis: Age and expertise. *Journal of Experimental Child Psychology*, 48, 190-211.

Means, M. L., and Voss, J. F. (1985). Star wars: A developmental study of expert and novice knowledge structure. *Journal of Memory and Language*, 24, 746-757.

Meister, D. (1985). *Behavioral Analysis and Measurement Methods*, New York: John Wiley & Sons.

Miyamoto, S., and Nakayama, K. (1984). A directed graph representation based on a statistical hypothesis testing and application to citation and association structures. *IEEE*, smc-14, 203-212.

Murphy, G. L., and Wright, J. C. (1984). Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 144-155.

Musen, M. A. (1989). *Automated Generation of Model-Based Knowledge-Acquisition Tools*, Pitman, London: Morgan Kaufmann Publishers, Inc.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*, Englewood Cliffs, New Jersey: Prentice-Hall.

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.

Olson, J.R. and Olson, G.M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Human-Computer Interaction*, 5, 221-265.

Olson, J. R., and Rueter, H. H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems*, 4, 152-168.

Palmer, E. S. (1978). Fundamental aspects of cognitive representation. In E. Rosch and B. B. Lloyd (Eds.), *Cognition and Categorization*, 259-303. Hillsdale, NJ: Earlbaum.

Payne, S. J. and Green, T. R. G. (1986). Task-action grammars: a model of mental representation of task languages. In T. P. Moran (Ed.), Human-Computer Interation, New Jersey: Lawrence Erlbaum Associates. 2, 93-103.

Pellegrino, and Kail (1982). Process of spatial aptitude. In R. S. Sternberg (Ed.), *Advances in the psychology of Human Intelligence*, New Jersey: Lawrence Erlbaum Associates. 1, 311-365.

Ploger, D. (1988). Reasoning and the structure of knowledge in biochemistry. *Instructional Science*, 17, 57-76.

Polzella, D. J., and Reid, G. B. (1989). Multidimensional scaling analysis of simulated air combat maneuvering performance data. *Aviation, Space, and Environment Medicine*, 60, 141-144.

Reitman, J. S. (1976). Skilled perception in go: Deducing memory structures from inter-response times. *Cognitive Psychology*, 8, 136-156.

Reitman, J. S., and Rueter, H. R. (1980). Organization revealed by recall orders and confirmed by pauses. *Cognitive Psychology*, 12, 554-581.

Rips, L., Shoben, E., and Smith, E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 12, 1-20.

Romesburg, H. C. (1984). *Cluster Analysis for Researchers*, California: Lifetime Learning Publications.

Schiffman, S. S., Reynolds, M. L., and Young, F. W. (1981). *Introduction to Multidimensional Scaling*, New York: Academic Press.

Schoenfeld, A. H., and Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 484-494.

Schvaneveldt, R. Durso, F., Goldsmith, T., Breen, T., Cooke, N., Tucker, R., and De Maio, J. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies*, 12, 699-728.

Schvaneveldt, R. W., Durso, F. T., and Dearholt, D. W. (1989). Network structures in proximity data. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, New York: Academic Press. 249-284.

Schvaneveldt, R. W., Goldsmith, T. E., Durso, F. T., Maxwell, K., Acosta, H. M., and Tucker, R. G. (1982). *Structures of Memory of Critical Flight Information*, (Report AFHRL-TP-81-46). Las Cruces, New Mexico: New Mexico State University, Department of Psychology.

Shaw, M. L. G. (1980). *On Becoming a Personal Scientist.* London: Academic Press.

Soloway, E., and Adelson, B. (1985). The role of domain experience in software design. *Transactions on Software Engineering*, SE-11, 1351-1360.

Sweller, J., Mawer, R. F., and Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology: General*, 112, 639-661.

Sweller, J., and Owen, E. (1985). What do students learn while solving mathematics problems. *Journal of Educational Psychology*, 77, 272-284.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, 69, 344-354.

Vessey, Iris (1985). Expert in debugging computer programs: A process analysis. *International Journal of Man-Machine Studies*, 23, 459-494.

Weiser, M., and Shertz, J. (1983). Programming problem representation in novice and expert programmers. *International Journal of Man-Machine Studies*, 19, 391-398.

Zeitz, C. M., and Spoehr, K. T. (1989). Knowledge organization and the acquisition of procedural expertise. *Applied Cognitive Psychology*, 3, 313-336.

ATTACHMENT 4

The Training of Knowledge Structures for Manufacturing Tasks: An Empirical Study

Richard J. Koubek, Timothy P. Clarkston, and Vance Calvez

In (PRESS): Ergonomics

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1256 and # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

The Training of Knowledge Structures for Manufacturing Tasks: An Empirical Study

Acknowledgment: This research was supported by the Office of Naval Research Cognitive Sciences Program under grants # N00014-90-J-1256 and # N00014-92-J-1153. The opinions expressed in this paper do not necessarily reflect the position of ONR.

AUTHORS

Richard J. Koubek[1], Timothy P. Clarkston[1], and Vance Calvez[2]

[1]School of Industrial Engineering
1287 Grissom Hall
Purdue University
West Lafayette, IN 47906-1287

[2]Department of Biomedical and Human Factors Engineering
Wright State University
Dayton, OH 45435

ABSTRACT

Knowledge structure refers to the manner in which a human organizes knowledge with a given domain. Research has identified knowledge structure as a determinant of the human ability to perform cognitive-oriented tasks. Yet uncertainty still exists about how to improve an individual's cognitive task performance through the controlled utilization of the individual's knowledge structure. The purpose of this study is to investigate whether the development of individual's knowledge structure in a particular domain can be manipulated through training. The experiment utilized the manufacturing domain of plastic extrusion machine operation. Sixteen subjects, having no previous knowledge of the domain, were randomly assigned to one of two experimental groups. Each of the experimental groups corresponded to a distinct training condition. Over a three day period, both training groups received the same instructional content; however, the sequence in which the training material was presented differed. One group initially received the abstract, conceptual relationships between domain concepts, followed by more detailed relationships associated with the lower level aspects of the domain. The other group received the training material in the reverse order; i.e. the lower level information followed by the abstract. Prior to and concluding the training sessions, each individual's knowledge structure was assessed along two dimensions, Hierarchical Levels and Multiple Relations, through a computer-based measurement technique entitled KSAT. The group which received the abstract relationships first showed significant improvement following training along both dimensions of knowledge structure. No significant changes in the knowledge structure dimensions were found for the group which received the lower level relationships first. This study suggests that an individual's knowledge structure can be manipulated through training, with a significant effect being attributed to the training sequence of abstract material followed by the more detailed material.

KEYWORDS: Knowledge Structure, Training, Skill Acquisition

INTRODUCTION

Background

The manufacturing environment is in a state of change. Mass production is giving way to new styles of manufacturing such as lean production, agile manufacturing, and customer-oriented manufacturing. As a direct result, jobs within the manufacturing facilities are also changing. The tightly controlled, narrow, highly repetitive manual tasks associated with mass production are being augmented or replaced by broad-based jobs with cognitive-oriented task components. Workers are sharing responsibility for plant functions while participating in production scheduling, maintenance, task assignments, and process control. Higher levels of decision making are being performed by the human operators. With this shift in task responsibilities, the ability of the human operator to perform cognitive-oriented tasks is becoming of increased importance.

Study of the physical aspects of work through ergonomics can maximize the performance of the human operator, while improving the overall situation of the human operator within the system. These gains are attained through the proper design of the tasks, tools, and workplace, as well as understanding the role of the human operator in the system (Corlett 1973). Just as with the physical domain, the cognitive components of tasks can be addressed to improve system productivity. Cognitive task performance can be attributed to many factors, such as knowledge content, innate cognitive skills and abilities, mental strategies, motivation, and degree of practice. In addition to these, recent research has shown the manner in which domain knowledge is organized to be a primary determinant of human cognitive task performance (Koubek and Salvendy 1991).

Previous research into the impact of knowledge organization on cognitive task performance has focused on the differences in the representation of domain knowledge between subject groups at different skill levels. Experts have been found to form abstract or conceptual representations of the domain. Novices, on the other hand, tend to focus on the surface or concrete features of the domain. For example, Adelson (1981) presented two groups of computer programmers, experts and novices, with randomized computer code. Based on a conceptual understanding of the domain, the experts reorganized the code into actual programs, while the novices, utilizing syntactic similarities, categorized the code according to the control words. Similar results have been found across a variety of domains: chess (de Groot 1965; Chase and Simon 1973), word processing (Koubek and Mountjoy 1991), problem solving (Hardiman et al. 1989; Chi et al. 1981; Lewis 1981; Bhaskar and Simon 1977; Schoenfeld and Herrmann 1977), software design (Barfield 1986; Jeffries et al. 1981), computer programming (McKeithen et al. 1981; Schneiderman 1977), and medical diagnosis (Hobus et al. 1987).

The above studies suggest that the higher levels of cognitive task performance are achieved, in part, through a more hierarchical or conceptual representation. Adelson (1984) confirmed that the expert's representation is more abstract and the novice's representation is more concrete. However, this study demonstrated that the performance of the novice group can be better than that of the expert group, in terms of correct responses. Of importance, the situation in which the novices exceeded the experts is attributed to the knowledge structure of the novices matching the requirements of the task better than that of the experts. Further evidence is found in a recent study by Koubek and Salvendy (1991) in which the representations, or organizations of domain knowledge, were examined for a group of experts and a group of super-experts while performing a computer program modification task. The super-experts were determined to be in the top 95th percentile of their field, while the experts were in the 70th to 80th percentile. The group of super-experts focused on abstract information attempting to understand the overall program prior to beginning the modification task. The experts tended to focus on task specific information limited to the modification task. It appears that the type of knowledge representation best suited to maximize the human operator's cognitive task performance may be task dependent, i.e. a function of the task characteristics and the attributes of the representation types.

In order to further examine the role of knowledge structure, a unified understanding of knowledge structure is needed. Current knowledge structure measurement techniques each have particular assumptions with respect to the components of human knowledge structure. Typically, studies in knowledge structure employ measurement techniques without having an *a priori* operational definition of the attributes they intend to measure. Various methodologies were developed specifically to measure certain features of knowledge structure. Consequently, the findings about human knowledge structure are quite dependent on the assessment technique employed and lack commonalty. Attempting to advance the systematic study of knowledge structure, Koubek (1991) formulated the following definition of human knowledge structure: "Human knowledge structure is defined as the structure of interrelationships between elements, concepts and procedures, in a particular domain, organized into a unified body of knowledge."

Elements are unique units of knowledge within a domain. The format of the elements can either be declarative or procedural. Declarative elements, referred to as concepts, contain knowledge in the from of facts. Procedural elements, or procedures, contain knowledge in the form of how to perform functions within the domain. Within the knowledge structure, the elements may exist at various levels of abstraction or hierarchy. For example, two elements within the discrete parts manufacturing domain may be "Milling machine" and "Flexible manufacturing cell". The milling machine is considered to be at a lower level of abstraction than, or subordinate to, the flexible manufacturing cell because it is part of the flexible manufacturing cell.

The structure of interrelationship denotes the relationships, or links, that can exist between the domain elements. As previously described, the interrelationships may be super- or sub-ordinate between two elements. It is possible that more than one relationship could exist between two elements. Given that a relationship does exist between a pair of elements, the strength of that relationship, i.e. the degree of to which that relationship applies, can vary. For example, "Volume" and "Temperature" are two possible domain elements a human operator may associate with the task of controlling a chemical manufacturing process. Given that both quantities are presented visually, the operator may relate volume and temperature in terms of their physical proximity on the visual display terminal. Yet, a second and stronger relationship may exist between volume and temperature concerning their interaction in the process, such that changes in volume are proportional to change in temperature.

Given the importance of knowledge structure to cognitive task performance, uncertainty still remains about how to capitalize on an individual's knowledge structure to improve productivity. The desirable goal would be to equip a workforce with the appropriate knowledge structure for a specific task assignments in order to maximize performance. At this point, however, it is still unknown whether an individual's development of knowledge structure within a domain can be manipulated through training or is solely a function of individual differences. This information is necessary to determine whether knowledge structure-task matching must rely primarily on selecting personnel with the ability to develop the appropriate knowledge structure, or whether personnel can be trained to obtain the most suitable knowledge structure. The purpose of this study is to examine the effect of training on the development of an individual's knowledge structure, as tested by the two hypotheses below.

Hypothesis one. The number of *Hierarchical Levels* developed within an individual's knowledge structure is a function of the training which an individual receives.

Hypothesis two. The number of *Multiple Relations* between concepts developed within an individual's knowledge structure is a function of the training which an individual receives.

METHODOLOGY

Subjects

Sixteen university students participated in this study. The subjects had no previous knowledge of the experimental domain of plastic extrusion machine operation. The range of ages for the subjects was 19 to 34 with a mean of 23. A post-hoc analysis of subject characteristics indicated that college major, age, and general manufacturing experience were balanced between the experimental groups.

Experimental Design

The experiment was performed as a 2 X 2 repeated measures ANOVA design with one between subjects variable, Training Condition, and one within subjects variable, Phase. Training Condition had two levels, Abstraction-First and Relations-First. Each of the two levels of Training Condition corresponded to a distinct training sequence; thus, forming two experimental groups. Subjects were randomly assigned to the experimental groups. Phase also had two levels, Initial Assessment and Final Assessment. These levels represent the point in each training sequence where the dependent variables were measured; i.e. when the individual's knowledge structure was assessed. The independent variables are summarized in Table 1.

Insert Table 1 about here.

The two dependent variables were Hierarchical Level and Multiple Relations. Since two dependent variables were used, the experimental design was essentially implemented twice, once for each dependent variable.

Measurement of the Dependent Variables with KSAT

Overview. The Knowledge Structure Assessment Technique (KSAT), developed for this study, is a computer-based methodology to measure two dimensions of human knowledge structure. The software elicits information from the subjects through the pairwise comparison of domain concepts. The software then analyzes the collected data and calculates values for the two knowledge structure dimensions of interest in this research, Hierarchical Level and Multiple Relations.

Assessment process. KSAT elicits information about an individual's knowledge structure through an exhaustive, pairwise comparison of domain concepts. Two lists are essential to the assessment process. One is the concept list, which contains 25 items relevant to the stimulus domain of plastic extrusion machine operation. The complete concept list is presented in Appendix 1. The second list, relationships, provides subjects with 68 types of relationships. The relationship list is presented in Appendix 2. The assessment process is as follows:

(1) One concept from the concept list is randomly selected to be the focal item.

(2) The focal item is paired with the remaining concepts. For each pair,

the subject must indicate with a 'Yes' or 'No' if at least one significant relationship exists between the two concepts.

(3) The items identified in step (2) as being related to the focal item are then compiled. Utilizing the relationship list, the subject views each pair and enters the relationship(s) which characterizes each pair.
(4) The subject repeats steps (1) - (3) with another concept serving as the focal item. The process continues until all of the concepts on the concept list have served as the focal item.

Calculation of Variables. Hierarchical Levels is equal to the number of levels of layers within an individual's knowledge structure based on a tree-like representation created by KSAT. For example, the knowledge structure depicted in figure 1, Hierarchical Levels is equal to 3. In order to quantify Hierarchical Levels, KSAT constructs the tree-like representation of a subject's knowledge structure from the pairwise comparison information elicited from the subject. KSAT utilizes a sorting algorithm which essentially reorganizes the domain concepts as provided by the subject in terms of super- and sub-ordinate relationships. Key relationships from the relationship list were determined to denote super- and sub-ordinate relationships. The three types of relationship are: example-class, part-whole, and attribute-object. When the sorting is complete, the algorithm determines the maximum number of levels in the structure.

Insert figure 1 about here.

Multiple relations is equal to the average number of relationships that exist between a pair of items. Based on the subject's response to the pairwise comparison of all the domain concepts, multiple relations is calculated by dividing the total number of relationships that the subject entered by the number of pairs for which at least one relationship was indicated to exist. The equation is shown below.

$$\text{Multiple Relations} = \frac{\text{Total \# of pairwise relationships for all pairs}}{\text{Total \# of pairs with at least one relationship}}$$

Development process. KSAT was developed in accordance with the model of knowledge structure previously described. The interrelationships in the model can be characterized by super- and sub-ordinate relationships between domain elements and by multiple relationships between two domain elements. KSAT was designed to quantify both types of interrelationships. Hierarchical Levels indicates the number of levels in a representation created from the super- and sub-ordinate relationships. Multiple relations corresponds to the average number of relations between two domain elements. KSAT was designed to produce non referent measures of an individual's representation of the domain, not the difference between the individual's representation and some ideal or assumed correct representation.

The following set of requirements were compiled at the beginning of the KSAT development process. The technique should be:

(1) Based on a theoretical model of the human knowledge structure.
(2) Designed to leverage from the strengths and minimize the weaknesses of the existing measurement techniques. (A complete review of existing knowledge structure measurement techniques is presented by Koubek, Benysh, and Calvez (1993)).
(3) Capable of producing measures which can be meaningfully compared across domains.
(4) Objective and quantitative.

(5) Reliable and operationally efficient.

KSAT was designed as a tool to be used for the assessment of an individual's knowledge structure. The current version of KSAT was proceeded by two prototypes in the development process. With each design iteration, significant design changes occurred as the result of pilot studies and an increased understanding of human knowledge structure assessment.

Key design features. Based on strengths and weaknesses of the existing techniques, the following KSAT features are imperative to the assessment process. Knowledge is elicited through an exhaustive process of pairwise comparisons; i.e., all possible combinations of items are compared. All of the judgments performed by the subjects are limited to a specific pair of items at a time. Throughout the assessment process, subjects have the ability to review and revise all information entered.

Of particular importance to the assessment process is the elicitation of the relationships between domain concepts from the subjects. The subject's initial decision regarding relatedness of a pair of items is a yes/no decision. The KSAT procedure allows for multiple relationships between a pair of items to be entered by the subject. Given two items are related, subjects can classify the link or links between the pair of items using a standardized list of relationship types. The KSAT relationship list incorporates declarative (semantic) relationships, as well as procedural (action-oriented) relationships. Relationships may be directional, that is apply only one way between a pair of concepts. In addition, the relationship types can be generic to any domain, or domain-specific, related only to the experimental domain.

Experimental Procedure.

Each subject participated in the experiment over a three-day period. Each day, the subjects received individual training which lasted approximately one hour and consisted of the subject studying the appropriate training document until they felt they had mastered the material. KSAT was administered immediately following the training session on the first day (Initial Assessment) and the third day (Final Assessment). Requiring about two hours per administration, the subjects entered data into KSAT according to the previously described assessment process.

On the first day, the same instructional material was presented to both experimental groups. The material was designed to familiarize the subjects with the domain of plastic extrusion machine operation. The two groups received different instructional material on the second day. One group studied material which stressed the abstract, or hierarchical, relationships between the domain concepts (Abstraction-First), while the other group read material which emphasized the multiple relationships between domain concepts (Relations-First). On the third day, the instructional material was identical for both groups, incorporating a combined emphasis on both the hierarchical and multiple relationships. Figure 2 displays a flowchart of the testing procedure.

Insert figure 2 about here.

Thus, by the end of third day, both groups had received the same instructional content. The only difference between the two groups was the order in which the instructional content was presented to them. Consequently, two distinct sequences of training, or Training Conditions, were formed.

Development of Training Documents

The training documents were systematically manipulated along two dimensions: abstraction and relations. The abstraction dimension refers to the degree to which a hierarchical structure between the domain concepts is presented within the training document. For the 25 domain concepts, a low level of abstraction within a training document is represented by the pattern of relationships previously shown in figure 1. For comparison, a high level of abstraction for the same 25 concepts is presented in figure 3. As stated above, abstraction was present at the low level to both groups on the first day.

Insert figure 3 about here.

The abstraction dimension was implemented by eliminating or including references to hierarchical relationships between domain concepts within the training documents, as previously illustrated in Figures 1 and 3. An example of a abstract relationship is as follows, "The shot chamber is actually the forward part of the barrel", with 'part of' denoting a hierarchical relationship between shot chamber and barrel.

The relations dimension indicates the number of citations to pairwise multiple relations within a given training document. The high level relations training document contained approximately three times the references to multiple relations than the low level relations training document. The document with the low level relations was read by both groups on the first day. The Relations-First Training Condition group was presented with the high level of relations document on the second day.

The relations dimension was implemented by removing or adding references which suggest more than one relationship between a pair of concepts. For example, the following passage is from the low level relations document, "...operator only prepares the product for shipping by packing the product in the shipping container." From the high level relations document, the same passage reads, "...operator only inspects the product for defects, removes the gates and sprue from the product, as well as any flash, packs the product for shipping by packing the product in the shipping container."

The relations dimension was also implemented by removing or adding references to multiple functions of a single concept. According to the low relations training document, for instance, the function of the screw is to inject and transport material. The same part in the high relations document is reported to have the functions of transporting, mixing, injecting, and halting material.

RESULTS

Hypothesis One. The number of *Hierarchical Levels* developed within an individual's knowledge structure is a function of the training which an individual receives.

The results for Hypothesis One are presented graphically in figure 4. The effect of Training Condition was not significant at the $p < .05$ level ($F(1,14) = .09$; $p > .76$); thus, no conclusion can be drawn regarding the difference between the two Training Conditions (Abstraction-first and Relations-first) on the development of Hierarchical Levels. As expected, the Phase effect was significant ($F(1,14) = 6.59$; $p < .023$), indicating that the number of Hierarchical Levels increased with training. Although not significant at the .05 alpha level, there does appear to be a trend toward an interaction between Training Condition and Phase ($F(1,14) = 3.85$; $p < .07$) as displayed in figure 4.

Insert figure 4 and table 2 about here.

The interaction between Training Condition and Phase was examined further through post hoc analysis. The purpose was to determine if the number of Hierarchical Levels for each group differed between the Initial and Final assessments. No evidence was found for a change within the group receiving the Relations-First Training Condition ($t(7) = .424$; $p > .66$). For the group receiving the Abstraction-First Training Condition, a significant increase was found from the Initial Assessment to the Final Assessment given appropriate adjustments for Type I error ($t\ (7) = 3.23$; $p < .05$). The post hoc analysis further explains the results for the main effects. The significant effect of Phase can be attributed to the increase in Hierarchical Levels for the Abstraction-first Training Condition alone. Therefore, subjects in the Abstraction-First Training Condition increased the number of Hierarchical Levels with training while those in the Relations Training Condition did not. All results for Hypothesis One are summarized in Table 3.

Insert table 3 about here.

Hypothesis Two. The number of *Multiple Relations* between concepts developed within an individual's knowledge structure is a function of the training which an individual receives.

The results for Hypothesis Two are presented graphically in figure 5. The main effect of Training Condition was found to be significant at the $p < .05$ level for Hypothesis Two ($F\ (1,14) = 5.04$; $p < .042$). The number of Multiple Relations which the subjects developed was a function of the two Training Conditions (Abstraction-First and Relations-First). The Phase effect was not significant ($F\ (1,14) = 2.17$; $p > .16$); therefore, there was no evidence for an overall increase in Multiple Relations between the Initial Assessment and the Final Assessment. Likewise, the interaction, Training Condition by Phase, was not significant ($F\ (1,14) = 3.03$; $p < .104$).

Insert figure 5 and table 4 about here.

Again, a post hoc analysis was performed to further examine the main effects. The purpose was to determine if the number of Multiple Relations of the Training Condition groups differed at the Initial and Final assessments. For the Initial assessment, there was no evidence for a significant difference between the two groups ($t\ (7) = .23$; $p > .59$). Yet, after subjects had received the complete training program, the groups differed significantly ($t(7) = 2.29$; $p < .05$). Subjects in the Abstraction-First Training Condition formed a greater number of Multiple Relations than the subjects in the Relations-first Training Condition, as evident in figure 5. The post hoc analysis helps explain the Training Condition main effect. All results for Hypothesis Two are summarized in Table 5.

Insert table 5 about here.

DISCUSSION

Evidence was found to support both hypotheses. The independent variable of Training Condition was expected to produce certain significant changes in the development of the subjects' knowledge structures. These changes would be evident through the dependent variables of Hierarchical Levels and Multiple Relations. More specifically, the subjects' number of Hierarchical Levels was expected to increase upon receiving the abstract portion of either Training Condition, due to the emphasis on the hierarchical structure of relationships between the domain concepts. In a similar manner, the relations portion of either Training Condition was expected to effect the number of Multiple Relations developed between domain concepts by the subjects.

Both of the expectations were correct for those subjects in the Abstract-first Training Condition prior to receiving relations training, as evident through changes the number of Hierarchical Levels and the number of Multiple Relations between the Initial and Final Assessment. The expected training effects were not evident in the subjects in the Relations-first Training Condition. At the Final Assessment, neither the number of Hierarchical Levels nor the number of Multiple Relations increased according to the data from Hypotheses One and Two. Thus, the impact of training on knowledge structure development is attributed, in part, to the sequence in which the training content was presented.

The sequence of training shown to have a significant impact, in terms of increased knowledge structure development, stressed the initial presentation of high level abstraction training material, followed by high level relations training material. This training sequence corresponds to the Abstract-First Training Condition. It is suggested that the initial presentation of abstract material and the subsequent development of hierarchical structure between domain concepts provides a framework for the subjects to organize the domain knowledge. The significant increase in Hierarchical Levels supports in existence of a framework of domain concepts. Additional support for this interpretation is provide by the fact that an increase in Multiple Relations was only found in the Abstract-First group who were presented with the hierarchical relationships between the domain concepts prior to high level relations training.

The results of this study also indicate that the sequence in which training content is presented may inhibit the development of an individual's knowledge structure. The subjects who initially received high level relations, followed by high level abstraction (Relations-First Training Condition) showed no development along either knowledge structure dimension of Hierarchical Levels and Multiple Relations. It is suggested that the subjects were unable to maintain multiple links between concepts and that the multiple relations training interfered with the latter formation of a hierarchical structure between domain concepts.

An individual's knowledge structure is, in part, a function of the training received within a specific domain. It is also known that knowledge structure is, in turn, a determinant of human cognitive performance. Differences in knowledge structure have been found to differentiate between groups of various skill levels. Although no direct performance measures were incorporated in this study, it is likely that training directed toward knowledge structure development will have an impact on the performance of cognitive-oriented tasks. This prediction is made with the underlying assumption that the type of knowledge structure development is appropriate for the given task is a function of task attributes.

The results of this study could potentially be implemented through the design of a such a training program. With the purpose of facilitating knowledge structure development, the improvements of the subjects in the Abstract-First Training Condition suggest that training should utilize a top down approach emphasizing the higher level, abstract aspects of the domain first, with progression to the lower level, more detailed relationships between domain concepts. However, it is expected that in order to achieve the desired performance benefits, the knowledge structure types should be matched specifically to the cognitive requirements of the task.

Given the changes in the manufacturing environment, an important goal is to provide the workforce with the appropriate knowledge structure for specific jobs in order to maximize the performance of the each human operator. The training implications of this study apply to new workers, existing workers with completely new jobs, as well as existing workers who have moved from simplified to enriched jobs. For existing workers, the potential exists for their current knowledge structure, resulting from previous training and experience within the domain, to interfere with their ability to achieve the appropriate knowledge structure.

This study supports the existence of two independent constructs within the theoretical model of knowledge structure, Hierarchical Levels and Multiple Relations, and provides empirical evidence for the ability to influence the development of knowledge structure along those dimensions. These findings have implications for the design of jobs, as well as employee selection and training programs. Further research is needed into methodologies for identifying task requirements, and knowledge structure dimensions, and the subsequent mapping of task characteristics to knowledge structure types. This knowledge can be then implemented into job design and selection and training programs to improve the productivity of the workforce.

REFERENCES

Adelson, B. 1981, Problem-solving and the development of abstract categories in programming languages, *Memory & Cognition*, **9**, 422-433.

Adelson, B. 1984, When Novices surpass experts: the difficulty of a task may increase with expertise, *Journal of Experimental Psychology*, **10**(3), 483-495.

Barfield, W. 1986, Expert-novice difference for software: Implications for problem solving and knowledge acquisition, *Behavior and Information Technology*, **5**, 15- 29.

Bhaskar, R., and Simon, H. A., 1977, Problem solving in semantically rich domains, *Cognitive Science*, **1**, 193-215.

Chase, W. G., and Simon, H. A. 1973, Perception in chess, *Cognitive Psychology*, **4**, 55-81.

Chi, M. T. H., Feltovich, P. J., and Glaser, R. 1981, Categorization and representation of physics problems by experts and novices, *Cognitive Science*, **5**, 121-152.

Corlett, E. N. 1973, Human factors in the design of manufacturing systems, *Human Factors*, **15**, 105-110.

De Groot, A. D. 1965, *Thought and Choice in Chess* (Basic Books, New York).

Hardiman, P. T., Dufresne, R., and Mestre, J. P. 1989, The relation between problem categorization and problem solving among experts and novices, *Memory & Cognition*, **17**, 627-638.

Hobus, P. P. M., Schmidt, H. G., Boshuizen, H. P. A., and Patel, V. L. 1987, Contextual factors in the activation of first diagnostic hypotheses: Expert-novice differences, *Medical Education*, **21**, 471-476.

Jeffries, R., Turner, A., Polson, P., and Atwood, M. 1981, The processes involved in designing software, in *Cognitive Skills and Their Acquisition*, ed. J. R. Anderson (Erlbaum, Hillsdale, NJ).

Koubek, R. J. 1991, *Toward a Model of Knowledge Structure and a Comparative Analysis of Knowledge Structure Measurement Technique* (Purdue University, School of Industrial Engineering).

Koubek, R. J., Benysh, D. V., and Calvez, V. 1993, A comparative review of knowledge structure measurement techniques for interface design, *International Journal of Human-Computer Interaction*, **5**, Number 3.

Koubek, R. J., and Mountjoy, D. N. 1991, The impact of knowledge representation on cognitive-oriented task performance, *International Journal of Human-Computer Interaction*, **3**, 31-48.

Koubek, R. J., and Salvendy, G. 1991, Cognitive performance of super-experts on computer program modification tasks, *Ergonomics*, **34**, 1095-1112.

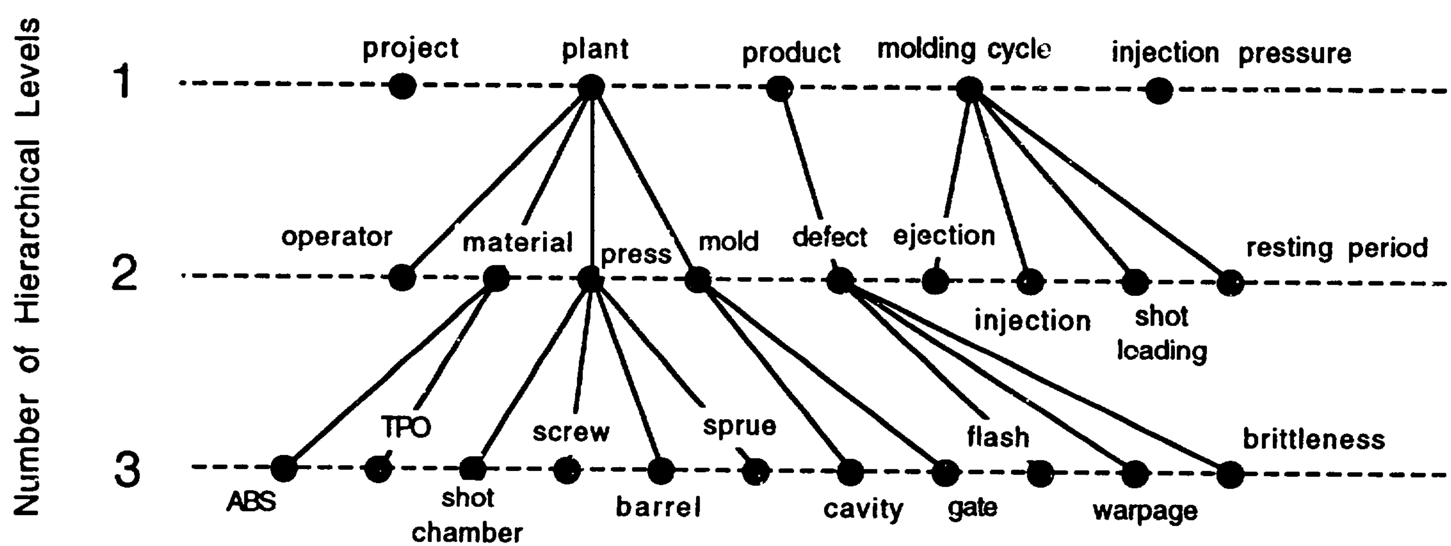
Lewis, C. 1981, Skill in algebra, in *Cognitive Skills and Their Acquisition*, ed. J. R. Anderson (Erlbaum, Hillsdale, NJ).

McKeithen, K. B., Reitman, J. S., Rueter, H. H., and Hirtle, S. C. 1981, Knowledge organization and skill difference in computer programmers, *Cognitive Psychology*, **13**, 307-325.

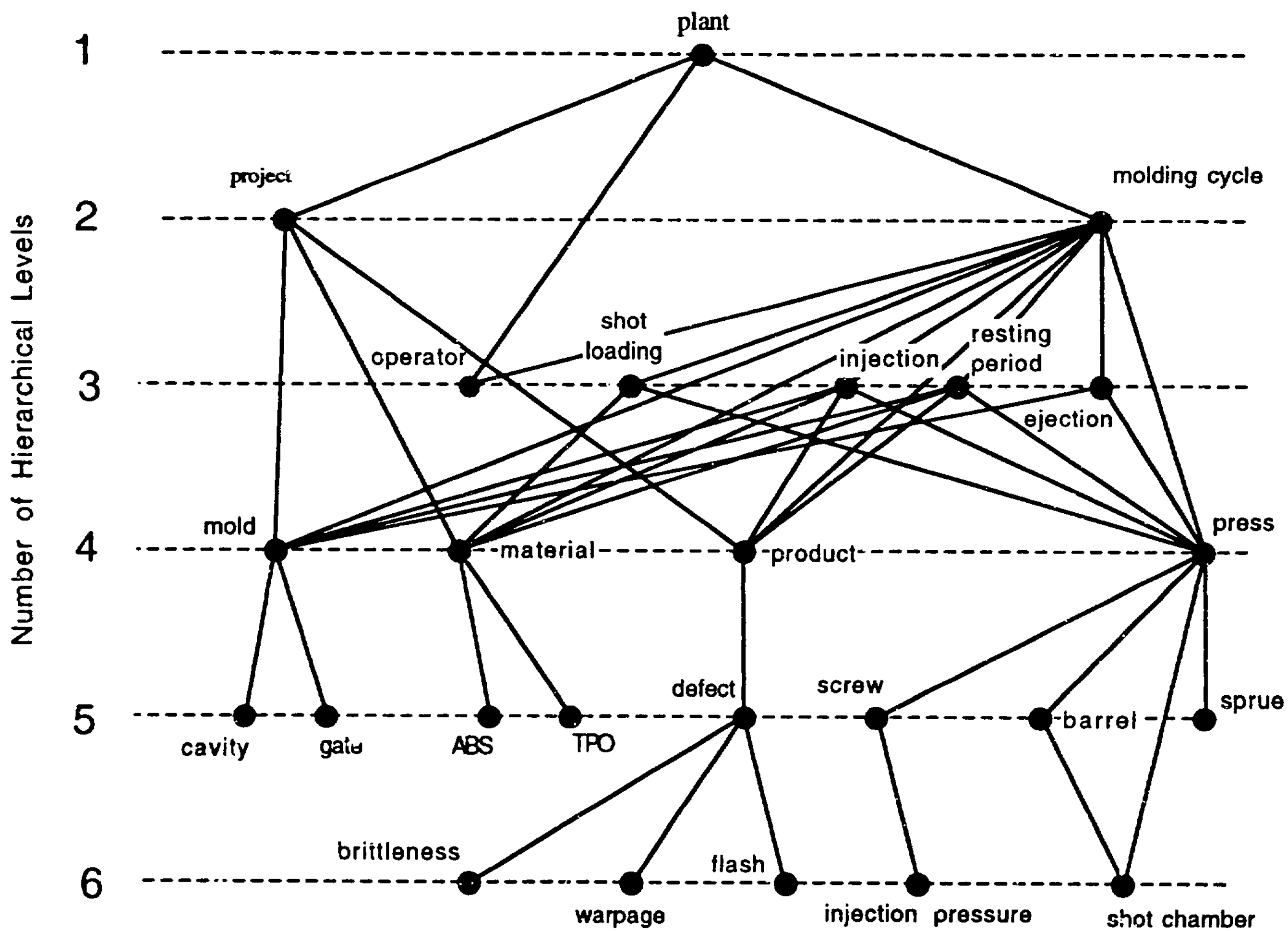
Schneiderman, B. 1977, *Measuring Computer Program Quality and Comprehension*, Tech. Report No. 16 (College Park, University of Maryland, Department of Information Systems Management).

Schoenfeld, A. H., and Herrmann, D. J. 1982, Problem perception and knowledge structure in expert and novice mathematical problem solvers, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **8**, 484-494.

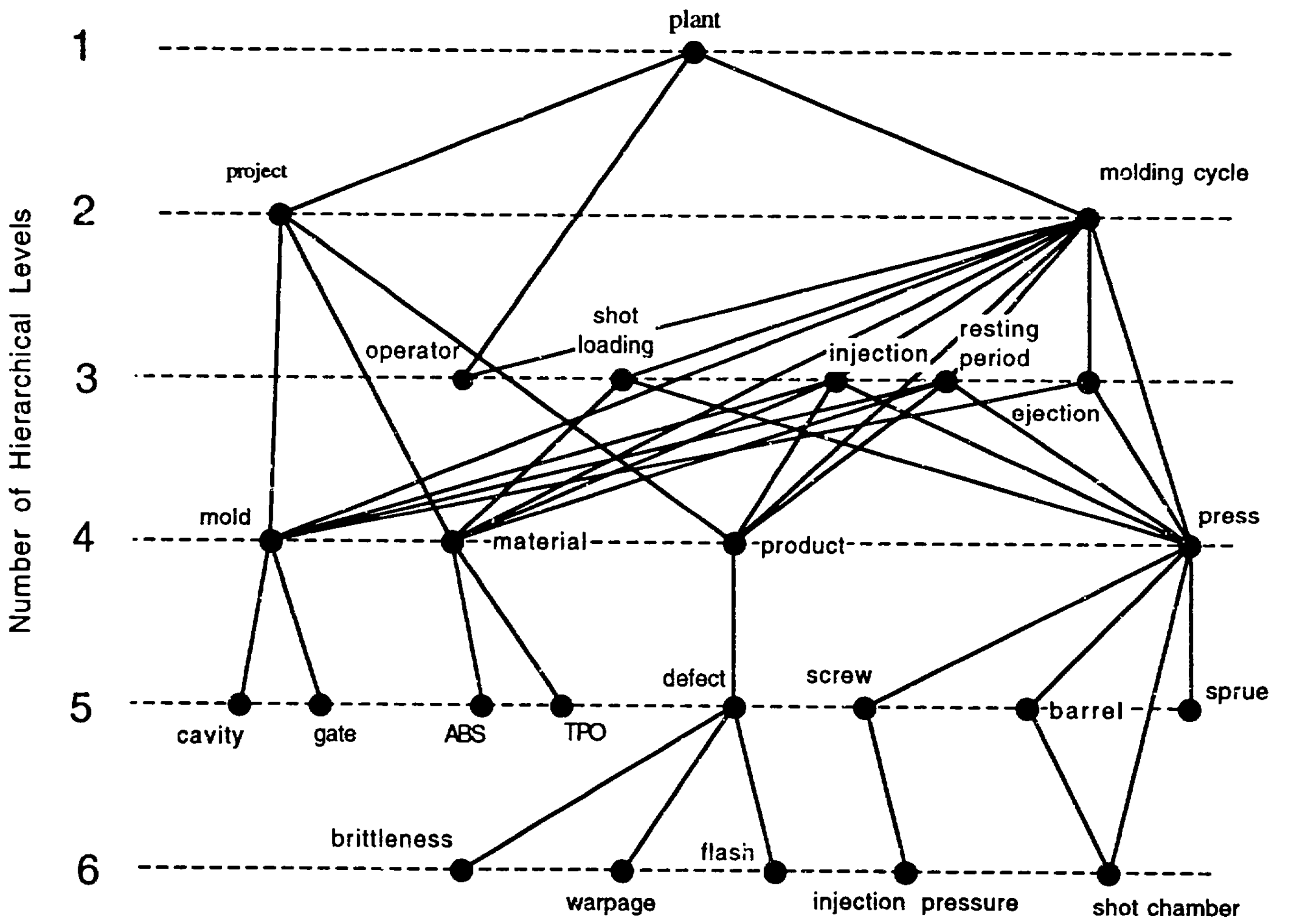
FIGURE CAPTIONS



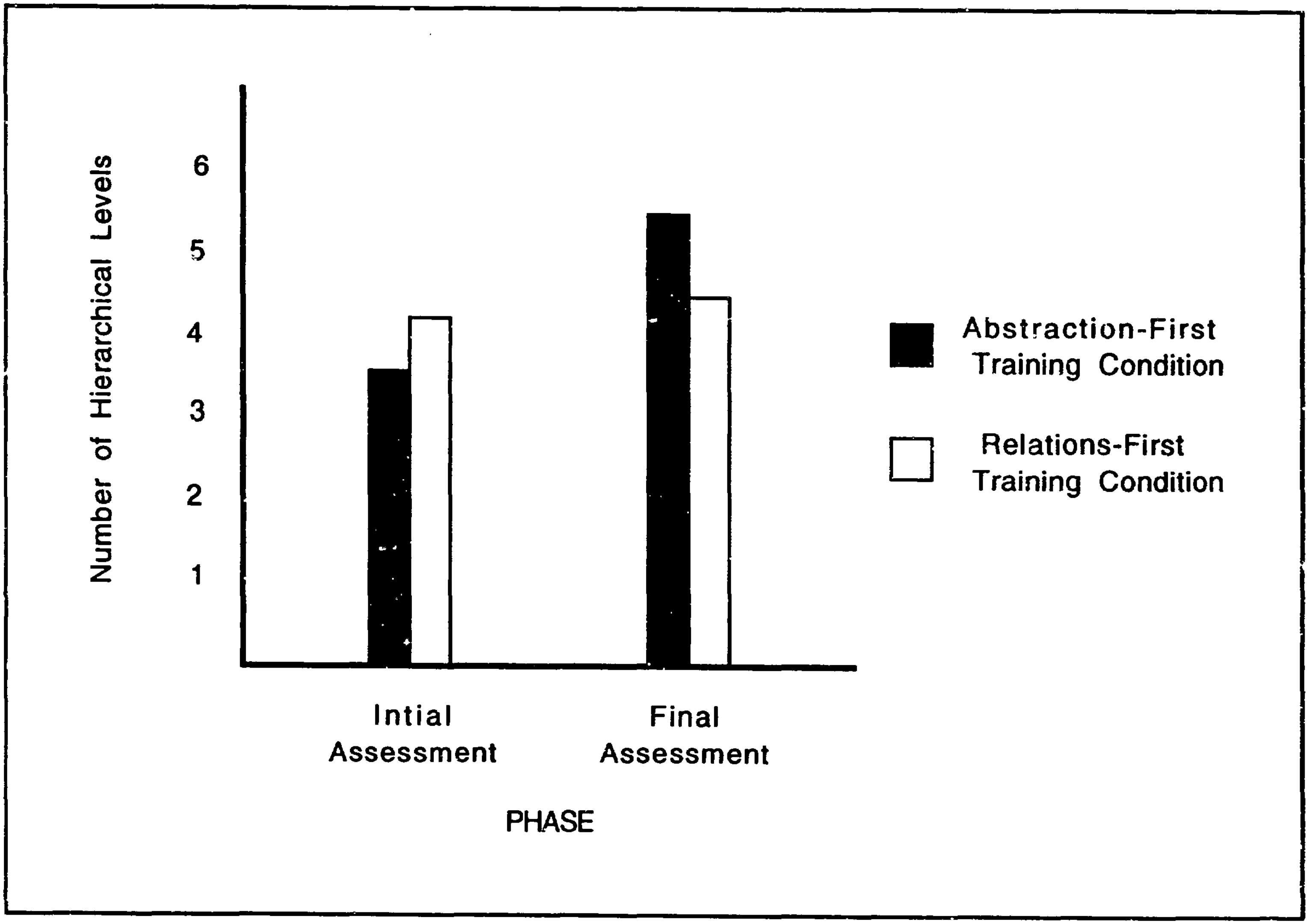
Note: The nodes represent domain concepts and the links represent hierachical relationships.

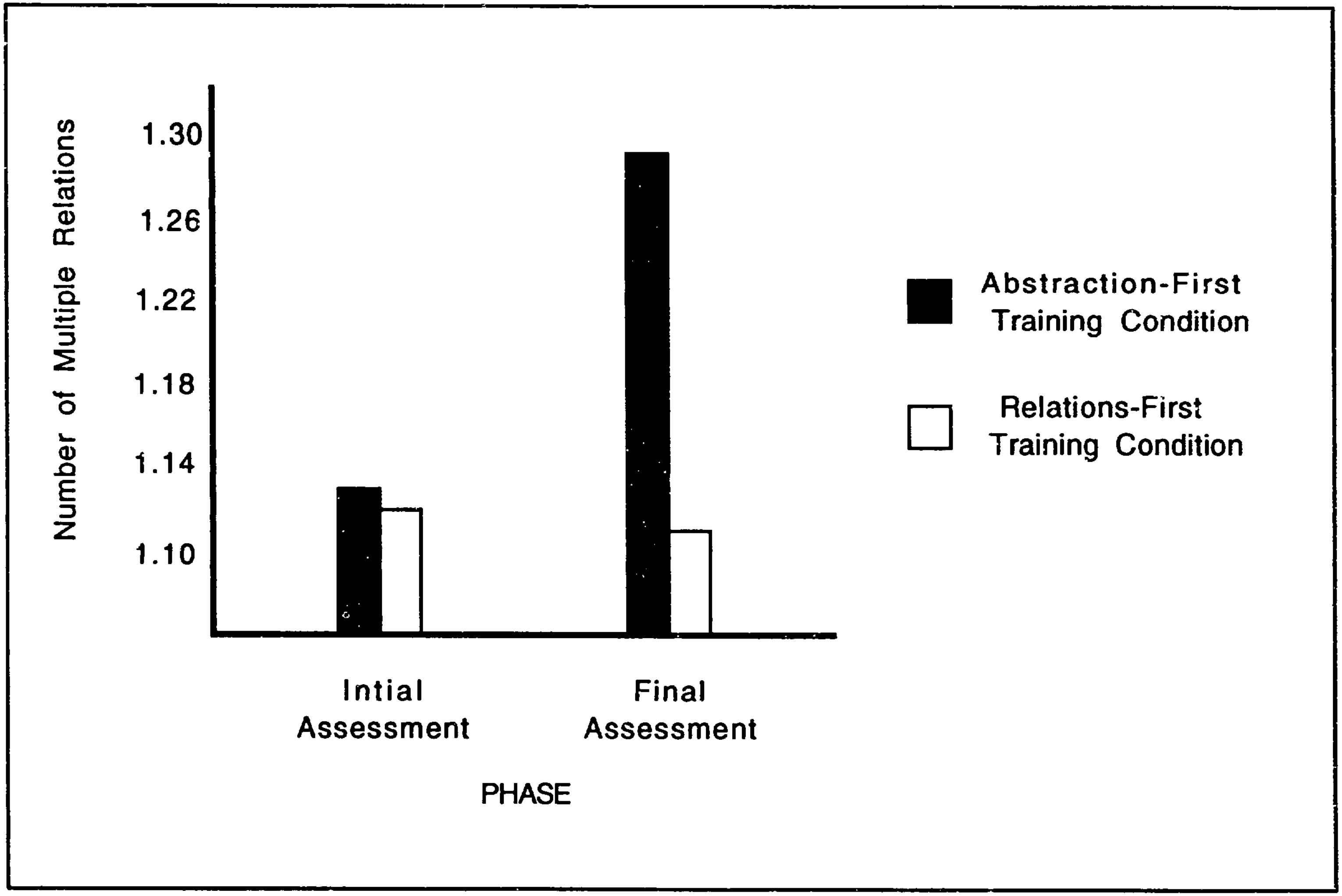


Note: The nodes represent domain concepts and the links represent hierachical relationships.



Note: The nodes represent domain concepts and the links represent hierachical relationships.


Number of Hierarchical Levels
6
5
4
3
2
1
Intial Assessment
Final Assessment
PHASE
Abstraction-First Training Condition
Relations-First Training Condition


Number of Multiple Relations
1.30
1.26
1.22
1.18
1.14
1.10
Intial Assessment
Final Assessment
PHASE
Abstraction-First Training Condition
Relations-First Training Condition

Appendix 1. KSAT Concept List for the Plastic Extrusion Domain.

KSAT Concept List for the Plastic Extrusion Domain	
operator	project
product	flash
screw	shot loading
gate	sprue
press	ABS
barrel	mold
resting period	ejection
plant	material
warpage	molding cycle
TPO	brittleness
shot chamber	injection pressure
defect	cavity
injection	

Appendix 2. KSAT Relationship List.

KSAT Relationship List		
advances during	is opposite of	requires low
advances toward	is same as	retracts during
are both characteristics	is within	retracts from
are both examples	made of	scheduled into
are both parts	matches	sealed off during
causes	melts	selects
closes	melts during	transforms
closes after	melts material during	transports
coincides with	mixes	transports material during
connected to	monitors	transports material to
controls	moves though	freezes before
determines	occurs within	halts
distributes	opens	hardens
distributes material into	opens before	hardens during
effects	operates	hardens material during
ejected from	packs	holds
empties during	prepares	influences
flows through	produces	injected into
follows	provides material for	injects
forces material into	removed from	inspects
forces material through	removes	is an example of
forms	requires high	is a part of
loaded into	is a characteristic of	

Table 1. Description of Independent Variables.

Independent Variables	Levels	Description
Training Condition	Abstraction-First	Training emphasizes hierarchical structure within domain concepts.
	Relations-First	Training emphasizes multiple relations between domain concepts.
Phase	Initial Assessment	Assessment of the subject's knowledge structure following the first training session.
	Final Assessment	Assessment of the subject's knowledge structure following the third training session.

Table 2. ANOVA for Hypothesis One.

SOURCE	dF	SS	MS	F	p
Between Subjects	15	42.4688			
Training Condition	1	0.2813	0.2813	0.09	0.7645
Ss within groups	14	42.1875	3.0134		
Within subjects	16	33.5066			
Phase	1	9.0313	9.0313	6.59	0.0224
Training Condition X Phase	1	5.2813	5.2813	3.85	0.0698
Phase X Ss within groups	14	19.1940	1.371		
Total	31	75.9754			

Table 3. Summary of Results for Hypothesis One.

Hypothesis One	Effect	Test Statistic	Significance Level	Conclusion
Hierarchical Levels develop as a function of training.	Training Condition	$F(1,14) = .09$	$p > .76$	No evidence for a difference between the Abstraction and Relations Training Conditions.
	Phase	$F(1,14) = 6.59$	$p < .023$	Hierarchical Levels increased with training.
	Interaction of Training Condition by Phase	$F(1,14) = 3.85$	$p < .07$	Suggestive evidence for interaction effect.
	Initial and Final assessments for the Abstraction-First Training Condition	$t(7) = 3.23$	$p < .05$	In the Abstraction Training Condition, Hierarchical Levels increased with training.
	Initial and Final assessments for the Relations-First Training Condition	$t(7) = .424$	$p > .66$	In the Relations Training Condition, no evidence for an increase in Hierarchical Levels with training.

Table 4. ANOVA for Hypothesis Two.

SOURCE	dF	SS	MS	F	p
Between Subjects	15	0.3592			
Training Condition	1	0.0950	0.0950	5.04	0.0415
Ss within groups	14	0.2642	0.0189		
Within subjects	16	0.4108			
Phase	1	0.0464	0.0464	2.17	0.1628
Training Condition X Phase	1	0.0648	0.0648	3.03	0.1035
Phase X Ss within groups	14	0.2996	0.0214		
Total	31	0.7700			

Table 5. Summary of Results for Hypothesis Two.

Hypothesis Two	Effect	Test Statistic	Significance Level	Conclusion
Multiple Relations develop as a function of training.	Training Condition	$F(1,14) = 5.04$	$p < .042$	Difference between Training Conditions.
	Phase	$F(1,14) = 2.17$	$p > .16$	Main effect for Phase not supported.
	Interaction of Training Condition by Phase	$F(1,14) = 3.03$	$p < .104$	No significant evidence for interaction effect.
	Abstraction-First and Relations-First Training Conditions at Initial Assessment	$t(14) = .23$	$p > .59$	No evidence for difference in Multiple Relations between groups.
	Abstraction-First and Relations-First Training Conditions at Final Assessment	$t(11)^{*} = 2.29$	$p < .05$	The Abstract Training Condition group had higher Multiple Relations than the Relations Training Condition group.

* Degrees of freedom vary due to adjustments for unequal variances

ATTACHMENT 5

Toward a Model of Knowledge Structure Development With Application to Cognitive task Performance in the Manufacturing Environment

Timothy Clarkston and Richard Koubek

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

Toward a Model of Knowledge Structure Development With Application to Cognitive Task Performance in the Manufacturing Environment

Timothy Paul Clarkston[1], Richard Koubek[2] and Darel Benysh[2]

[1]Harris Corporation
Government Communication Systems Division
P.O. Box 9100
Melbourne, FL 32902-9100

[2]School of Industrial Engineering
1287 Grissom Hall
Purdue University
West Lafayette, IN 47907-1287

This research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-92-J-1153. The opinions expressed in this paper do not necessarily reflect the position of ONR. This work was performed while the first author was at Purdue University.

ABSTRACT

The first objective of this study was to identify factors and their interactions which affect the development of knowledge structure and formulate a predictive model of knowledge structure development. The second objective examined the relationship between knowledge structure dimensions and actual performance in a master production scheduling task. These two objectives were designed to provide information which will facilitate the integration of human capabilities in production systems. Evidence was found to indicate that cognitive style and knowledge structure are independent constructs of human cognitive ability. In addition, the impact of knowledge structure on cognitive task performance was validated..

1. INTRODUCTION

The manufacturing environment is in a state of change. Mass production is giving way to new styles of manufacturing such as lean production, agile manufacturing, and customer-oriented manufacturing. Companies are competing in a global marketplace along the four performance dimensions of quality, time, cost, and flexibility. To remain competitive, manufacturing systems are shifting away from traditional methods of operation, management strategies, and organizational structures.

As a direct result, the jobs are changing within the manufacturing system. The tightly controlled, narrow, highly repetitive, manual tasks associated with mass production are being augmented or replaced by broad-based jobs with cognitive-oriented task components. The current movements toward worker empowerment, job enrichment, and group technology have resulted in a shift toward higher levels of decision making for the human operators at the shop floor level. Workers are sharing the responsibility for plant functions while participating in such activities as production scheduling, maintenance, task assignments, and process control. Although the physical aspects of work at the shop floor level may still exist, the cognitive aspects of the new tasks require the workers to use their knowledge of the manufacturing system in order to solve problems in a dynamic environment, often in situations of uncertainty. Thus, the ability of the human operator to perform cognitive-oriented tasks is becoming of increasing importance to the overall performance of the manufacturing system.

The capability to perform cognitive-oriented tasks can be attributed to many factors, such as knowledge content, innate cognitive skills and abilities, mental strategies, motivation, and degree of practice. In addition to these, research has shown knowledge structure to be a primary determinant of human cognitive task performance (Koubek and Salvendy 1991). Knowledge structure refers to the manner in which an individual organizes knowledge of a specific domain.

In other words, knowledge structure is the representation of a domain based on the units of knowledge within that domain.

1.1. Objectives

To improve the performance of the manufacturing system, the goal should be that each individual within the system possess the appropriate cognitive attributes required to perform their tasks in the most effective manner. Based on the research into cognitive task performance, the appropriate cognitive attributes should include the best suited knowledge structure for each task. The goal of equipping the workforce with the appropriate knowledge structure requires awareness of at least two important aspects. First, the factors (and their interactions) which influence the development of a knowledge structure in a given domain must be identified and related to specific dimensions of knowledge structure. Second, the relationship(s) between the knowledge structure dimensions and certain task characteristics which result in superior task performance must be determined.

The first objective of this study is to identify the factors and interactions affecting the development of knowledge structure along specific dimensions; then, formulate a predictive model of knowledge structure development. Factors being tested are individual differences in cognitive style, various training paradigms, and related experience.

The second objective of this study is to conduct an exploratory experiment to examine the relationships between the knowledge structure dimensions and actual task performance. This experiment is viewed as cursory since it is limited to one type of task; master production scheduling.

The objectives of this study are represented in Figure 1. First, the study will attempt to identify the knowledge structure development factors and formulate a predictive model of knowledge structure development. Second, the study will explore the relationship between knowledge structure and cognitive task performance. The 'Other determinants' in the figure indicate additional factors which affect cognitive task performance, such as degree of practice, motivation, and environmental surroundings.

Please Insert Figure 1 About here

1.2. Background

1.2.1. Description of Knowledge Structure

The following operational definition of knowledge structure will be utilized in this experiment in order to further the systematic study of knowledge structure. Proposed by Koubek (1991), "Human knowledge structure is defined as the structure of interrelationships between elements in a particular domain, organized into a unified body of knowledge." According to this definition, knowledge structure has two primary components, elements and interrelationships, each with its own characteristics.

Elements are unique units of knowledge, or concepts, within a domain. Domain elements can exists in either a declarative or procedural format. Declarative concepts contain knowledge in the form of facts and meanings within the domain. Procedural concepts contain knowledge in the form of how to execute functions associated with the domain. Within a knowledge structure, elements may exist at various levels of hierarchy. For example, two concepts within the domain of discrete parts manufacturing may be "Milling machine" and "Flexible manufacturing cell". The milling machine is considered to be at a lower level than, i.e. subordinate to, the flexible manufacturing cell because it is part of the flexible manufacturing cell.

Interrelationships in the knowledge structure definition indicate that relationships, or links, exist between the domain elements. As previously described, two elements may be at different levels of hierarchy, this is because a super- or sub-ordinate relationship exists between them. Given that a relationship does exist between a pair of elements, the strength of the relationship, or the degree to which the relationship applies, can vary. It is possible that more than one relationship could exist between the same two elements. For example, "Volume" and "Temperature" are two possible domain elements a human operator may associate with the task of controlling a chemical manufacturing process. Assuming that both quantities are presented visually, the operator may relate volume and temperature in terms of their physical proximity on the visual display terminal. Yet, a second and stronger relationship may exist between volume and temperature concerning their interaction in the process, such that changes in volume are proportional to changes in temperature.

The two primary components of knowledge structure and their characteristics are summarized in Table 1.

Please Insert Table 1 About Here

Essential to this study is the ability to quantify specific features of knowledge structure. Three knowledge structure dimensions (Hierarchical Levels, Multiple Relations, and Information Content) are developed to quantify the features previously described in the operational definition of knowledge structure. The dimensions are based on the assumption of a tree-like network representation of knowledge structure formed by the domain elements and their interrelationships. The organization of domain knowledge through the decomposition of the domain concepts according to super- and sub- ordinate relationships is incorporated in the dimension of Hierarchical Levels. Multiple Relations measures the number of relationships existing between two domain elements. Adapted from information theory, Information Content quantifies the amount of information within a knowledge structure to indicate the breadth and depth of the development of an individual's knowledge structure. Although Information Content is a newly proposed knowledge structure dimension, evidence for the existence of Hierarchical Levels and Multiple Relations is provided by Koubek (1991) and Koubek, Clarkston, and Calvez (in press). Further explanation of the knowledge structure dimensions will be presented in the Methodology section.

1.2.2. Expertise as a Function of Knowledge Structure

Previous knowledge structure research has focused on the differences in the manner by which groups of different skill levels represent knowledge of the same domain. The two groups usually correspond to experts and novices and are selected by some criteria which is often external to the experiment, such as years of experience in a field of work. In general, research indicates that the higher skill level group represents the domain in a more abstract or conceptual manner, while the lower skill level group tends to focus on the concrete or elemental aspects of the domain.

A classical example of this type is the study performed by Adelson (1981). Two groups of computer programmers, experts and novices, were presented with randomized computer code and asked to organize it in any manner. Based on a conceptual understanding of the domain, the experts reorganized the computer code into working programs. The novices, with a more surface understanding of programming, organized the computer code into categories according to similarities in key control words. Similar results have been found across a variety of domains: chess (de Groot 1965; Chase and Simon 1973), word processing (Koubek and Mountjoy 1991), problem solving (Hardiman et al. 1989; Chi et al. 1981; Lewis 1981; Bhaskar and Simon 1977; Schoenfeld and Herrmann 1977), software design (Barfield 1986; Jeffries et al. 1981), computer programming (McKeithen et al. 1981; Schneiderman 1977), and medical diagnosis (Hobus et al. 1987).

Another study performed by Adelson (1984) investigated the interaction between domain representations and cognitive task performance. This study confirmed that the expert's representation is more abstract and the novice's representation is more concrete. However, this study demonstrated that the performance of the novice group could exceed that of the expert group in terms of correct responses. The condition in which the novices exceeded the experts is attributed to the knowledge structure of the novices matching the requirements of the task better than that of the experts. This study suggests that the knowledge structure best suited for a task is dependent upon the characteristics of that task.

Review of existing research on the relationship between knowledge structure and cognitive task performance reveals little work exploring the practical link between knowledge structure development and subsequent task performance. The development of knowledge structure is suspected to be a function of training, cognitive style, and experience, among other factors.

1.2.3. Training

The fact that knowledge within a specific domain is partly a function of training within that domain is intuitive. This study focuses on the quantitative development of knowledge structure within a specific domain as a result of training in that domain.

The effect of training on the representation of a task was examined by Koubek and Mountjoy (1991). In that study, two training styles were used to train two groups in a word processing domain. Hierarchical training featured word processing commands ordered by functions, while alphabetical training featured commands listed by alphabetic order. Based on the dependent variables used to classify the domain representation, no significant differences were found in the representations of the two training groups.

A more recent study by Koubek, Clarkston, and Calvez (in press) demonstrated that the development of knowledge structure could be manipulated through training. The experiment utilized the manufacturing domain of plastic extrusion machine operation. Sixteen subjects, having no previous knowledge of the domain, were randomly assigned to one of two distinct training groups. Over a three day period, both training groups received the same instructional content; however, the sequence in which the training material was presented differed. One group initially received the abstract, conceptual relationships between the domain concepts, followed by more detailed relationships associated with the lower level aspects of the domain. The other group received the training material in the reverse order; i.e. the lower level information followed by the abstract information. Each individual's knowledge structure was assessed along two dimensions, Hierarchical Levels and Multiple Relations, which are equivalent to the two of the dependent variables of this study. The group which received the abstract relationships first showed significant improvement along both dimensions of knowledge structure following training. No significant changes in the knowledge structure dimensions were found for the group which received the lower level relationships first. This study suggests that an individual's knowledge structure can be manipulated through training, with a significant effect being attributed to the training sequence of abstract material followed by the more detailed material. Unfortunately, the development of the training documents was ad-hoc, leaving little more than global guidelines for practitioners desiring to apply these results.

One intent of this study is to test the effect of training on knowledge structure development based on systematic application of a fundamental theory of training. Although there is not a global theory of instructional design, Gagne's theory of instructional design is widely accepted. The main feature of Gagne's theory of instruction is the decomposition of the material to be learned into smaller units of knowledge and prerequisite skills, referred to as a hierarchical task analysis. The design of the instruction then begins with the lowest level of the hierarchical task analysis, i.e. the lowest level of the domain, and proceeds until the highest level of the hierarchical task analysis have been presented (Gagne 1987; Gagne 1985; Brower and Hilgard 1981).

This study will test the effect of two distinct training approaches on the development of knowledge structure. The two training approaches are based on different perspectives of the hierarchical task analysis. Initially, a hierarchical task analysis must be performed on the instructional domain. The first training approach, in accordance with Gagne's theory, starts with the lowest level of the hierarchical task analysis, i.e. the lowest level of the domain concepts, and proceeds to the highest level of the domain concepts. This approach is referred to as the Bottom-Up training approach. Based on the research findings that groups of higher skill level tend to have a more conceptual or abstract representation of the domain, the second training approach is aimed at developing that conceptual or abstract knowledge structure. This Top-Down training approach is designed by starting at the highest level of hierarchical task analysis, i.e. the highest level domain concepts, and proceeding to the lowest levels of the domain concepts. The implementation of both approaches in developing corresponding training document will be discussed in the Methodology.

1.2.4. Cognitive Style

Cognitive style is a construct developed to explain individual differences in the human ability to organize and structure information received from the environment or surroundings. Cognitive style has been studied and defined experimentally in a multitude of ways and shown to affect an individual's representation of their surrounding environment (Goldstein and Blackman 1978). The potential effects of cognitive style on the development of knowledge structure appears to be straightforward. However, specific research on the effect of cognitive style on knowledge structure development is inconclusive (Koubek and Mountjoy 1991).

This study will test the effects of two particular cognitive styles, Speed of Closure and Associational Fluency, on the development of knowledge structure. Speed of Closure refers to the human ability to identify a single unified concept in a perceptual field of apparently unrelated stimuli. This cognitive style approach is similar to field dependence, the most frequently tested approach, except for the primary difference that the target to be identified is unknown to the subject; thus, relying solely on the individual's ability to link the stimuli into a coherent object (Erkstrom et. al. 1976). Speed of Closure is suspected to affect knowledge structure development by influencing the ability of the individual to internally organize the domain information, specifically in a hierarchical manner, when presented within the training documents. Associational Fluency, the second approach to be tested, reflects the human ability to produce concepts or words with similar meanings or semantic properties (Erkstrom et. al. 1976). With respect to knowledge structure development, Associational Fluency is expected to influence the ability to form more than one relationship between two domain concepts and capture the role of memory in knowledge structure development.

The two cognitive styles approaches, Speed of Closure and Associational Fluency, will be assessed quantitatively through the appropriate tests from the Kit of Factor Referenced Cognitive Tests (Erkstrom et. al. 1976). The actual testing procedures will be discussed in the

Methodology section. For the purpose of classifying subjects into experimental groups, criteria for a 'High' and 'Low' level of each cognitive style approach will be based on large sample distributions for each, excluding twenty five percent of the population about the sample mean.

1.2.5. Experience

It is suspected that knowledge structure development is effected by previous experience. Previous experience includes knowledge of the same domain, knowledge of the same domain with some modified tasks, and knowledge of a similar domain with some common concepts and procedures. The suspected effects of experience on knowledge structure development is unclear. The effect may be positive, i.e. facilitate development of a suitable knowledge structure, or adverse, i.e. interfere with development of a suitable knowledge structure.

To implement previous experience in this study, an experience variable would require a minimum of two groups of varying experience levels plus pre- and post- assessment of each subject's knowledge structure. Experience is not being tested in this study due to resource constraints.

1.3. Hypotheses

It is proposed here that knowledge structure development is partially a function of training and cognitive style. The impact of these factors on knowledge structure development can be assessed through the knowledge structure dimensions of Hierarchical Levels, Multiple Relations, and Information Content. Hypotheses One, Two, and Three below reflect the first objective; to determine a predictive model of knowledge structure development. Hypothesis Four corresponds to the second objective of this study; to determine the relationship between knowledge structure dimensions and task performance.

1.3.1. Hypothesis One

The development of an individual's knowledge structure along the dimensions of Hierarchical Levels, Multiple Relations, and/or Information Content is a function of the training which the individual receives.

1.3.2. Hypothesis Two

The development of an individual's knowledge structure along the dimensions of Hierarchical Levels, Multiple Relations, and/or Information Content is a function of individual differences in the cognitive styles of Speed of Closure and Associational Fluency.

1.3.3. Hypothesis Three

The development of an individual's knowledge structure along the dimensions of Hierarchical Levels, Multiple Relations, and/or Information Content is a function of the interaction between the training which the individual receives and individual differences in the cognitive styles of Speed of Closure and Associational Fluency.

1.3.4. Hypothesis Four

An individual's cognitive task performance on the master production scheduling task, as assessed through time and errors, is a function of the individual's knowledge structure within the task domain according to the dimensions of Hierarchical Levels, Multiple Relations, and/or Information Content.

The development factors being considered are training and individual differences in cognitive style. The knowledge structure dimensions being assessed through Hierarchical Levels, Multiple Relations, and Information Content. Cognitive task performance is measured through Completion Time and Error.

2. METHODOLOGY

2.1. Experimental Domain

The experiment utilized the domain of master production scheduling to test the hypotheses involving training, task performance, and knowledge structure assessment. Master production scheduling is usually associated with the initial step of material requirements planning (MRP). Given a demand for a final product, master production scheduling involves the calculation of a planned order for the final product and all of the parts required to manufacture that final product. The planned order for a part includes both the quantity to produce and the time to start production. The master production scheduling process involves the calculation of a series of standardized records with reference to a product database.

Master production scheduling was chosen for the experimental domain to be representative of a general type of cognitive-oriented task associated with manufacturing. It was assumed to have the complexity required to develop multiple training approaches, as well as provide sufficient domain content to assess knowledge structure development.

2.2. Subjects

As a control measure, subjects were limited to pure novices; i.e. subjects who had no previous knowledge of the domain. This constraint was due to the results of pilot studies which showed the effect of related experience. Determination of a subject's experience status was achieved via a questionnaire administered prior to scheduling an appointment. Twenty undergraduate students participated in the study, all being classified as pure novices. Subject demographics were balanced. Subjects were paid five dollars per hour for participating in the experiment. In order to participate, subjects signed an informed consent prior to the cognitive style assessment and the actual experiment.

2.3. Experimental Design

The experiment was originally to be performed as a 2^3 factorial design with the three independent variables corresponding to Training Condition, Speed of Closure, and Associational Fluency. The design was modified due to the inability to find any subjects within the low Associational Fluency experimental condition (based on national norms). Given the nature of the Associational Fluency test, it is suspected that this occurred because the subject pool had a distribution with a mean higher than that of the test criteria distribution. The Associational Fluency data was collected to be used in post hoc analysis.

The experiment was performed as a 2X2 completely crossed factorial with the two independent variables of Training Condition and Speed of Closure. Training Condition had two levels, Top-Down and Bottom-Up. Each of the levels of Training Condition corresponded to a distinct training approach. Subjects were randomly assigned to the training groups. Speed of

Closure also had two levels, Low and High. Subjects were classified as Low Speed of Closure or High Speed of Closure as a result of the cognitive style assessment.

Data was collected for a total of 5 dependent variables, with the experimental design being repeated five times; once for each dependent variable. Three of the dependent variables correspond to the knowledge structure dimensions of Hierarchical Levels, Multiple Relations and Information Content. The other two dependent variables measure task performance in terms of Completion Time and Error.

2.4. Experimental Procedure

Each subject completed the experimental procedure in one day requiring approximately three hours. Initially, a cognitive style assessment was completed by each subject. Whether or not the subject continued in the experiment was dependent upon the subject's scores on the cognitive style assessment. If the subject was in an experiment group, the subject would then be randomly assigned to one of the two Training Conditions. During the training phase of the experiment, the subject studied the appropriate training document for an unlimited amount of time; however, before proceeding to the next phase of the experiment, each subject was required to pass a proficiency exam. If a subject's answers were not satisfactory, the subject was asked to re-read the training document and a different proficiency exam was given, for a maximum of four times. In next phase of the experiment, the subject was asked to perform a master production scheduling task. During the final phase of the experiment, data was elicited about the subject's knowledge structure within the master production schedule domain via a computer-based questionnaire. This knowledge elicitation phase consisted of two parts. One questionnaire emphasized declarative knowledge, while the other emphasized procedural knowledge. The focus of this study is limited to declarative knowledge structure; thus, this study only utilizes the data from the declarative knowledge elicitation. The order in which the two elicitation questionnaires were administered was randomized.

2.5. Cognitive Style Assessment

The two cognitive styles being examined is this study are Speed of Closure and Associational Fluency. The ETS Kit of Factor Reference Cognitive Tests provides efficient and reliable methods of assessing both cognitive styles. The Gestalt Completion test was used to identify individual differences in Speed of Closure. The test presents a subject with twenty incomplete stimuli (pictures of objects) which must be identified within a four minute time limit. A subject's score on this test is equal to the total number of correct identifications. The selection procedure is designed to eliminate twenty-five percent of the population about the mean in order to produce two distinct groups. Therefore, the selection criteria for 'High' Speed of Closure was a score greater than or equal to eighteen and for 'Low' Speed of Closure, the score had to be less or equal to fourteen.

The cognitive style of Associational Fluency as measured through the Controlled Associations Test. A subject is presented with eight target words and asked to generate as many similar or related words in a twelve minute time limit. The score on this test is the sum of all correct responses. Although not used as a subject selection criteria, the original criteria corresponding to twenty-five percent around the mean was a score greater than or equal to 28 to be considered "High' and a score less than or equal to 20 to be considered 'Low'.

2.6. Training Document Development

As previously discussed in the introduction, the two training approaches begin tested correspond to a top-down and bottom-up adaptation of Gagne's Hierarchical Task Analysis. Thus, the starting point of both training documents was a hierarchical task analysis performed on the master production scheduling task. The purpose is to decompose the master production scheduling domain into the elemental concepts of knowledge required to perform the task. Below the lowest level of the hierarchical task analysis presented are prerequisite skills required to perform the task, such as basic mathematics and reading. Given the subject pool of undergraduate students, these skills are assumed to be at a sufficient level.

At this point, each document was developed as described below. It should be noted that both training documents were revised to include missing information required to perform the task and clarify misconceptions about the domain concepts and procedures as revealed through the evaluation of the task performance of the pilot subjects, as well as interviews conducted with the pilot subjects.

2.6.1. Top-Down Training Document

Development of the Top-Down training document started with the highest level of the hierarchical task analysis, the concept of master production schedule, and proceeded to present the concepts of planned order and product structure, until the lowest levels of the domain, such as lead time and batch size, were presented. The Top-Down training document was written from a conceptual perspective relating the domain concepts to the goal of the master production scheduling task. Following the conceptual explanation, the document features a complete example of the master production scheduling task emphasizing the lowest level aspects of the domain and completing the record. The Top-Down training document was referred to as Training Document A within the confines of the experimental sessions.

2.6.2. Bottom-Up Training Document

As the title implies, the Bottom-Up training document is essentially the opposite of the Top-Down training document. The Bottom-Up training document began with the quality and time aspect of the planned order and record completion, and then proceeded to present the higher level domain concepts. To facilitate this approach, an example was developed stepwise throughout the Bottom-Up training document, relating the domain concepts to the example as necessary to perform the task. For consistency, the same example was used in both training documents, only the method of presentation differed. For experimental purposes, the Bottom-Up training document was label Training Document Z. Figure 2 illustrates both training approaches with respect to the hierarchical task analysis of the master production scheduling domain.

Please Insert Figure 2 About here

2.7. Proficiency Examination

A proficiency examination was conducted following the training session and prior to performing the master production scheduling task. Each subject had to pass the same exam in order to proceed to the task portion of the experiment.

Initially, the proficiency examination consisted of a list of questions which tested the subjects semantic knowledge of selected domain concepts. For example, a question was "What is batch size?". However, the results of the pilot study indicated that it was possible for a subject to sufficiently answer the proficiency questions, but still be unable to complete the master productions scheduling task.

As a result, the proficiency examination was changed to more directly reflect the knowledge required to perform the task. The subject was asked to complete an individual master production schedule record given the required information. The record was assessed by the experimenter. Each subject was expected to complete the record correctly and unassisted with relative ease.

2.8. Task

The subjects were presented with a manufacturing scenario in a text format, then asked to perform operations associated with the master production scheduling task. First, the subjects were asked to interpret the product structure as presented in the text and draw a graphic representation as illustrated in the training documents. Next, the subjects filled in missing values in the product database with information contained in the text. Finally, the most significant portion of the task, the subject had to formulate a complete master production schedule to meet the demand for the final product. This involved the calculation of nine master production schedule records. The subjects were instructed to perform the task as quickly as possible but without sacrificing the accuracy of the responses.

2.9. Dependent Variables for Task Performance

The performance of each subject was assessed through two dependent variables, Completion Time and Error. Completion Time was simply the time elapsed from when the subject was presented the entire task by the experimenter until the subject indicated that he/she was finished. The time required for the experimenter to provide instructions was excluded from the Completion Time.

The Error dependent variable is a summation of inaccurate responses over the complete master production scheduling task for each subject. The possible errors include both errors of omission and errors of commission. An error of omission is not performing the correct action, such as not calculating the current inventory for a given week. Errors of commission involve performing the correct action, but in some inadequate manner, such as calculating the current inventory for a given week without including the scheduled receipts in the calculation. Extraneous errors, such as pure mathematical mistakes, were not included in the Error dependent variable. Similarly, the replication of an error thorough the remaining portion of the task was counted only as a single error. All tasks were graded by the same experimenter for consistency.

2.10. Knowledge Structure Elicitation and Analysis Method (KSEAM)

Knowledge Structure Elicitation and Analysis Method (KSEAM) is a computer-based technique to assess an individual's knowledge structure within a specified domain. The assessment process is comprised of two steps, knowledge structure elicitation followed by knowledge structure analysis.

2.10.1. Knowledge Structure Elicitation

Information about an individual's knowledge structure within a domain is elicited through the pairwise comparison of domain concepts. Two lists are essential to the elicitation process. The concept list as present in the appendix (Table A1) contains fifteen units of knowledge, or concepts, from the master production scheduling domain as determined by the hierarchical task analysis. The relationship list contains twenty-two possible relationships that could exist between domain concepts (see Table A2).

Subjects were presented with pairs of domain concepts in the following format:

Final Product --> Component

The subject must decide if a relationship exists between the two concepts. If the subject does not think that the concepts are related, the subject proceeds to the next pair of concepts. If the subject decides that a relationship does exist between the two concepts, the subject can then classify the relationship by entering the numeric code of the relationship from the relationship list which applies to the pair. At this point, the subject can either enter another relationship which applies to the same pair or proceed to the next pair of concepts.

The pairs of domain concepts were presented to the subjects in random order. Two domain concepts could be presented in two possible sequences. For example,

Backward Scheduling --> Lead Time

OR

Lead Time --> Backward Scheduling

However, each concept pair was only displayed in one sequence during the knowledge structure elicitation process.

During the elicitation process, ten pairs of concepts were repeated in exactly the same manner to assess the reliability of the elicited knowledge. Reliability was calculated by comparing the difference between the subject's initial response to the pair and the response to the repeated pair. The concept pairs repeated were randomly selected from all possible concept pairs.

2.10.2. Dependent Variables for Knowledge Structure Assessment

Based on the data elicited from the each subject, KSEAM produces three measures of an individual's knowledge structure corresponding to the three dependent variables of this study, Multiple Relations, Information Content, and Hierarchical Levels.

Multiple relations is equal to the average number of relationships that exist between a pair of items. Based on the subject's response to the pairwise comparison of all the domain concepts, Multiple Relations is calculated by dividing the total number of relationships that the subject entered by the number of pairs for which at least one relationship was indicated to exist. The equation is shown below.

$$\text{Multiple Relations} = \frac{\text{Total \# of pairwise relationships for all pairs}}{\text{Total \# of pairs with at least one relationship}}$$

The dependent variable of Information Content is derived from Information Theory. The underlying assumption is the equating of a tree-like network of an individual's knowledge to a tree-like decision network. Each concept represents a decision node and the relationships between nodes represent links to decision alternatives. The amount of information (H) for a given decision node is calculated by

$$H = \log_2 N$$

where N is equal to the number of decision alternatives available at the node. Adapted to knowledge structure, the data elicited through the pairwise comparison of domain concepts is converted by KSEAM to a matrix format, such that the concepts are listed on both the rows and columns and any relationship(s) that exist between a pair of concepts are listed in each cell. The matrix is equivalent to a 'From/To' chart, such that the rows are 'From' concepts which links are leaving and the columns are 'To' concepts which links are entering. In creating the matrix, KSEAM utilized only super- and sub-ordinate relationships corresponding to relationships one through fourteen on the relationship list. N is equal to the total number of links exiting from a concept; i.e. the summation of a row. Given N for a row, the calculation of $\log_2 N$ gives the information (H) for that row. The summation of the information (H) for the rows is equal to the Information Content for the entire network. According to the definition of information as the reduction of uncertainty, the higher the calculated value, the greater the amount of information contained in the individual's structure of concepts, as opposed to the content of each concept.

Hierarchical Levels quantifies the hierarchical manner by which an individual organizes the domain knowledge. Hierarchical Levels is equal to the number of levels, or layers, within an individual's knowledge structure resulting from the existence of super- and sub- ordinate relationships between domain concepts. For example, a possible knowledge structure from the master production scheduling domain is shown in Figure 3. The number of Hierarchical Levels equal to six.

Hierarchical Levels is determined by constructing a tree-like representation of an individual's knowledge structure from the pairwise comparison data. According to the super- and subordinate relationships entered, KSEAM sorts the data for each concept to determine the concepts that should be located above and below it in the hierarchy of the individual's knowledge structure. The experimenter then maps the sorted data from KSEAM into the actual tree-like representation and counts the maximum number of levels.

Please Insert Figure 3 About here

2.10.3. KSEAM Design and Development

KSEAM was developed with the following design features. The method was developed in accordance with the proposed operational definition of knowledge structure and design to leverage off the strengths and weakness of the existing knowledge structure assessment techniques. A complete review of existing techniques is provided by Koubek, Benysh, and Calvez (1993). KSEAM was intended to produce non-referent measures of an individual's representation of a domain; not the difference between the individual's representation and some ideal or assumed correct representation. The elicitation process can accommodate both declarative and procedures concepts and/or relationships. The analysis process is based on the ability to classify relationships between domain concepts through a standardized list of possible relationships.

3. RESULTS

The distributions for all five of the dependent variables, both the knowledge structure dimensions and the performance measures, were normally distributed, with homogeneity of variance between the experimental conditions. The average reliability for the information elicited from all the subjects was .925 with a standard deviation of .030.

3.1. Hypotheses One, Two, and Three

Hypotheses One, Two, and Three proposed that knowledge structure development is partially a function of training and cognitive style. The impact of these factors on knowledge structure development was assessed through the knowledge structure dimensions of Hierarchical Levels, Multiple Relations, and Information Content. Hypotheses One, Two, and Three were each tested with a two-way ANOVA. There is no evidence for a significant effect, at the $p < .05$ level, on the development of any of the knowledge structure dimensions due to the main effects of Training Condition and Speed of Closure or the interaction between Training Condition and Speed of Closure. Based on the results of the ANOVA(s), the regression analysis of each knowledge structure dependent variable was unlikely to yield any significant relationships with the model of Training Condition, Speed of Closure, and their interaction. The R-squared values for each of the dependent variables is as follows: .082 for Hierarchical Levels, .002 for Multiple Relations, .037 for Information Content. The ANOVA(s) and the regression analyses had low power. The power for each of the dependent variables is as follows: .13 for Hierarchical Levels, .05 for Multiple Relations, .08 for Information Content. The low power is attributed to a large within subjects' variance relative to the total number of observations. The total number of observations required to achieved an average power of .50, with alpha equal to .05, for the three knowledge structure dependent variables is approximately 350. At this point, in accordance with the original approach to the experimental design, it was not feasible to increase the total number of observations within the constraints of this study.

A regression analysis was performed on the knowledge structure dimensions of Hierarchical Levels, Multiple Relations, and Information Content to determine if a causal relationship existed with the cognitive styles of Speed of Closure and Associational Fluency. With an R-squared value of .07 for Information Content, .0004 for Multiple Relations, and .06 for Hierarchical Levels, it is concluded that no such causal relationship exists. In addition, a correlation analysis was performed including all the knowledge structure dimensions and both of the cognitive styles to determine if any relationship existed between the two constructs. No significant correlation was found at the $p < .05$ level.

3.2. Hypothesis Four

The purpose of Hypothesis Four was to examine the relationship between the dimensions of knowledge structure and actual task performance. This analysis of the data yielded a significant effect on task performance due to the interaction of Training Condition and the cognitive style, Speed of Closure. At $p < .05$ level, the interaction was significant in terms of both Completion Time ($F(1,16) = 4.5881$; $p < .0479$) and Error ($F(1,16) = 4.6621$; $p < .0464$). The ANOVA tables for Completion Time and Error are presented in Table 2 and Table 3, respectively. The effect of the significant interaction between Training Condition and Speed of Closure is evident graphically on Completion Time (Figure 4) and Error (Figure 5).

As another analysis for testing Hypothesis Four, regression analysis was performed on each of the task performance dependent variables (Completion Time and Error) separately with respect to the knowledge structure dimensions in order to determine the percentage of variance in performance that could be accounted for by the knowledge structure dimensions (the R-squared value). The three knowledge structure dimensions of Hierarchical Levels, Multiple Relations, and Information Content accounted for approximately twenty-two percent of the variance in the Completion Time and approximately fifteen percent of the variance in the Error.

Please Insert Tables 2 and 3 About Here

Please Insert Figures 4 and 5 About Here

The regression analysis was expanded to include the cognitive styles, Speed of Closure and Associational Fluency. The R-squared value was determined for the cognitive styles alone and in conjunction with knowledge structure dimensions. All of the regression analyses are summarized in Table 4.

Please Insert Table 4 About Here

Although the p-values for the regression analyses were not significant at the alpha equal .05 level, it is apparent from Table 4 that the percentage of variance accounted for increased in an approximately additive manner given the combination of the knowledge structure dimensions and the cognitive styles.

4. DISCUSSION

The first objective of this study was to identify the factors and their interactions affecting knowledge structure development, then formulate a predictive model of knowledge structure development. Based on the regression analysis of the data from Hypotheses One, Two, and Three, this study was unable to formulate a model of knowledge structure development. However, insight was gained into the possible factors responsible for knowledge structure development.

Although a significant effect was not found for Training Condition, the manner in which instruction is presented can not be discounted due to other studies as a factor contributing to knowledge structure development. The attempt of this study was to test the effect of training on knowledge structure development according to a theoretical base, utilizing Gagne's theory of instructional design. It is possible that the two training approaches used in this study were not distinct enough given the domain to effect the development of the knowledge structure dimensions; i.e. an effect may have been evident if the domain was more complex or contained more concepts for knowledge structure elicitation.

This study, as well as the studies by Koubek and Mountjoy (1991) and Koubek, Clarkston, and Calvez (in press), varied training by presenting the instructional material from both a conceptual and elemental perspective. The study by Koubek, Clarkston, and Calvez (in press) found a significant effect for training on knowledge structure development, while the other two studies did not. The primary difference between the study which found significance and the two studies which did not was the sequence in which the knowledge content was presented to the subjects. The study by Koubek, Clarkston, and Calvez (in press) attributes the significant knowledge structure development to the initial presentation of abstract information followed by more detailed, elemental information, suggesting that the conceptual instruction provides the individual with a conceptual framework to manage the domain knowledge. The sequence of training should be further investigated as a possible factor influencing the development of knowledge structure.

The analysis of the data from Hypothesis One, Two, and Three provides evidence that knowledge structure and cognitive style are independent constructs of the human ability to perform cognitive tasks. Their independence would explain the inability to find a causal relationship or any other type of relationship between the cognitive style scores and the knowledge structure dimensions. Hypothesis Four also supports the independence of knowledge structure and cognitive style and even suggests a complementary or additive effect in accounting for the variance in cognitive task performance. This study, in conjunction with other studies such as Koubek and Mountjoy (1991) and Koubek and Salvendy (1989) would serve to eliminate cognitive style as a possible factor contributing to knowledge structure development. Thus, knowledge structure and cognitive style can be treated as independent constructs in both theoretical research and practical application.

Although previous experience was not a dependent variable in this study, the potential effects of previous effects on knowledge structure development were evident from cursory observations of the pilot studies. Future research should consider previous experience as a possible factor influencing knowledge structure development.

The second objective of this study was to examine the relationship between knowledge structure and cognitive task performance. The analysis of Hypothesis Four validates the influence of knowledge structure on cognitive task performance by showing the variance in task performance that could be accounted for by the three dimensions of knowledge structure.

The significant interaction of Training Condition and Speed of Closure provides two conditions in which complex task performance, in terms of Completion Time and Error, can be maximized regardless of individual differences in cognitive style. An equivalent performance level was achieved by the 'Low' Speed of Closure group who received the Bottom-up training document as the 'High' Speed of Closure group who received the Top-Down training document. These performance results are attributed to the match between the various training approaches and the differences in cognitive style. The elemental, building approach of the Bottom-Up training document, as opposed to the Top-Down training document, is suspected to support the learning of the 'Low' Speed of Closure group who, by definition, have less ability to extract and organize the domain knowledge from the training documents. The conceptual, abstract approach of the Top-Down training document appears to compliment the ability of the 'High' Speed of Closure group to extract and organize the domain knowledge from the training documents. This conclusion is consistent with the findings of Koubek and Mountjoy (1991). In that study, a field dependent group showed superior performance on complex tasks having received a structured training approach, while the field independent group showed equal performance on the structured and unstructured training conditions.

The analysis of the collected data indicates a large within subjects variance across the knowledge structure dimensions, as well as the task performance measures. Given the long and tedious nature of the experimental procedure, the large variance may be attributed to subject fatigue and lack of subject motivation during the knowledge structure elicitation stage. The large variance is evident in the low power and perhaps the inability to find significant results. This points to the importance of finding a new technique other than paired comparisons for eliciting knowledge structures.

The original figure conceptualizing the objectives of this study has been modified to reflect the conclusions of this study. At this point, the factors influencing knowledge structure development is still unknown. Knowledge structure and cognitive style are now determined to be independent determinants of cognitive task performance. Although a clear match between the knowledge structure dimensions and superior cognitive task performance was not achieved, the impact of knowledge structure on cognitive task performance was validated.

This study provides evidence that knowledge structure is an independent and primary factor in human cognitive task performance. The theoretical implication for cognitive engineering, cognitive modeling, and learning theory is that knowledge structure must be recognized and considered as such a factor in research into and modeling of human cognitive task performance and skill acquisition. The need still exists to identify the factors which affect knowledge structure development and the relationship to task performance for a range of task characteristics.

A model of this type could then be implemented into employee selection and training programs tailored for the development of the most suitable knowledge structure, as well as incorporated into the design of tasks and human-machine interfaces. Two practical requirements exists for such an implementation. Compared to current techniques, efficient methods of knowledge structure assessment must be developed. The knowledge structure types, as characterized by the knowledge structure dimensions, must be accurately matched with a spectrum of task characteristics. At this point, the goal of equipping the workforce with the appropriate knowledge structure to maximize the overall performance of the manufacturing system would be feasible.

LIST OF REFERENCES

Adelson. B. 1981, Problem-solving and the development of abstract categories in programming languages, *Memory & Cognition*, **9**, 422-433.

Adelson, B. 1984, When Novices surpass experts: the difficulty of a task may increase with expertise, *Journal of Experimental Psychology*, **10(3)**, 483-495.

Barfield, W. 1986, Expert-novice difference for software: Implications for problem solving and knowledge acquisition, *Behavior and Information Technology*, **5**, 15-29.

Bhaskar, R., and Simon, H. A., 1977, Problem solving in semantically rich domains, *Cognitive Science*, **1**, 193-215.

Brower, G. H., and Hilgard, E. R., 1981, *Theories of Learning* (Prentice-Hall, Inc, Englewood Cliffs, NJ).

Chase, W. G., and Simon, H. A. 1973, Perception in chess, *Cognitive Psychology*, **4**, 55-81.

Chi, M. T. H., Feltovich, P. J., and Glaser, R. 1981, Categorization and representation of physics problems by experts and novices, *Cognitive Science*, **5**, 121-152.

De Groot, A. D. 1965, *Thought and Choice in Chess* (Basic Books, New York).

Ekstrom, R.B., French, J.W., Harman, H.H., and Derman, D. 1976, *Manual for Kit of Factor-Referenced Cognitive Tests.* (Educational Testing Service, Princeton NJ).

Gagne, R. B., 1985, *The Conditions of Learning and Theory of Learning* (CBS College Publishing, New York, NY).

Gagne, R. B., 1987, *Instructional Technology: Foundations* (Erlbaum, Hillsdale, NJ).

Goldstein, K. M., and Blackman, S., 1978, *Cognitive Style* (John Wiley & Sons, New York, NY).

Hardiman, P. T., Dufresne, R., and Mestre, J. P. 1989, The relation between problem categorization and problem solving among experts and novices, *Memory & Cognition*, **17**, 627-638.

Hobus, P. P. M., Schmidt, H. G., Boshuizen, H. P. A., and Patel, V. L. 1987, Contextual factors in the activation of first diagnostic hypotheses: Expert-novice differences, *Medical Education*, **21**, 471-476.

Jeffries, R., Turner, A., Polson, P., and Atwood, M. 1981, The processes involved in designing software, in *Cognitive Skills and Their Acquisition*, ed. J. R. Anderson (Erlbaum, Hillsdale, NJ), pp. 255-284.

Koubek, R. J. 1991, *Toward a Model of Knowledge Structure and a Comparative Analysis of Knowledge Structure Measurement Techniques* (Purdue University, School of Industrial Engineering).

Koubek, R. J., Benysh, D. V., and Calvez, V. 1993, A comparative review of knowledge structure measurement techniques for interface design, *International Journal of Human-Computer Interaction*, **5**(3), 211-238.

Koubek, R. J., Clarkston, T. P., and Calvez, V. (in Press), The Training of Knowledge Structures for Manufacturing Tasks: An Empirical Study, *Ergonomics*.

Koubek, R. J., and Mountjoy, D. N. 1991, The impact of knowledge representation on cognitive-oriented task performance, *International Journal of Human-Computer Interaction*, **3**, 31-43.

Koubek, R. J., and Salvendy, G. 1991, Cognitive performance of super-experts on computer program modification tasks, *Ergonomics*, **34**, 1095-1112.

Lewis, C. 1981, Skill in algebra, in *Cognitive Skills and Their Acquisition*, ed. J. R. Anderson (Erlbaum, Hillsdale, NJ), pp. 85-110.

McKeithen, K. B., Reitman, J. S., Rueter, H. H., and Hirtle, S. C. 1981, Knowledge organization and skill difference in computer programmers, *Cognitive Psychology*, **13**, 307-325.

Schoenfeld, A. H., and Herrmann, D. J. 1982, Problem perception and knowledge structure in expert and novice mathematical problem solvers, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **8**, 484-494.

Appendix

Table A1. Concept List for Master Production Scheduling Domain

Master Production Schedule
Product Structure
Final Product
Component
Subassembly
Record
Part ID
Planned Order
Requirement
Demand
Current Inventory
Product Database
Batch Size
Lead Time
Backward Scheduling

Table A2. Relationship List for Master Production Scheduling Domain

1	is composed of
2	is decomposed into
3	affects
4	is affected by
5	is calculated with
6	is used to calculate
7	utilizes
8	is utilized by
9	requires
10	is required by
11	results in
12	is the result of
13	includes
14	is within
15	identifies
16	stores
17	is added to
18	is subtracted from
19	must satisfy
20	proceeds
21	supports
22	is the goal of

Table 1. Proposed Operational Definition of Knowledge Structure.

Element	A unique unit of knowledge, or concept, within a domain.
	Domain elements can exist in either a declarative or procedural format.
	Domain elements can exist at different levels of hierarchy within a knowledge structure.
Interrelationship	Relationships, or links, exist between domain elements.
	More than one relationship can exist between two domain elements.
	Relationships can vary in strength, or degree of applicability.

Table 2. ANOVA for Completion Time.

SOURCE	dF	SS	MS	F	p
Training Condition	1	.20	.20	.0025	.9609
Speed of Closure	1	16.20	16.20	.2010	.6599
Training Condition X Speed of Closure	1	369.80	369.80	4.5881	.0479
Error	16	1289.60	80.60		
Total	19	1675.80			

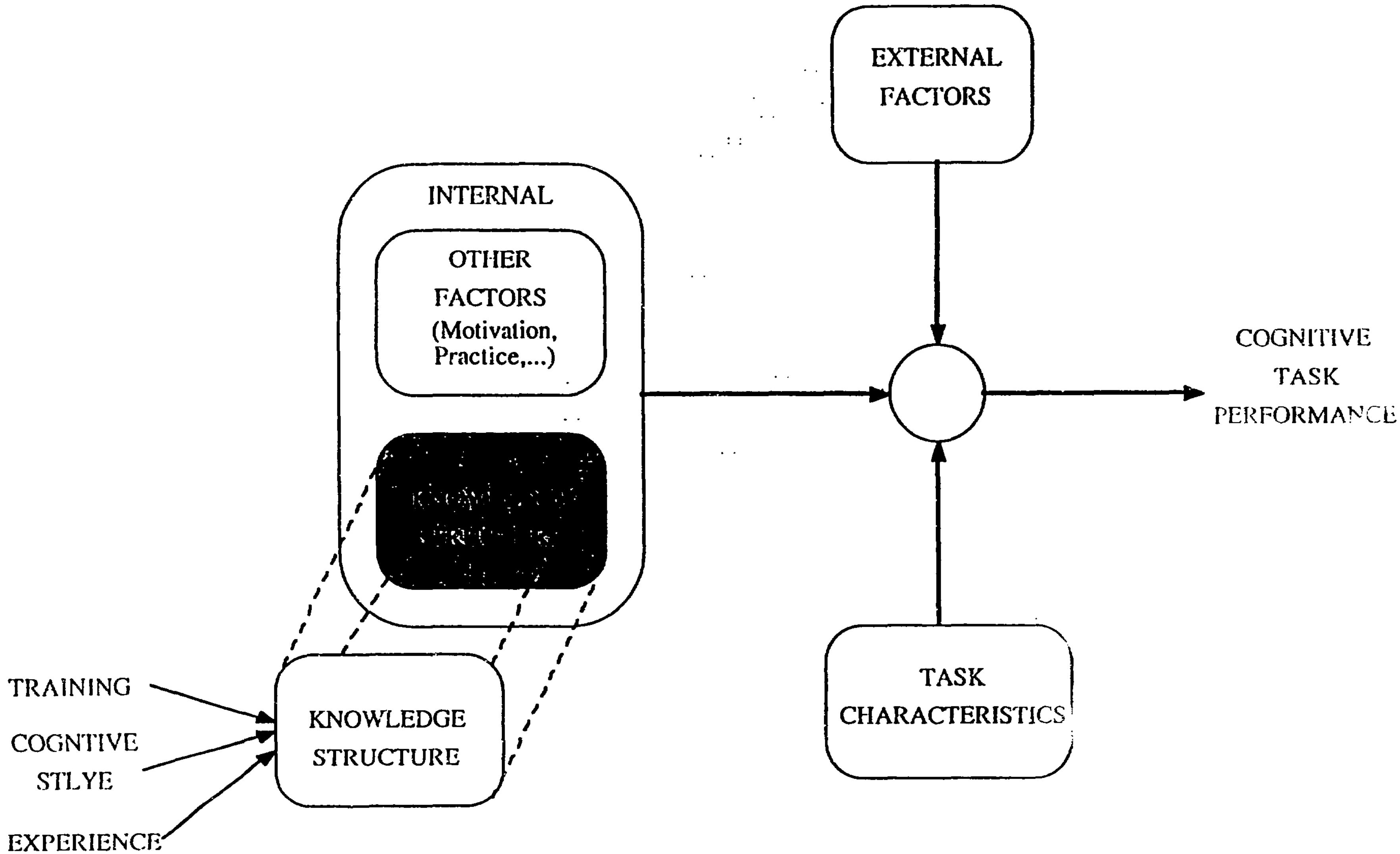
Table 3. ANOVA for Error.

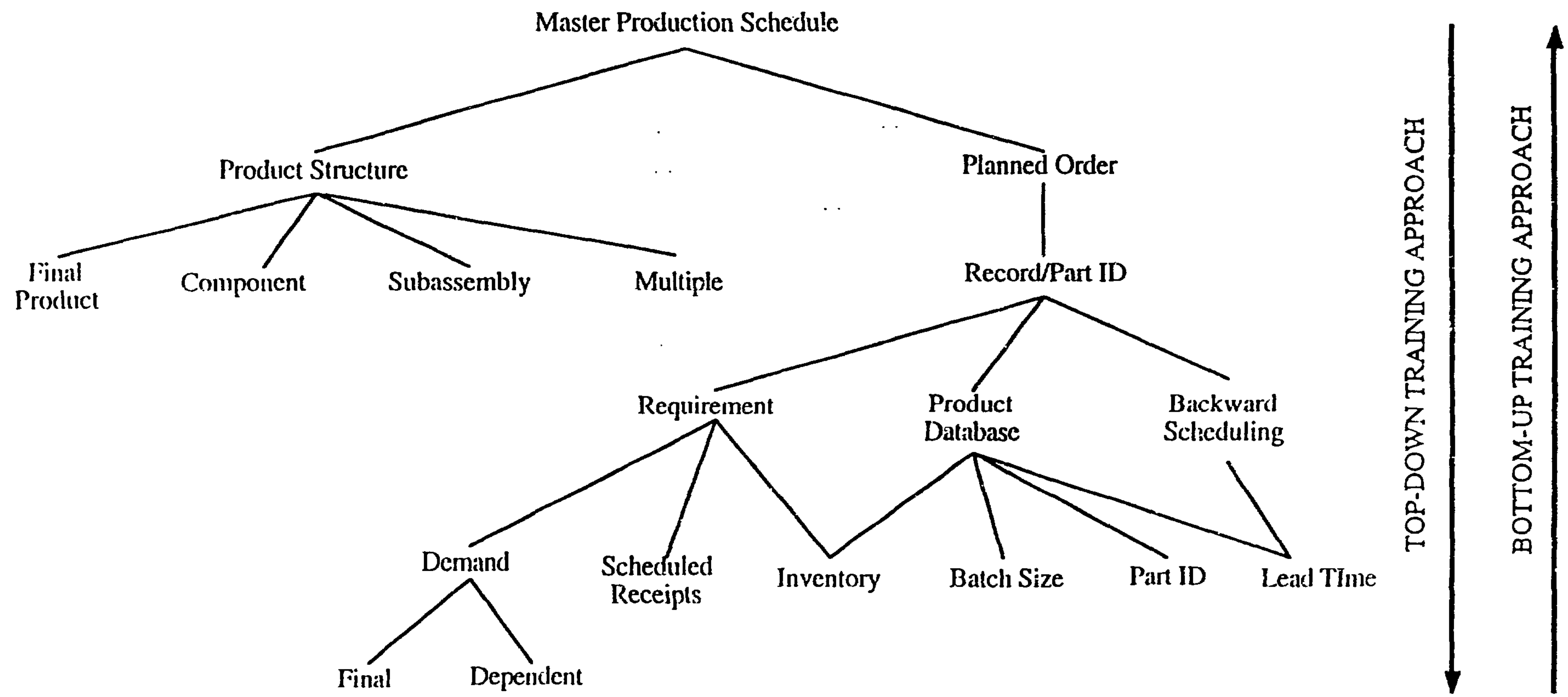
SOURCE	dF	SS	MS	F	p
Training Condition	1	5.00	5.00	.0431	.8382
Speed of Closure	1	39.20	39.20	.3379	.5691
Training Condition X Speed of Closure	1	540.80	540.80	4.6621	.0464
Error	16	1856.0	116.0		
Total	19	2441.0			

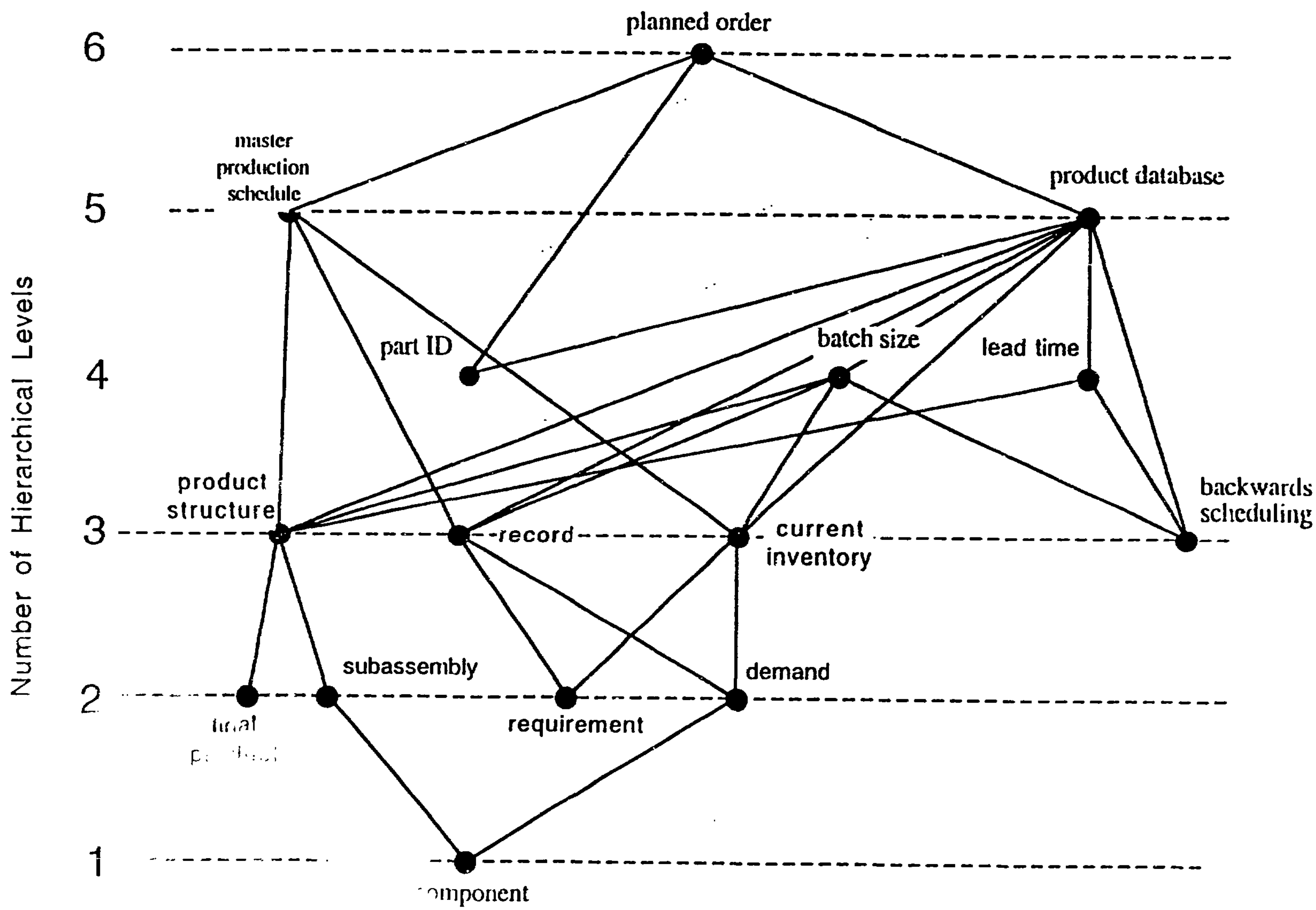
Table 4. Regression Analyses Summary.

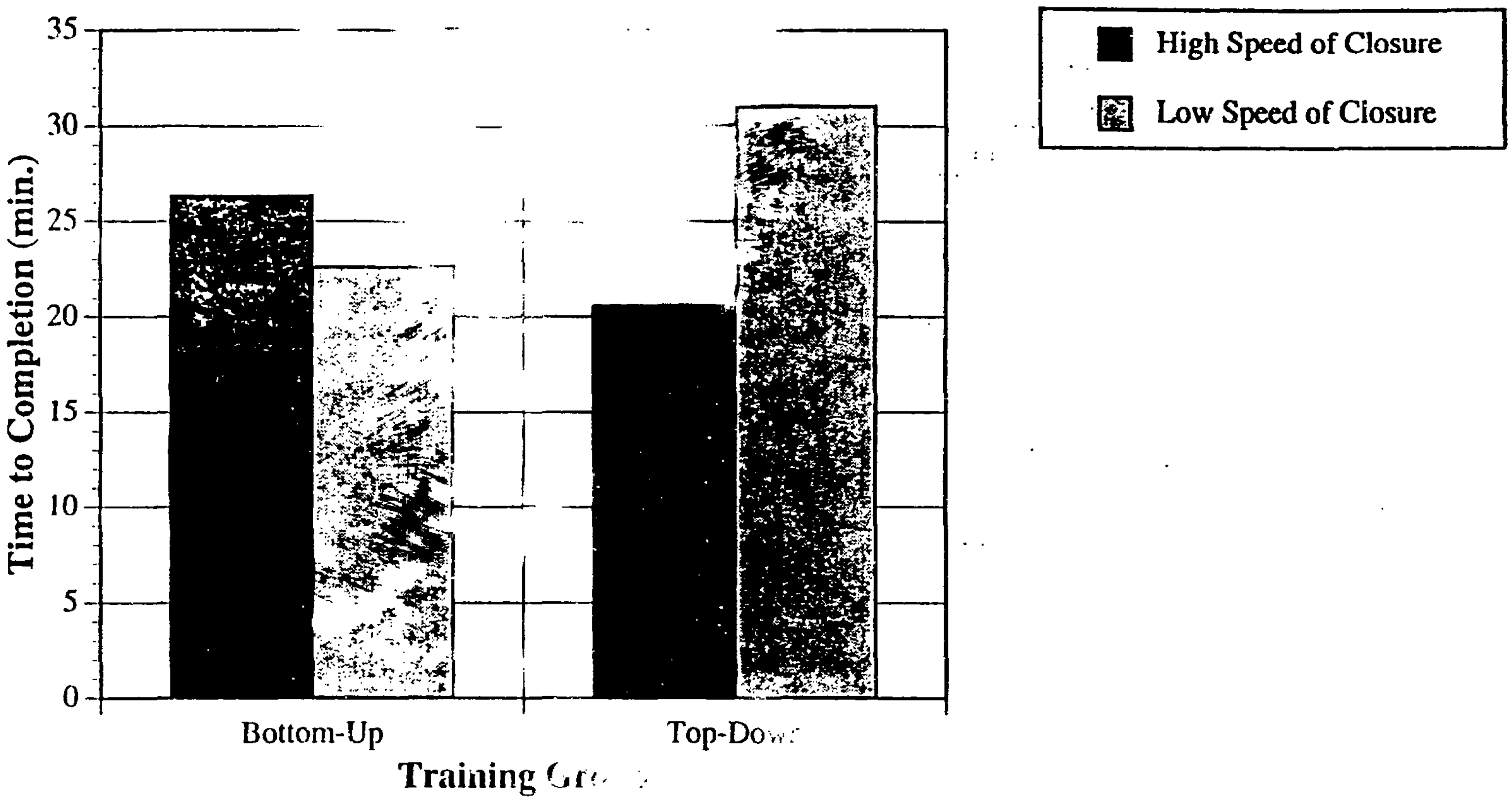
Dependent Variables	Independent Variables	R-squared
Completion Time	Cognitive Style	.08
	Knowledge Structure	.22
	Cognitive Style and Knowledge Structure	.33
Error	Cognitive Style	.24
	Knowledge Structure	.15
	Cognitive Style and Knowledge Structure	.33

LIST OF FIGURES

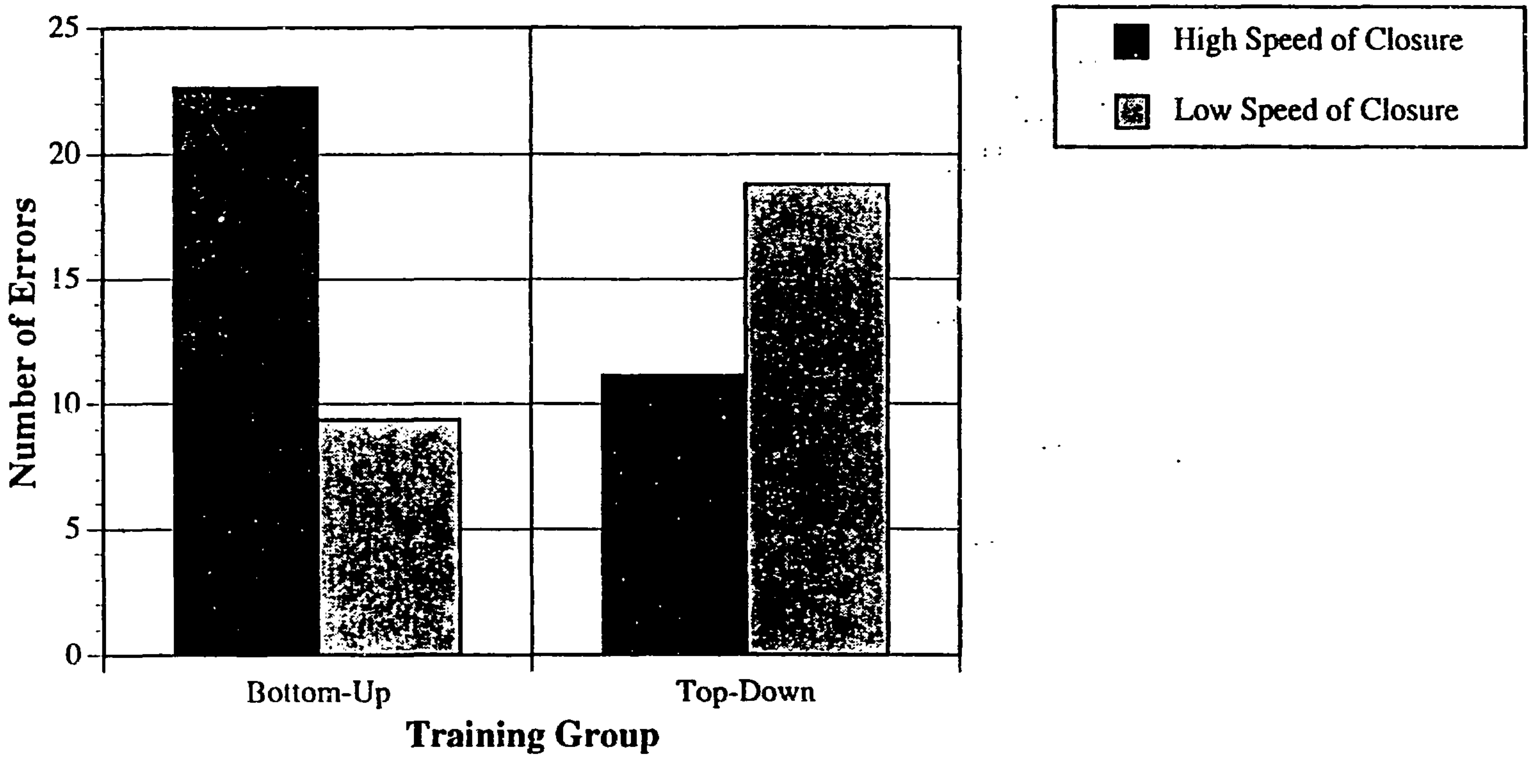

EXTERNAL FACTORS
INTERNAL
OTHER FACTORS (Motivation, Practice,...)
COGNITIVE TASK PERFORMANCE
TRAINING
COGNTIVE STLYE
EXPERIENCE
KNOWLEDGE STRUCTURE
TASK CHARACTERISTICS


Master Production Schedule
Product Structure
Planned Order
Final Product
Component
Subassembly
Multiple
Record/Part ID
Requirement
Product Database
Backward Scheduling
Demand
Scheduled Receipts
Inventory
Batch Size
Part ID
Lead TIme
Final
Dependent
TOP-DOWN TRAINING APPROACH
BOTTOM-UP TRAINING APPROACH


Number of Hierarchical Levels
6
5
4
3
2
1
planned order
master production schedule
product database
part ID
batch size
lead time
product structure
record
current inventory
backwards scheduling
subassembly
demand
requirement



35
30
25
20
15
10
5
0
Time to Completion (min.)
Bottom-Up
Training
High Speed of Closure
Low Speed of Closure



25
20
15
10
5
0
Number of Errors
Bottom-Up
Top-Down
Training Group
High Speed of Closure
Low Speed of Closure

ATTACHMENT 6

A Model of Procedural Knowledge Structure Representations

Darel V. Benysh and Richard J. Koubek

This Research was supported by the Office of Naval Research Cognitive Sciences Program under grant # N00014-92-J-1153. The opinions expressed here do not necessarily reflect the position of ONR.

A Model Of Procedural Knowledge Structure Representations

Darel V. Benysh and Richard J. Koubek

School of Industrial Engineering Purdue University West Lafayette, Indiana 47907, USA

This research proposes a model combining the outstanding features of Cognitive Modeling techniques into a Knowledge Organization framework in order to improve the accuracy in predicting human performance. The resulting Procedural Knowledge Structure Model (PKSM) has been evaluated to assess likely structural dimensions which have an effect on task performance. These dimensions are defined in terms of quantifiable measures, which are empirically validated. Results indicate that the PKSM measures, and thus model dimensions, are significant indicators of aspects of task performance. Furthermore, these measures provide more predictive power than traditional knowledge structure dimensions, and a combined model (with both sets of measures included) provides yet even stronger predictions.

This work was supported by the Cognitive and Neural Sciences Division, Office of Naval Research, under grant number N00014-92-J-1153. The opinions expressed in this paper do not necessarily reflect the position of ONR.

RUNNING TITLE: Procedural Knowledge Structures

Mailing Address:

D. V. Benysh or Dr. R. J. Koubek
School of Industrial Engineering,
1287 Grissom Hall,
Purdue University,
West Lafayette, Indiana,
47906-1287, USA.

1 OVERVIEW

It is widely recognized that recent trends in work activities have moved from manual oriented labor to include supervisory and cognitive tasks. The demands imposed by these new work related activities is still not adequately known. This has produced the need to develop greater understanding of the factors contributing to higher levels of skilled cognitive task performance.

A number of research areas have identified factors to account for, or explain, particular features of cognitive task performance. These areas include cognitive abilities, skill acquisition, cognitive modeling, and knowledge organization. Research in cognitive factors has identified certain basic cognitive abilities that can be associated with task performance. Skill acquisition has focused on how people develop task related skill. Cognitive modeling focuses on simulating human cognitive performance in an effort to understand it. Knowledge organization research is concerned with how humans structure information and how that structure affects performance.

Each of these approaches have been able to successfully account for a portion of task performance and expertise. However, research in each of these areas tends to focus on task performance exclusively from a single vantage point. It is most likely that skill and expertise are multidimensional, in that no one approach is complete. For example, cognitive abilities can be combined with skill acquisition research to determine whether there are fundamental differences between abilities need to *learn* a task versus abilities need to *perform* a task.

A multi-factor approach of skill and expertise is proposed in this study. This statement implies that these multiple factors are complementary, rather than additive, when viewed in terms of understanding skill and expertise. Of consequence to this premise is the necessity of a framework or model which can provide for the integration of a multi-attribute approach to skill.

The purpose of this research is to develop a hybrid model of knowledge organization and cognitive modeling (CM), combining aspects of some of the more popular models in each of these areas in an effort to provide a foundation for further study. It pursues this goal through the adaptation of Knowledge Structure (KS) representation techniques to model human structure of procedural knowledge and validating this model in an experimental setting. The rationale for selecting KSs and CMs will be discussed in the following section.

2 BACKGROUND

In order to develop a structural model of procedural knowledge as described in Section 1 current frameworks of both procedural and declarative knowledge are reviewed as a basis for derivation of the proposed model. Namely, evaluation of the contributions that may be provided through integrating the cognitive modeling and knowledge organization bodies of research is performed. Work in this conjoined area is limited due to the nature of the two fields. Knowledge organization is fundamentally a science of declarative memory. On the other hand, cognitive modeling research has emphasized performance of tasks and the procedural knowledge associated with such tasks.

Many theories have been proposed to describe the various types of human memory and knowledge used in cognitive tasks. These theories are usually specific to the areas in which they are used, such as psychology, cognitive science, or human factors engineering. The two main classifications of knowledge used in the domain of cognitive modeling are declarative and procedural knowledge. Declarative knowledge is considered to be long-term conceptual knowledge about the world. It is knowledge about facts, principles and relations. In contrast to declarative knowledge, procedural knowledge is knowledge of *how* to do something. Once learned and proceduralized, it is also considered to be relatively static in long-term memory.

The advantages of each basic representation are summarized as follows (Bonnet, 1985; Winograd, 1975). Declarative representations are readable and easier to modify, especially by non-specialists. Second, they are economical and flexible. A declarative statement may be written once, and then used in different ways on different occasions. A procedural form must be repeated in every procedure that uses it, because usage may vary from one to the next. Next, declarative representations are more easily modified. This is particularly important for the evolution of the system and for providing the capability to learn from experience, that is, to modify itself.

Procedural representations, on the other hand, include constructs such as Meta-knowledge, future use dependence, and necessity. Meta-knowledge, or knowledge about knowledge, is more easily expressed in procedural form. The declarative form does not suit such information well since it requires such information to be highly situation dependent. Procedural representations also consider dependence on future use. Since clauses are often interpreted in the order in which they are written means that a second "and if..." will only execute if the first one does. Even though, in theory, statements can often be written without regard for the use that will be made of them later in the program, in practice the programmer will always have this in mind. Finally, ultimate necessity indicates that there is always a final level at which any declarative statement has to be interpreted and executed by some procedure. Therefore, there is always an irreducible kernel that must be programmed and must therefore be of procedural form.

Researchers attempting to model human behavior and performance have found the procedural / declarative distinction to be useful in describing extensive complex knowledge bases. Their use in this field stems primarily from the distinctly different types of knowledge they convey. Further, these classifications of knowledge are put to various uses in the different fields. Of interest at this time are the fields of cognitive science and cognitive modeling. Cognitive scientists have developed procedural and declarative models primarily for the purpose of examining, understanding, and quantifying various aspects of human behavior. In contrast, cognitive modeling researchers have pursued these two types of knowledge for the purpose of replicating the outcome of human cognition (i.e. AI and Expert Systems). Table 1 provides an overview of their different methods of cognitive modeling and knowledge organization, as a function of their form. The next three sections below will discuss these approaches in detail in order to identify features for incorporation into a hybrid model.

Insert Table 1 about here.

2.1 Knowledge Organization Frameworks

Traditionally, General Weighted Networks (GWN) (Schvaneveldt, Durso & Goldsmith, 1985), Multidimensional Scaling (MDS) (Schiffman, Reynolds & Young, 1981; Davison, 1983; Schvaneveldt, Durso & Goldsmith, 1985; Rips, Shoben & Smith, 1973; Olson & Rueter, 1987), Hierarchical Clustering Scheme (HCS) (Koubek, 1991; Benysh & Koubek, 1993; Olson & Rueter, 1987; Johnson, 1967), and Semantic Nets (Quillian, 1968; Neves & Anderson, 1981; Anderson & Bower, 1973; , 1989; Bonnet, 1985; Hendrix, 1979) have been used for representation of knowledge organization. However, a model of human knowledge structure has been proposed and validated by Koubek (1991). The model identified and defined elements and their interrelationship as important components in this structure. Elements can be procedural or declarative concepts in a particular domain which may exist within multiple levels of abstraction. The strength of these relationships can vary, as well as can multiple relations exist between concepts. This model has provided a functional description of the features and attributes of human knowledge representation. It is from this foundation that this study draws upon in the development of a model of procedural knowledge structure.

2.2 Cognitive Modeling

In contrast to knowledge organization, cognitive modeling methods focus on attempting to simulate human performance. These simulations can be described as either Exploratory, if they attempt to mimic human processes, or Functional, in the case of using any tools that work for the simulation. This section reviews a number of cognitive modeling frameworks which are considered important for consideration.

2.2.1 Exploratory Models

This field is primarily concerned with modeling *how* humans organize, learn, and use knowledge (typically procedural). This leads to a concentration on extremely specific problem domains or general broad sweeping approaches intended to account for many aspects of human cognition at once. They focus on highly specific aspects or incorporate various theoretical works into a proposed complete comprehensive "holistic" theory of cognition. Work in exploratory modeling is generally performed with tools such as GOMS (Card, Moran & Newell, 1983; Marchionini & Sibert, 1991), NGOMSL (Kieras, 1988; Kieras & Elkerton, 1991), HAM/ACT (Anderson & Bower, 1973; Anderson, 1976; Anderson, Kline & Lewis, 1977; Solso, 1979; Gardner, 1985; Solso, 1979; Singley & Anderson, 1989; Anderson, 1983; Neves & Anderson, 1981), GPS (Newell & Simon, 1972; Card, Moran & Newell, 1983; Gardner, 1985; , 1989), TAG (Koubek, 1991; Benysh & Koubek, 1993), CCT (Bovair, Kieras & Polson, 1990; Polson & Lewis, 1990; Polson, Bovair & Kieras, 1987), or SOAR (Newell, 1992).

2.2.2 Functional Models

In contrast to work in exploratory cognitive modeling, where the goal is to replicate cognitive processes, researchers in this area tend to work in performance frameworks. This performance approach attempts to build systems that behave intelligently with little consideration for how humans do so. Researchers in simulation attempt to accomplish the same system goals except through emulation of human characteristics, but are often hindered by limitations of the machines or languages available (Weizenbaum, 1976). Functional modeling is typically performed within the Rule-based, Logic-based, Object Oriented programming, or Frames & Scripts frameworks.

2.3 Summary

This section has reviewed numerous knowledge representation schemes including varieties of knowledge organization, and functional and exploratory cognitive modeling techniques. These models include both declarative and procedural forms of knowledge encoding.

It is observed that the primary drawback to these theories is that typically only the declarative structures have quantifiable variables that have proven to be significant determinants of task performance. Depending upon their usage, these procedural representations, being performance based, either don't make predictions or they make predictions for very specific task situations.

Additionally it is noted that typical method of circumventing or compensating for the shortcomings of these various representation schemes is to use multiple representations simultaneously. This is often convenient and sometimes necessary. However, this practice results in a few notable problems (Bobrow, 1975).

First, if different representations are used, routines are required to transform information from one representation to the other. Additionally, sometimes the encoding scheme in one representation may not have an equivalent scheme in the other representation. Also for different data types in the same database it may be more convenient to answer questions and perform some calculations in one scheme but other operations in the other scheme. Second, updating changes in data becomes a more serious problem. It may be required to maintain current data values in a declarative representation then incorporate routines in the procedural representation to retrieve and transform the data as needed. Third, as can be deducted from the above issues, computation efficiency also becomes a factor.

One method to avoid these difficulties would be to have a single representation that is capable eliminating the need for multiple concurrent representations. Towards this end, a list of common attributes found over the whole range of representation schemes could constitute an ideal list of desirable attributes for a comprehensive representation scheme. A summary of these features are as follows:

- procedural modularity (as in GOMS, NGOMSL, frames)
- increasing specificity when descending a hierarchical representation (as in NGOMSL)
- class / object-attribute inheritance (as in Semantic Nets, Frames)
- all encompassing classification scheme (as in Production Systems, Semantic Nets)
- increasing proceduralization at lower levels (as in NGOMSL, GOMS, Frames)
- can handle differing situations / initial states (as in rule based)
- represent general "proximity" of concepts (as in GWN, MDS, HCS)
- can represent individual differences in ability / expertise (as in GWN, HCS)

3 Theoretical Model

In this section a model is developed in pursuit of the first objective in this study which integrates human task knowledge structure with the computerized cognitive modeling practice into a hybrid representation scheme.

As stated, during the development of computational Cognitive Models (CMs), it appears that structure of the knowledge plays a very minor role. However, it has been found that the structure of task specific knowledge is an important factor in task performance. Thus, inclusion of a reasonable KS in the CM would seem to provide a model with these two aspects complementary to one another, rather than over lapping.

The model developed here is first and foremost a procedural model. It is a structural organization of procedural knowledge. It is comprised of a taxonomy of methods or procedures, specific to a particular task environment. However, it uses a visual representation scheme similar to that of declarative knowledge theory in order to bring out the structural and organizational aspects of procedural knowledge. This combination benefits from the advantages of each of the representation models previously discussed by incorporating the identified desirable aspects. Indeed, a number of elements in the current models have been incorporated into what is termed here as the Procedural Knowledge Structure Model (PKSM). It has a significant amount of structure as is found in human cognitive representations. Also, it is used to capture procedural knowledge used in task execution, as in the production systems. Finally, it demonstrates the proceduralization, automation, or modularity capabilities present in many of the previous models. The primary contribution is the model's focus on the structural aspects of knowledge while simultaneously incorporating important procedural aspects.

3.1 Horizontal Component of PKSM: methods

Like the theory of declarative knowledge structures, the PKSM is best illustrated through its visual representation scheme. Unlike traditional KS representations or assessments, the elements of this model are, for the most part, procedural. It is always possible to represent the declarative elements and their relations in the traditional manner in a separate declarative representation if need be.

The model is best visualized as a three dimensional pyramid similar to that shown in Figure 1. The pyramid is sectioned into "floors" or levels. Each of these levels contains a flowchart representation of the task steps. The difference between the flowcharts of each level is that with each descending level, the items in the chart are broken up into smaller and smaller sub-units or sub-tasks, as to be described in Section 3.2. Each individual subtask at one level will be broken up into its constituent sub-tasks at the next level. This decomposition continues, level by level, until at the very bottom level, the flowchart is a very complex structure consisting of basic task actions (keystrokes, hand or eye motions, and decisions) which cannot be further decomposed except at the human neurological level. Objects in the flow chart may be task goals (which will be further decomposed), basic task actions (which can be performed or executed), and decision nodes (which control the flow through the chart).

Insert Figure 1 about here.

In Figure 1 task goals are represented by rectangles. These goals will be subdivided into smaller goals, task elements, or decision nodes at the next lower level, much like hierarchical task analysis. In this same figure, basic task actions are represented by the ellipses and decision nodes by diamonds. These task actions and decision nodes, once encountered, will be perpetuated through to lower levels since they cannot be further decomposed. These individual levels have the advantage that each level's flow chart, taken separately, can be executed from start to finish in order to perform the task. This can be accomplished by proceeding through the chart executing each basic task action and branching at decision nodes.

Horizontally, objects are related by "occurs before" or "occurs after" type relations for simple items and "may lead to" or "is dependent upon" for decision items and the items that follow decisions. These are considered the link types of the horizontal portion of the model.

The procedural elements in the task are laid out in the flowchart type organization. A horizontal chart displays the actions, computations, decisions, and outputs necessary in performing the task. Unlike traditional flowcharts though, it has the capability of simultaneously representing many different ways of doing the same task, as Figure 2 illustrates. This is known as the principle of equifinality. It is not restricted to one particular strategy and does not focus on any best way of doing the work. In this manner, it is able to model the various different approaches used by different people ranging from novice to expert level performance with differing experience and ability. Furthermore, these decision nodes mimic production rule systems by defining a direction to take based on facts in the environment, or NGOMSL selection rules by selecting path to take. Also, this expands upon the usefulness of scripts by incorporating methods for each unique task situation.

Insert Figure 2 about here.

The decision node aspect of the model follows from assumption that people approach tasks in a goal directed manner, proceeding in an organized fashion deciding at any particular stage what the next action should be. Alternatively, a series of steps without a decision node represents a set routine where no goal evaluations need to be made. Decision nodes should accumulate at the earlier stages of the task procedure where the solution is still being resolved; similar to NGOMSL, GOMS, or production systems.

3.2 Vertical Component of PKSM: organization

Sub-task proceduralization capabilities are found by moving upwards through the pyramid structure. Working from the bottom up, groups of elemental task actions can be grouped together to form a single task item (or goal) at the next level up as shown in Figure 3.

The elemental items are collected or said to be proceduralized into the next higher level. Each higher structural level represents an increase in the level of task proceduralization than the one below it. The highest level will simply be the singular task statement or task goal. Vertically, objects are related through goal – sub-goal classifications or subordinate – superordinate relationships. These are considered to be the vertical link types of the model.

Insert Figure 3 about here.

Proceduralization, or development of a set of routines for a common set of actions, is a natural aspect in human task performance (Shiffren & Shneider, 1977; Shneider & Shiffren, 1977). This is also indicated in Rasmussen's shift from rule-based knowledge to skill-based knowledge in his theory of behavior categories in industrial performance. This has been incorporated into various knowledge-based systems and computerized cognitive modeling systems as learning by deduction, where a series of steps will be treated as a unit once it is determined that the end result is always the same (Tseng, Law & Cerva, 1992).

Goals are often subdivided into smaller goals and operators similar to GOMS or Gonéz's methods. This decomposition continues downward until lowest terminating elements are all operators.

However, in these goal hierarchys there is no stratification of proceduralized levels, indicating uniform levels along which operators or goals are at a common level of proceduralization. Further, in typical goal hierarchies, links between temporally related consecutive operations are not presented and not emphasized at any point in the GOMS goal hierarchy. One objective of the PKSM is to treat the process as a list or flowchart where order of operations is apparent.

3.3 Extended Example

Clearly, the PKS model is one of scale or scope. This is evidenced by the fact that any individual proceduralized sub-task can be separated from the structure to stand alone as a separate task. This procedure will then become the overall task with its sub-goals and sub-tasks self contained. Likewise, the original task structure can be subsumed into a larger overall task and become one of the many sub-goals within it. For example, the task of driving to North Carolina can be seen to include individual sub-tasks of packing the car, starting the car, driving the car, etc. Driving the car can be further broken down to many sub-task of staying on the road, picking the route, stopping for gas as needed, eating, changing the radio station, and many other things. Each of these can continue to be broken down until they reach the basic decisions, and other elemental task items such as fundamental eye, hand, and foot movements necessary to complete the task. Further, this task example of driving to North Carolina can be included into a larger task of making your fiancée happy on an anniversary. This larger task could include such immediate sub-goals as, picking a gift, deciding on a departure time, getting prior commitments done in advance, driving to North Carolina, going to dinner, etc...

Insert Figure 4 about here.

A small fragment of this structure may been seen in Figure 4, illustrating the goal of starting a car. In this illustration, the difficulties of representing a 3–dimensional structure are bypassed by displaying the contents of each level separated with a dashed line. For clarification, a profile structure is presented in Figure 5 indicating the proceduralization aspects of the model but losing some of the specific objects at each level. These figures show how the goal of Starting Car follows Entering Car and precedes the goal of Engaging Gear. Furthermore, they indicate how starting a car can be composed of a decision, and two alternative methods. Each of these methods is then decomposed at the next level down into their respective components. The goals at this level, Prepare Ignition and Check Gear, would be further specified at the next level down, presumably approaching an elemental action level. The action at this third level, Turn Ignition, is an elemental action and is represented in an ellipse. This ellipse will be continued down to the next level in order to produce a complete flow chart representation at that level as well, similar to how the decision nodes are carried from the second to third level. Note also how the procedural link from Start Car to Engage Gear are perpetuated downwards through the structure to connect to lower levels of the Engage Gear goal partially represented here.

Insert Figure 5 about here.

3.4 Individual Differences

It is quite apparent from the previous example that such a structure can become quite complex. The previous example did not attempt to convey any special conditions which may be required to account for in the goal, other than transmission type. For instance, manual chokes, cold weather starting, parking brakes, etc could also have been included. Also, this example illustrated only one particular method of performing this goal. Another individual's structure may have the elements in a slightly different order or their structure could be more or less extensive depending on their needs. For instance, if the individual had never driven a car equipped with a manual transmission, there may only be a vague notion of the actions in that branch of the structure, or there may be no decision and branch at all. This leads to the possibility of individual differences between models and their impact on behavior.

Individual differences may take almost any conceivable form. The most likely of which are:

- differences of object order.
- differences in content.
- objects at different levels.
- differences in level of conceptualization.
- incomplete knowledge in some areas.

- differing control structures.
- differing control requirements.

These differences provide the potential for a broad spectrum of structures for the same task as might be held by many different individuals, thereby producing different behaviors. This spectrum provides the thrust for the interest in measuring difference PKS features in these individuals and how they affect individual performance.

3.5 Measurement of Model Dimensions

Since humans often perform tasks differently and each has a unique performance level, it is possible that certain featural measures of their representation would be indicative of certain aspects of that performance. The intent of this research is in understanding and assessing basic structural features, analogous to traditional KS measurement techniques. In this way, a set of tools to ascertain the postulated PKS dimensions can be developed in accomplishing the second objective of this study; to develop a set of quantitative, non-referent, model parameters which can be objectively measured and used to understand and predict human behavior. There are many potential aspects of PKS which are of interest, as will be discussed below. This section will define these measures, as well as refer to a few others which may be of future interest.

3.5.1 Number of Decisions

The number of decisions (ND) in a given structure is an indication of the amount of differentiation of control, or flow, through the environment. Although the total number of decisions does not distinguish between decisions at different levels or the continuation of decisions downward through the structure, it still reflects a certain level of control differentiation.

3.5.2 Number of Hierarchical Levels

A measure of the number of hierarchical levels (HL) is a typical measurement in the KS theory that the PKSM has evolved from. However, since a HL measure would only indicate a perceived level of task complexity in this case, it is not a reliable or useful measure with respect to performance. Although interesting in its own right, HL will be substituted by a measure more sensitive to hierarchical decomposition, namely the number of super or subordinate links identified in the structure. This measure provides insight into the degree of task decomposition held by the individual. Note that individuals with a relatively low level of conceptual knowledge, consisting only of elemental task procedures, would have a low value on this measure, as would someone who only knows the upper level organization of the task (high conceptual knowledge).

3.5.3 Path Length

A measurement of path length through the structure, or the number of steps required in the execution in a task, is a measure very similar to counting steps in the execution of a GOMS or NGOMSL representation. This step count can then be used to make predictions of time, working memory load, learning time, and learning transfer as in NGOMSL. However, this count would be misleading with respect to the links present in the structure as a whole, including those present in parallel paths and at different levels. It is this whole structure that is of current interest in verifying the existence of a structural model similar to that of traditional knowledge structures. Thus, at this time, the measure of the number of procedural links between concepts found in the model structure will be used to examine the breadth and procedural emphasis contained therein. With regard to the individual, this measure would be high if they perceive many connections between steps.

At some point however, after model validation, analyzing the structure by way of path length, and comparing results (of time, etc.) to those of NGOMSL would be a worthwhile endeavor.

3.5.4 Object Measures

Potential measures as to the ordering and variety of the objects in the structure may also prove to be enlightening. Order of operators can be presumed to have a significant impact on human task behavior, but not necessarily task performance since many different strategies, techniques, or methods may be applied to the same ends. Once again, this area is indicated for further study.

3.6 Summary

This section has proposed a theory of representation of procedural and declarative knowledge and how such individual's representations can be evaluated with respect to certain model measures. These measures, taken together, cover the prime dimensions indicated in the PKS model. The next step is to examine the validity of the model by exploring how these differing representations affect cognitive task performance.

4 Validation Study: Objectives

The structural features of the developed procedural model are believed to indicate or represent important aspects of cognitive processing which effect task performance. This section, and the next two, address the third and fourth objectives of this study. Namely, validation of the model through these measures or dimensions in predicting and complementing traditional KS predictions of performance.

In the traditional KS model, the measure indicating the level of abstraction (HL) and and the measure representing multiple relations (MR) have been correlated with task performance. Furthermore, in both forms of structural representation, a new measure, the amount of uncertainty (AU), is hypothesized as a significant performance determinant. In the procedural KS model a measure of the number of hierarchical links (superordinate or subordinate), the number of procedural links, a ratio of hierarchical links to procedural links (h/p), and the number of decision nodes (ND) are hypothesized as being significant indicators of task performance. Each of these measures, in addition to the measures of performance, is formally defined within the the discussion of experimental methodology (Section 5.1). For now, increasing performance is defined as reduced execution time (TM), reduced number of errors (ER), and reduced error rate (E/T).

4.1 General Hypotheses

The overall hypothesis for the study is that a structural representation of task specific procedural knowledge will provide a representation of human task knowledge. As such, that representation contains measurable dimensions which will predict human performance on tasks.

More specifically:

1. The wide range of potential representations in the PKSM provides the ability to distinguish between individual differences in task representation (through measurement of these dimensions) and thus performance.
2. Furthermore, this capability should be sensitive enough to be able to distinguish between features of representations held by task experts and novices.
3. Next, the PKS model measurements, when combined with the traditional KS measures, will be better predictors of task performance than traditional KS model measures alone. This is expected due to the task specific content of the procedural model over that of the declarative content of previous KSs.

4.2 Formal Experimental Hypotheses

1. The PKS model measures (AU, h/p, ND) are able to predict performance (ER, TM, E/T).
2. The PKS model measures (AU, h/p, ND) are able to discriminate between experts and novices.
3. The PKS model is a better predictor of task performance than the traditional KS model.
4. The PKS model is better able to discriminate between experts and novices than the traditional KS model.
5. The combined measures of the PKS (AU, h/p, ND) and traditional KS (AU, HL, MR) models are able to increase the accuracy in performance prediction over that of each individually.
6. The combined measures of the PKS (AU, h/p, ND) and traditional KS (AU, HL, MR) models are able to increase the accuracy in expert / novice discrimination over that of each individually.

5 Validation Study: Method

The experimental design this study is a simple one-way ANOVA at two levels. Two groups constituted the levels, and each of these groups was assessed on a number of task and model variables. A wide range of data was collected including performance measures, declarative measures, and procedural measures.

5.1 Variables

The is only one primary independent variable used in this study, which is called Group. It is used to classify users of differing expertise levels in the task domain. The subjects were classified as being in one of two groups, either Experienced or Novice. Qualifications distinguishing between the two groups is further discussed in Section 5.3.

There were a large number of dependent variables in the study which can be classified as either performance, declarative, or procedural variables.

5.1.1 Performance Variables

The dependent variables associated with performance were time to complete the task, the errors that occurred in task performance, and the overall error rate.

The task completion time (TM) is a simple measure of the time required to read and complete the task. It is measured to a resolution of seconds and includes the time required to read the instructions as well as any time required to make corrections. While subjects who recognize errors and attempt to correct them may penalized through this measure, they could perhaps compensate for this through a lower number of errors in the final result, captured in the next measure.

The number of errors (ER) is a measure of the total number of errors committed during task execution. Errors are classified as errors of commission, omission, and extraneous. Errors of commission are defined as doing something incorrectly that is required in the method to perform the task. Errors of omission are defined as not doing something that is required. Extraneous errors are errors not included in the other types, these were expected to be (and proved to be) primarily mathematical calculation errors. The indirect effect of the extraneous errors, as perpetuated throughout the task, were ignored. The three categories of errors are summed up to determine a total error measure. Total error is used based on the rationale that no matter what the source of an error, it is still an incorrect action and results in an incorrect final state. To measure these values, all the tasks were graded by one person, who was not the primary experimenter, using a defined, standard process in order to obtain an unbiased measure with a reasonable level of consistency.

The overall error rate (E/T) is the ratio of the total number of errors to the task time.

$$\mathrm{E/T} = \frac{\mathrm{ER}}{\mathrm{TM}}$$

This ratio provides an indicator of the number of errors that could be expected in a unit of time (one minute). This combines the effects of the other two performance measures, yet can be deceiving if used alone. Therefore, all three will be used in the final analysis of performance.

5.1.2 Traditional Declarative Variables

The dependent variables associated with the declarative structure are hierarchical level (HL), multiple relations ratio (MR), and the amount of uncertainty (AU). The first two measures were developed and used by Mountjoy and Calvez, but have been formalized here to become independent of their specific usage (Koubek, 1991; Koubek, Clarkston & Calvez, 1993).

The hierarchy variable is a measure of the number of levels in the subject's elicited declarative knowledge structure. It is roughly a measure of the depth of hierarchical relationships in the subject's data (Koubek, Clarkston & Calvez, 1993). The actual relationship types will be discussed in Section 5.2.2. This measure is determined by hierarchal tree reconstruction based on the experimental data.

An operational definition of the multiple relations variable (MR), another traditional KS measure, is the average number of relationships for all concepts which have at least one relationship. It is computed as follows:

$$\text{MR} = \frac{\text{total \# of all relationships between all concepts}}{\text{\# of all concept pairs with at least 1 relationship}}$$

This calculation results in a number ranging from one to infinity (the number of multiple relations is theoretically unlimited, but is actually limited to the number potential relationships presented to the subject in the relationship list. A value of $MR = 1$ indicates that all pairs of concepts with any relationship had exactly one relationship (Koubek, Clarkston & Calvez, 1993).

The final dependent declarative measure is the amount of uncertainty in the network. This is a measure of the sum of the information required to distinguish between alternate outgoing relationships for each node, summed over the whole structure. This measure is an adaptation of Information Theory which states that the information required to make a decision between equally likely alternatives is:

$$H = \log_2(N)$$

where H is the amount of information required (measured in bits) and N in the number of alternatives. In the case of a series of decisions, as in tracing all the links or paths available in a hierarchical tree, this equation becomes:

$$H = \sum_{i=1}^{K} \log_2(N_i)$$

where K is the total decisions, events, or paths traversed. Although this equation is intended to measure the information required in a series of items, it also provides a reasonable measurement of the total uncertainty of a hierarchical tree when summed over the entire structure. When used for this purpose, K becomes the total number of concepts with outgoing relationships and N_i is the number of outgoing relationships for concept i (Clarkston, 1993).

5.1.3 Proposed Model Variables

The dependent variables associated with the procedural structure are number of decision nodes (ND), number of hierarchical links (#h), number of procedural links (#p), the ratio of hierarchical links to procedural links (h/p), the amount of uncertainty (AU), and the multiple relations ratio (MR).

The first four measures are based on theoretic features of the PKS model and are discussed in Section 3.5. The last two measures are discussed above, and are applicable here as this is also a structural model of human knowledge representation and contain similar features.

The computation of the PKS measures, #h, and #p are quite straight forward counts of differing link types. The ratio h/p is similarly uncomplicated. It is defined as the ratio #h/#p. The number of decision nodes (ND) will be determined through assessment of the number of links associated with a special decision object in the procedural concept list used during the elicitation stage. As mentioned in Section 3.5, these measures are selected for their powerful objective representation of various structural features in the PKSM, however over simplified they may appear.

5.2 Stimuli

During experimentation, the subject was presented with a small variety of different stimuli, each related to separate parts of the study. A training document was used in a training section and concept and relationship lists were used during the KS elicitation section.

5.2.1 Training Document

This study made use of a training document to train or review the subject on the task of Master Production Scheduling (MPS). It was developed as a part of a set of documents used in another study (Clarkston, 1993) examining the effects of different training methodologies on KSs. Part of this study involved the organization of the content of the information in the training documents. The top down version of this training document was chosen because of the orderly presentation which should facilitate learning for the novices as well as refresh the experts adequately. It was organized in a top down fashion, from discussion of higher level concepts working to more detailed items.

5.2.2 Relationship Lists

Two relationship lists were developed, one each for the declarative and procedural elicitation sessions. The particular relationships for depicting potential hierarchy relations was developed by Clarkston (1993).

The procedural relationships list consisted of a set of relations which would enable the subject to indicate superset or subset and temporal relations. The development of the list followed from the work of Quillian (1968) and that of Rodriguez et al (1991). The relationships used to define potential hierarchical links were:

- X is superordinate to Y. (X contains substep Y)
- X is subordinate to Y. (X is a subset of Y)

These relationships were presented to the subject during elicitation along with the parenthetical explanation.

The remaining relationship types were defined in terms of procedural relations. Rodriguez and Anger (1991) began work in this area by introducing the concept of temporal relations to describe events that occurred in time. This led to the development of procedural relationships which linked concepts or steps which occurred in sequence:

- X occurs immediately before Y.
- X occurs immediately after Y.

Additionally, two more relationships were created to handle the conditional procedural relations which would be characteristic of loop structures in the subject's KS:

- X may lead to Y. (Depending on certain situations)
- X may follow Y. (Depending on certain situations)

These three pairs of responses should be able to describe any relationship which exists between concepts in a serial, procedural environment.

5.2.3 Concept Lists

Similar to the relationship lists, two concept lists were also developed for eliciting links. The declarative concept list was comprised of concepts extracted out of a hierarchical task analysis. The procedural list was composed of concepts obtained from a procedural task analysis. The result of these differing techniques was the creation of a set of terms / concepts meaningful for the particular elicitation portion they were to be used for.

5.3 Subjects

A total of twenty (20) subjects participated in this study. There were ten subjects in each of the novice and experienced categories. Novices were individuals with no knowledge in the task domain. Experienced individuals, on the other hand, were selected based on their knowledge in the task domain. The experienced group consisted of manufacturing/production graduate students and one senior in Industrial Engineering (IE) who demonstrated proficiency through course work in the domain.

The requirements for novices were as follows: (1) that they had no experience with Master Production Scheduling (MPS) or Materials Resource Planning (MRP), (2) they had no knowledge of the terms used in either of these areas, and (3) they had no prior manufacturing experience. All this information was obtained through a preliminary phone interview as a form of prescreening.

5.4 Apparatus

During experimentation, the subject was provided with certain items in order to perform either the task or elicitation phase of the experiment.

5.4.1 Task Phase

For the task phase of the experiment, a standard hand held calculator, a pencil with erasure, and the task sheets were provided.

5.4.2 Elicitation Phase

For the KS elicitation phase, a computer program was used that collected the data on each individual. The software, developed for this experiment, elicited directed relationships for pairwise comparisons of items from the appropriate concept list. This data collection was performed on a Sun 3/50 workstation.

Also for the elicitation phase, a sheet of paper with either the declarative or procedural relationships and concepts was placed on the table next to the workstation for the subject to refer to. Each concept and relationship set was on its own sheet of paper and only the sheet required for the current elicitation session was available.

5.4.3 Knowledge Structure Elicitation Scheme and Tools

In addition to this model development, a number of programs have been constructed to facilitate it's analysis. The Knowledge Structure Elicitation and Analysis Method (KSEAM) tools were developed specifically for this experiment, and consisted of two separate programs. The first program uses pairwise similarity ratings to extract the relationship(s) between concepts or items. This program exists in two versions, one for the traditional declarative KS and one for procedural KS model discussed. Using the procedural KSEP, subjects are asked to indicate relationships between concepts (task items) with respect to temporal of hierarchical relationships.

Following the relation extraction stage, the resulting data sets are analyzed by the procedural KSAP. This program calculates the quantitative measures of the model's structural features which were discussed in Sections 5.1.2 and 5.1.3.

5.4.4 Program Interaction

The KS Elicitation Program (KSEP) was written in the "C" programming language and incorporated the "curses" screen manipulation routines to format the display. The program is used to elicit relationships between pairs of concepts presented in random order.

For each concept pair, the subject decides if a relationship existed between the concepts from the list of relationships provided. To enter a relationship, the subject would enter the relationship number followed by the letter "i" (for insert). To enter additional relationships for the current concept pair, the same procedure was performed. To remove a relationship which the subject entered but decided against, the subject would enter the negative sign ("–") followed by the appropriate relationship number and the letter "i". If there were no relationships or after entering all applicable relations, typing "n" (for next) would proceed to the next concept pair. After all concept pairs are presented, the program would end.

5.5 Procedure

The testing of subjects consisted of three main phases performed in a single session. The session consisted of a training phase, a task phase, and a structure elicitation phase.

At the beginning of the session, the experimenter introduced the training phase by presenting the training document and giving specific instructions. The instructions included mentioning that there was no time limit for training, they should read the document thoroughly, and that the subject would be required to pass a proficiency test before proceeding to the next phase. Passing the proficiency exam required that all answers be correct. If the subject did not pass the proficiency test, they would be required to review the training document, then attempt the proficiency test again using different questions. The proficiency test consisted of a set of calculations (an MPS record) constituting a basic skill required to perform the task. This cycle could continue up to a total of four times before the subject would have to be rejected. However, no subject required more than three proficiency tests.

After the individual had demonstrated proficiency in the task domain, they were presented with the task: the second experimental phase. The task was comprised of a manufacturing scenario and the associated data required to perform the MPS task. The task was a pencil and paper task with a calculator available as well. Subjects were instructed to read the scenario and perform the MPS task they just learned or, as in the case of the experts, reviewed. They were also informed that the task would be graded with respect to time and errors. The experimenter emphasized that accuracy should not be sacrificed for speed.

The third phase of the experiment entailed eliciting the subject's structure of MPS task knowledge. This required eliciting the the declarative structure as well as the procedural structure. The order in which these elicitations occurred was randomized such that the subject performed the declarative then the procedural structure elicitation, or vice versa.

The interaction with the elicitation program is described in Section 5.4.4.

6 Validation Study: Results

In this chapter, the last phase of the third and fourth objectives will be presented. This will be accomplished through an evaluation of the model measures, on a per hypothesis basis, in order to examine identified model parameters.

Preliminary to this evaluation, it is important to point out that a couple underlying assumptions were also tested. First, although results are not indicated here, it was verified that the experienced group did indeed perform better than the novice group. This is important to validate that the task provided sufficient level of difficulty to distinguish between the groups without being too challenging. Second, there was no significant effect of the order in which the elicitations (procedural and declarative) were performed. This was important to examine due to the possibilities of recency / primacy or fatigue factors entering into the results.

The first pair of hypotheses were designed to evaluate the ability of the model to predict performance and distinguish between experience levels based on model measures. To test these hypotheses, an "all possible regressions" technique was applied, then the model with the highest adjusted R^2 was selected as the best model. The adjusted R^2 technique is used in order to account for the different number of variables possible in the predictor equations.

Hypothesis 1 The PKS model measures are able to predict performance.

To test this hypothesis, the above regression selection technique was applied to the all dependent procedural variables for each of the dependent performance variables. The results of this technique are shown in Figure 6.

The first column of these results show that the dependent variables Number of Decision Nodes and Multiple Relations were able to account for 39.47% of the variance in the dependent performance error variable with a p-value of .018 for the resulting model. Additionally, the beta parameters for Number of Decision Nodes and Multiple Relations were significantly different from zero in the resulting model at $p < 0.0110$ and $p < 0.0206$ respectively.

Figure 6 shows that the dependent model measure, Hierarchy/Procedural Links was able to account for 19.57% for the time variance in task performance with a p-value of .0579.

The third column of Figure 6 shows that the dependent model measures Number of Decision Nodes and Multiple Relations were capable of accounting for 64.79% for the variance in error rate associated with task performance. The beta parameters for these measures significantly different from zero with $p < 0.0003$ for either variable. Furthermore, the model's p-value (.0002) is also significant.

Insert Figure 6 about here.

In sum, the PKSM dimension measures have significant value in predicting performance with most of this ability coming from the Number of Decision Nodes and Multiple Relations measures.

Hypothesis 2 The PKS model measures are able to discriminate between experts and novices.

To test this hypothesis, a between means comparison for each dependent procedural variable was performed. The results are presented in Table 2. It is seen that Amount Uncertainty and Number of Procedural Links are highly significant between groups.

Insert Table 2 about here.

Actual estimates indicate that the Number of Procedural Links and Amount of Uncertainty both have higher means for the novices. This confirms that the novices perceive the task more in terms of the process than the experienced group does, and that they have more uncertainty (in terms of information theory used to define the measure). Additionally it can be noted that Reliability, a measure of the consistency in the structure, was also significantly different between the two groups, with the experienced group displaying more consistency. Reliability was measured as a proportion of identical responses to a repeated randomly selected subset of concept pairs.

Additionally, a regression was performed for skill level, resulting in the values presented in the fourth column of Figure 6. This indicates that 44.08% of the variance in skill level can be accounted for by Amount Uncertainty. Further, this result is significant at $p < .0019$.

The second pair of hypotheses in this study was designed to compare the abilities of the model with those of the traditional KS model.

Hypothesis 3 The PKS model is a better predictor of task performance than the traditional KS model.

To test these hypotheses, the same "all possible regressions" technique was applied as before, except using the dependent declarative variables to predict the various dependent performance measures. The resulting R^2s from both models were then compared to test for a significant difference using the Fisher transformation technique to test the difference between two independent rs.

Figure 7 indicates that the procedural measures were able to account for 39.47% of the variance in error ($p<.0180$) while the declarative measures were able to account for 36.75% ($p<.0693$).

The Fisher test resulted in a p-value of .4562 for the difference between the R^2s. Although the regression indicates that the procedural measures were more accurate in predicting errors than the declarative measures, there is not significant difference in the R^2 values so this hypothesis is not supported.

Figure 7 also shows that the procedural measures were able to account for 19.57% of the variance in time ($p<.0579$) while the declarative measures were able to account for 24.72% ($p<.0303$).

The Fisher test indicates that there is no significant difference between the R^2 values (p-value $< .5871$). This indicates that the declarative measures were slightly more accurate in predicting times than the procedural measures, but not significantly so. Still, this hypothesis is not supported.

Figure 7 indicates that the procedural measures were able to account for 64.79% of the variance in error rate ($p<.0002$) while the declarative measures were able to account for 45.83% ($p<.0074$).

The Fisher test shows that there is no statistically significant difference between the R^2s at $p < 0.1867$, however it does tend to show a direction of improvement. This indicates that the procedural measures were more accurate in predicting error rates than the declarative measures, but not in a statistically significant manner, thus this hypothesis is not supported.

Insert Figure 7 about here.

Hypothesis 4 The PKS model is better able to discriminate between experts and novices than the traditional KS model.

These results, indicated in Figure 7, show that the procedural measures were able to account for 44.08% of the variance in expertise ($p<.0019$) while the declarative measures were able to account for 54.34% ($p<.0070$).

The Fisher test resulted in an insignificant p-value of 0.6736 indicating that there is no difference in the amount of variance account for in the two models. This, in addition to the fact that the regression indicates that the procedural measures were not as accurate in predicting error rates as the declarative measures results in this hypothesis not being supported.

The third pair of hypotheses in this study was designed to evaluate the benefit of combining the measures of the two models. The techniques used to do this is the same as that used by the previous set of hypotheses, namely building a model then testing for a significant increase of R^2 over that of each separate model.

Hypothesis 5 The combined measures of the PKS and traditional KS models are able to increase the accuracy in performance prediction over that of each individually.

The results of this hypothesis test are shown in Figure 8. The first three sections of this figure show that the combined models account for 51.05% of error performance variance, 47.89% of time variance, and 73.10% of the error rate variance, with p-values of .0308, .0179, and .0021 respectively. The resulting combined models contained four measurement variables for error performance, three variables for time, and five variables for error rate. These results indicate that the combined models are better than either individually, and that the procedural measures are able to significantly contribute to the accuracy of the declarative KS approach. Furthermore, it can be noted that in each case the procedural measures entered the regression first (indicating better fit) or had more terms (measures) in the resulting fit equation.

Insert Figure 8 about here.

The primary result here is that the procedural measures comprise over half the R^2 in the predictions of error rate and that the combined model demonstrates a significantly improved R^2 value over the over the Traditional KS model (at the $\alpha = .10$ level).

Hypothesis 6 The combined measures of the PKS and traditional KS models are able to increase the accuracy in expert / novice discrimination over that of each individually.

These results, also shown in Figure 8 as the last division, indicate that the combined procedural and declarative measures account for 61.17% of the variance in predicting or discriminating between expertise groups with four measurement variables in the model. Furthermore, the model resulted in an overall p-value of .0070. However, the respective p-values for the difference of R^2 terms are not significant. Therefore, this hypothesis is supported.

6.1 Summary

Table 3 is a summary of all the regression results discussed in this section including R^2s, p-values, the number of variables in the resulting model, and the increase in R^2 when the procedural measures are added to the traditional declarative measures.

Insert Table 3 about here.

This table show evidence for supporting many of the hypotheses in this study. However, for the final four hypotheses concerning comparisons between the individual R^2s indicates that none are significantly different at an $\alpha = .05$ level. Even as this is the case, many of the individual p-values for the R^2 comparisons indicate a marginal level of significance at the $\alpha = .10$ to $\alpha = .30$ levels. These values demonstrate that although the regression models do not have statistically different R^2s, the models still have some practical applicability.

7 CONCLUSIONS

In summary, this study examined relevant literature in the various knowledge representation fields related to cognitive skilled task performance, including Knowledge Organization, Exploratory Cognitive Models, and Functional Cognitive Models. The results of this review provided a compilation of a list of attributes common in many of the models, which were then specified as desirable in a new multi-factor model. This multi-dimensional approach was adopted due to the assumption that the factors associated with the various models are complimentary and that this integration would provide a more complete understanding of skill and expertise.

Following this, the study proposed a hybrid model of procedural knowledge structure representation, designed to unite these attributes into a common framework. This model embodied a number of structural dimensions which were hypothesized as being indicative of aspects of cognitive performance. A set of procedural knowledge structure measures were then defined based on these structural dimensions. Finally, this study validated that model with respect to performance on skilled cognitive task performance and traditional knowledge structure measures.

This validation indicated that the procedural model of knowledge structure is capable of predicting aspects of performance as well as expertise. There is also suggestive evidence that it can account for more variance than traditional knowledge structures when predicting errors and error rates. Further, results also indicate that the model contributes to the predictive capability of the traditional knowledge structure model with respect to performance and expertise level. The procedural model independently did not perform better than the traditional model of knowledge structure when making predictions of expertise level.

Overall, the value of incorporating procedural knowledge into a structural framework has value for user modeling in the areas of Human-Computer Interaction, task design, personnel selection and training, and task analysis. Within these areas, the procedural model can provide a better description of how humans structure task specific knowledge and how that structure affects performance. This includes predicting or explaining cognitive task behaviors. As a result, the use of such a model during software design will provide knowledge and interaction procedures which more closely match the user's knowledge of the objects and procedures in the computer environment. Therefore, such software and its associated interface will operate in a fashion more natural to the user and will be considered more "user friendly."

REFERENCES

Anderson, J. R. (1976). *Language, Memory, and Thought.* Hilldale, NJ: Lawrence Erlbaum Publishers.

Anderson, J. R. (1983). *The Architecture of Cognition.* Hilldale, NJ: Harvard University Press.

Anderson, J. R. & Bower, G. H. (1973). *Human Associative Memory.* New York: John Wiley and Sons.

Anderson, J. R., Kline, P. J., & Lewis, C. H. (1977). A production system model of language processing. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive Processes in Comprehension* (pp. 271–311). Hilldale, NJ: Lawrence Erlbaum Publishers.

Barr, A. & Feigenbaum, E. A. (Eds.). (1989). *The Handbook of Artificial Intelligence.* New York: Addison-Wesley.

Benysh, D. V. & Koubek, R. J. (1993). A comparative review of knowledge structure measurement techniques for interface design. *International Journal of Human–Computer Interaction, 5*(3), 211–238.

Bobrow, D. G. (1975). Dimensions of representation. In D. G. Bobrow & A. Collins (Eds.), *Representation and Understanding* (pp. 1–34). New York: Academic Press.

Bonnet, A. (1985). *Artificial Intelligence: promise and performance.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Bovair, S., Kieras, D. E., & Polson, P. G. (1990). The acquisition and performance of text-editing skill: a cognitive complexity analysis. *Human Computer Interaction, 5,* 1–48.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human–Computer Interaction.* Hilldale, NJ: Lawrence Erlbaum Publishers.

Clarkston, T. (1993). Towards a model of knowledge structure development with application to cognitive task performance in the manufacturing environment. Master's thesis, Purdue University, School of Industrial Engineering, West Lafayette, IN.

Davison, M. L. (1983). *Multidimensional Scaling.* New York: John Wiley and Sons.

Gardner, H. (1985). *The Mind's New Science: a history of the cognitive revolution.* New York: Basic Books Inc.

Hendrix, G. (1979). Encoding knowledge in partitioned networks. In N. Findler (Ed.), *Associative Networks.* New York: Academic Press.

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika, 32*(3), 241–253.

Kieras, D. E. (1988). Towards a practical goms model methodology for user interface design. In M. Helander (Ed.), *Handbook of Human–Computer Interaction* (pp. 135–157). North-Holland: Elsevier.

Kieras, D. E. & Elkerton, J. (1991). How to do a goms analysis for interface and documentation design. Tutorial in Computer–Human Interface 1991.

Koubek, R., Clarkston, T. P., & Calvez, V. (1993). The training of knowledge structures for manufacturing tasks: An empirical study. Ergonomics, At press.

Koubek, R. J. (1991). The impact of knowledge representation on cognitive-oriented task performance. *International Journal of Human-Computer Interaction*, *3*(1), 31–47.

Marchionini, G. & Sibert, J. (1991). An agenda for human-computer interaction: Science and engineering serving human needs. Technical report, The George Washington University. School of Engineering and Computer Science.

Neves, D. M. & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automation of cognitive skills. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 57–84). Hillsdale, NJ: Erlbaum.

Newell, A. (1992). Unified theories of cognition and the role of soar. In J. A. Michon & A. Akyurek (Eds.), *Soar: A cognitive architecture in perspective* (pp. 25–79). Boston: Kluwer Academic Publishers.

Newell, A. & Simon, H. A. (1972). *Human Problem Solving.* Englewood Cliffs, NJ: Prentice Hall Inc.

Olson, J. R. & Rueter, H. H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems*, *4*(3), 152–168.

Polson, P. G., Bovair, S., & Kieras, D. E. (1987). Transfer between text editors. In *Human Factors in Computing Systems and Graphics Interface Conference Proceedings*, (pp. 27–32).

Polson, P. G. & Lewis, C. H. (1990). Theory based design for easily learned interfaces. *Human-Computer Interaction*, *5*, 191–220.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic Information Processing* (pp. 227–270). Cambridge, Mass: MIT press.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*, 1–20.

Rodriguez, R. V. & Anger, F. D. (1991). Temporal reasoning: a relativistic model. *International Journal of Intelligent Systems*, *6*(3), 237–254.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to Multidimensional Scaling.* New York: Academic Press.

Schvaneveldt, R. W., Durso, F. T., & Goldsmith, T. E. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies*, *23*, 699–728.

Shiffren, R. M. & Shneider, W. (1977). Controlled and automatic human information processing: II. perceptual learning, automatic attending, and a general theory.

Shneider, W. & Shiffren, R. M. (1977). Controlled and automatic human information processing: I. detection, search, and attention. *Psychological Review*, (84), 1–66.

Singley, M. K. & Anderson, J. R. (1989). *The Transfer of Cognitive Skill.* Hilldale, NJ: Harvard University Press.

Solso, R. L. (1979). *Cognitive psychology.* Boston: Allyn and Bacon.

Tseng, M. M., Law, P.-L., & Cerva, T. (1992). Knowledge-based systems. In G. Salvendy (Ed.), *Handbook of Industrial Engineering* (pp. 184–210). New York: John Wiley and Sons.

Weizenbaum, J. (1976). *Computer Power and Human Reason*. San Francisco: Freeman and Company.

Winograd, T. (1975). Frame representations and the declarative/procedural controversy. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding : studies in cognitive science* (pp. 185–210). New York: Academic Press.

Table 1: Breakdown and organization of representation schemes.

Knowledge Organization Frameworks	Approaches of Cognitive Modeling	
	Exploratory	Functional
GWN	GOMS	Rule Based
MDS	NGOMSL	Logic Based
HCS	HAM / ACT	Object Oriented
Semantic Nets	GPS	Frames and Scripts
	TAG	OAV / OVP

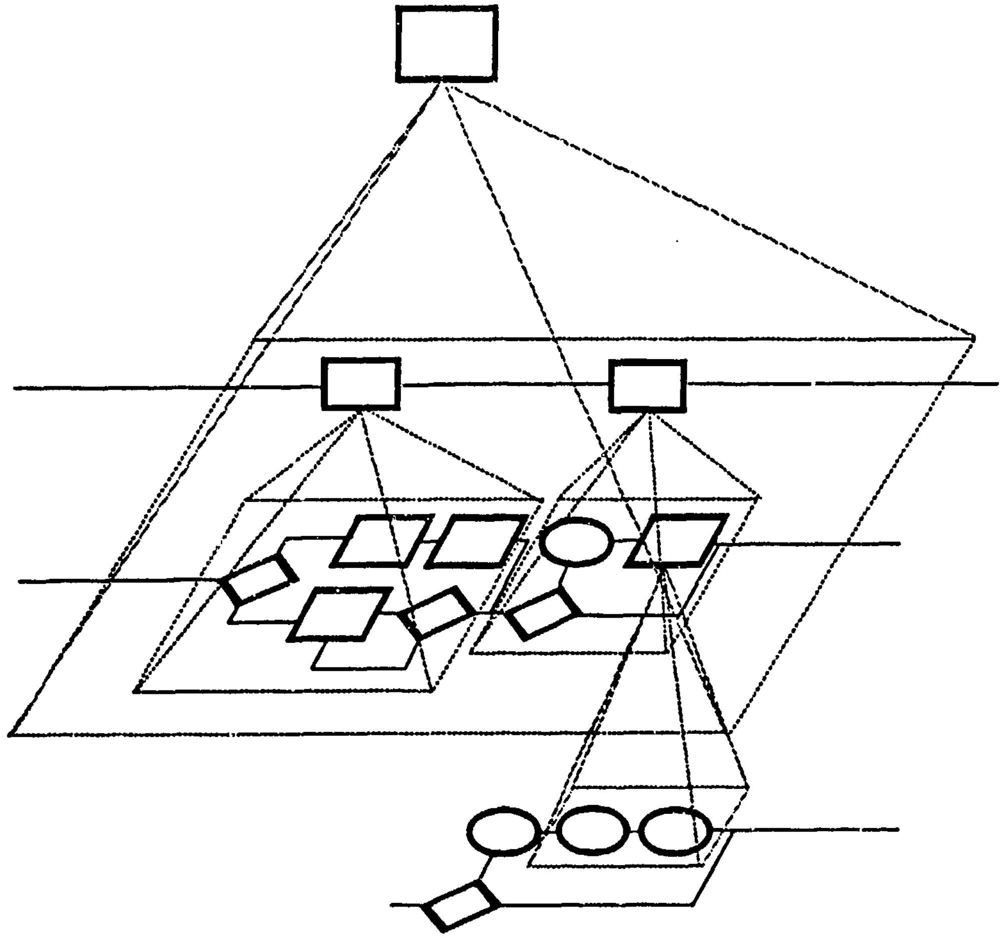
Table 2: t-test table.

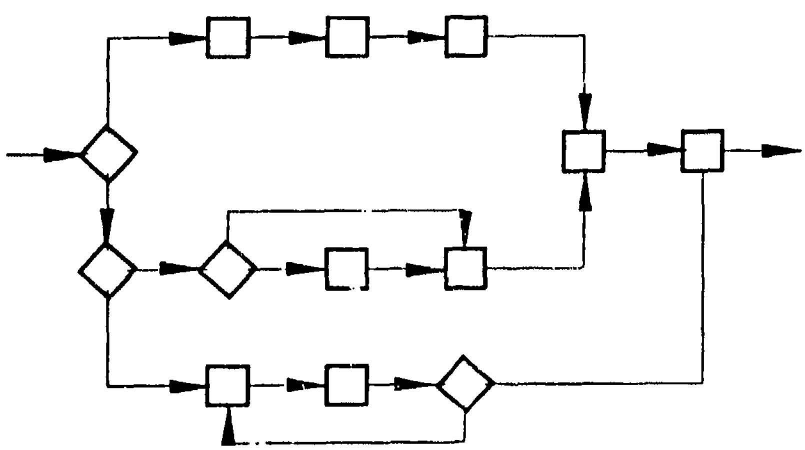
Variable	Novice $\overline{X}$	Novice SD	Experienced $\overline{X}$	Experienced SD	t-value	prob > $\|t\|$
Number of Hierarchy Links	37.556	3.455	37.000	18.251	0.0748	.9413
Number of Procedural Links	85.889	26.601	52.500	19.868	3.1215	.0062
Hierarchy/Procedural Links	.5013	.2828	.8818	.6410	1.639	.1195
Amount of Uncertainty	43.621	4.791	35.058	5.345	3.6605	.0019
Number of Decision Nodes	13.667	4.555	11.400	2.459	1.370	.1885
Multiple Relations	1.235	.4404	1.043	.0631	1.3675	.1893
Reliability	.8019	.04456	.8705	.0662	2.6184	.0180

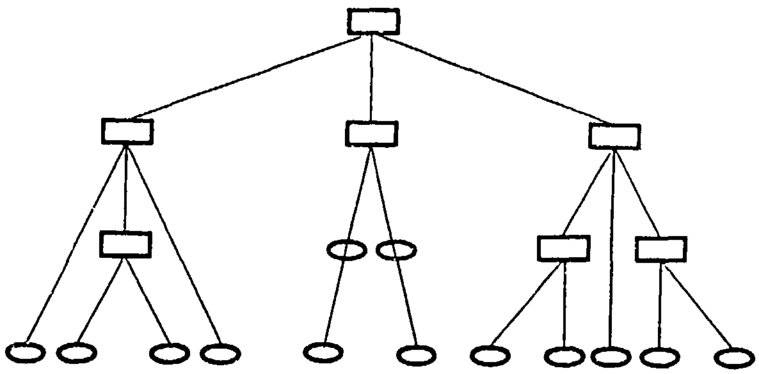
Table 3: Regression results summary.

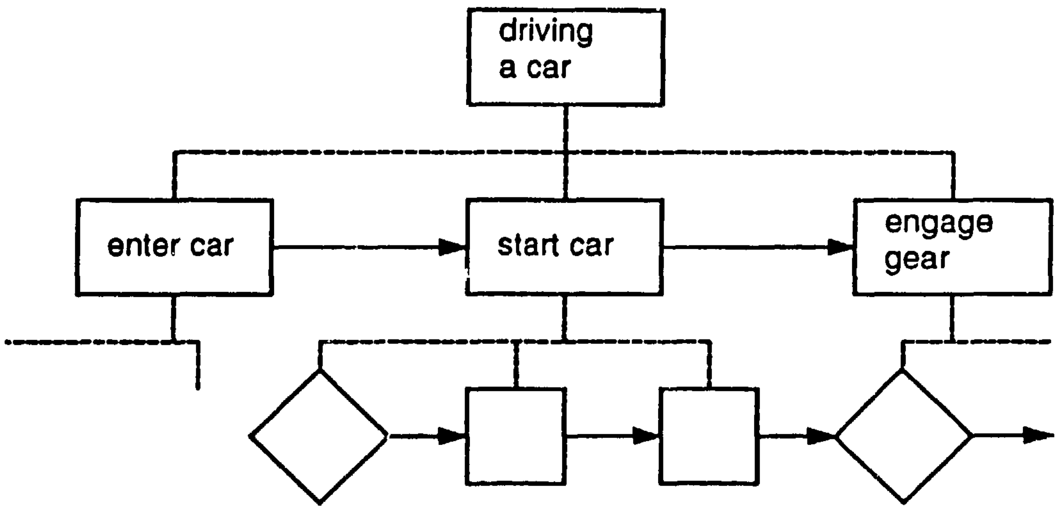
Dependent Variable	Procedural Alone R^2	Procedural Alone prob > $\|t\|$	Procedural Alone # vars	Declarative Alone R^2	Declarative Alone prob > $\|t\|$	Declarative Alone # vars	Combined R^2	Combined prob > $\|t\|$	Combined # vars
Error	.3947	.0180	2	.3675	.0693	3	.5105	.0308	4
Time	.1957	.0579	1	.2472	.0303	1	.4789	.0179	3
Error Rate	.6479	.0002	2	.4583	.0074	2	.7310	.0021	5
Group	.4408	.0019	1	.5434	.0070	3	.6117	.0070	4

List of Figures:

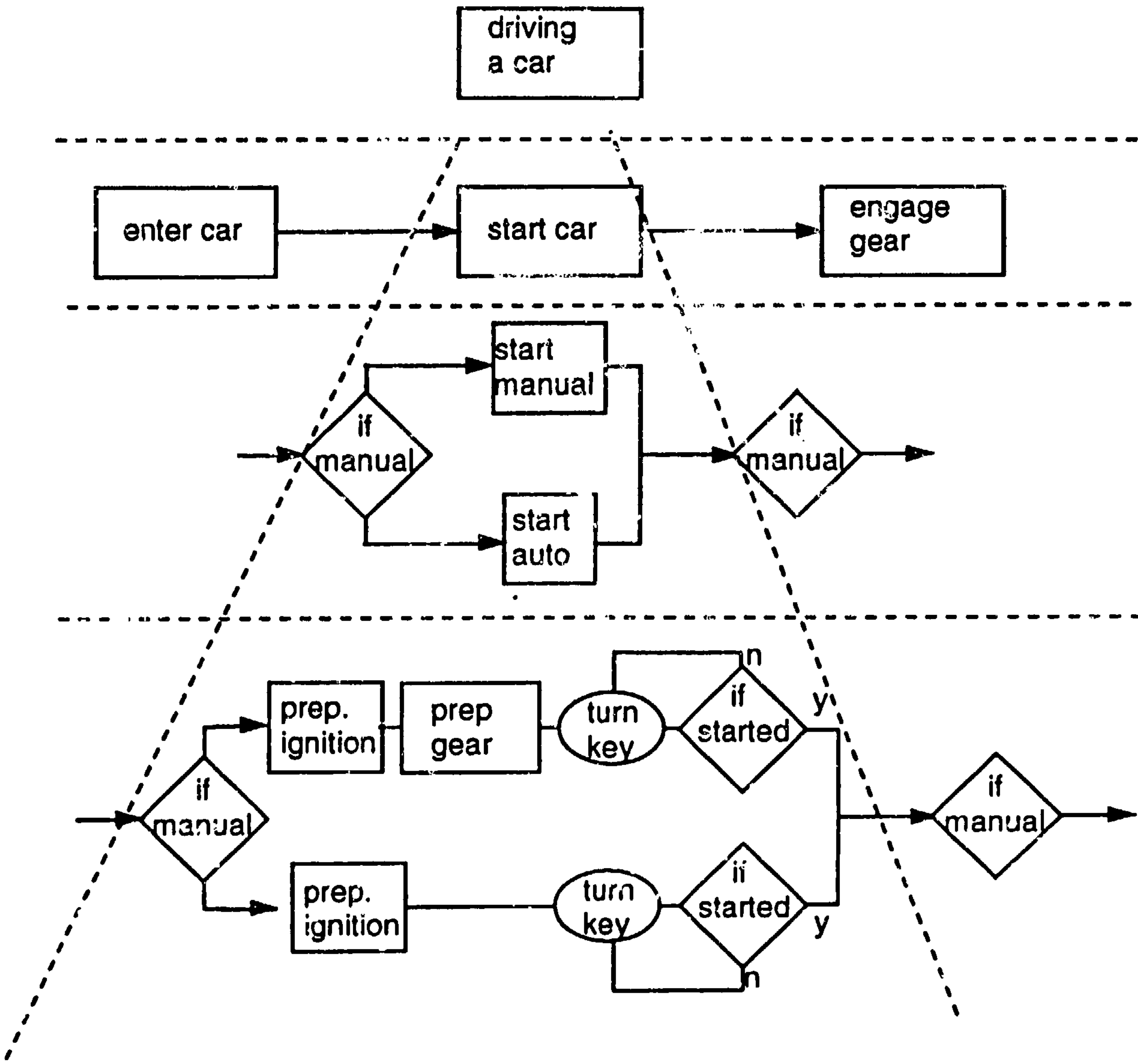


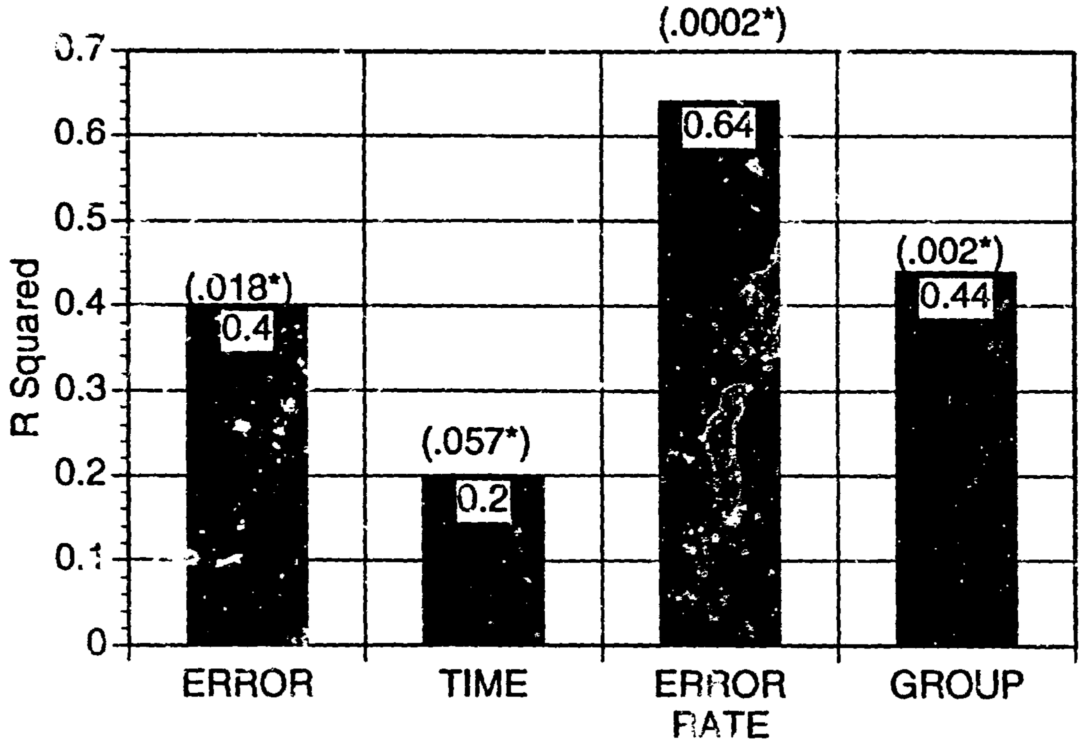




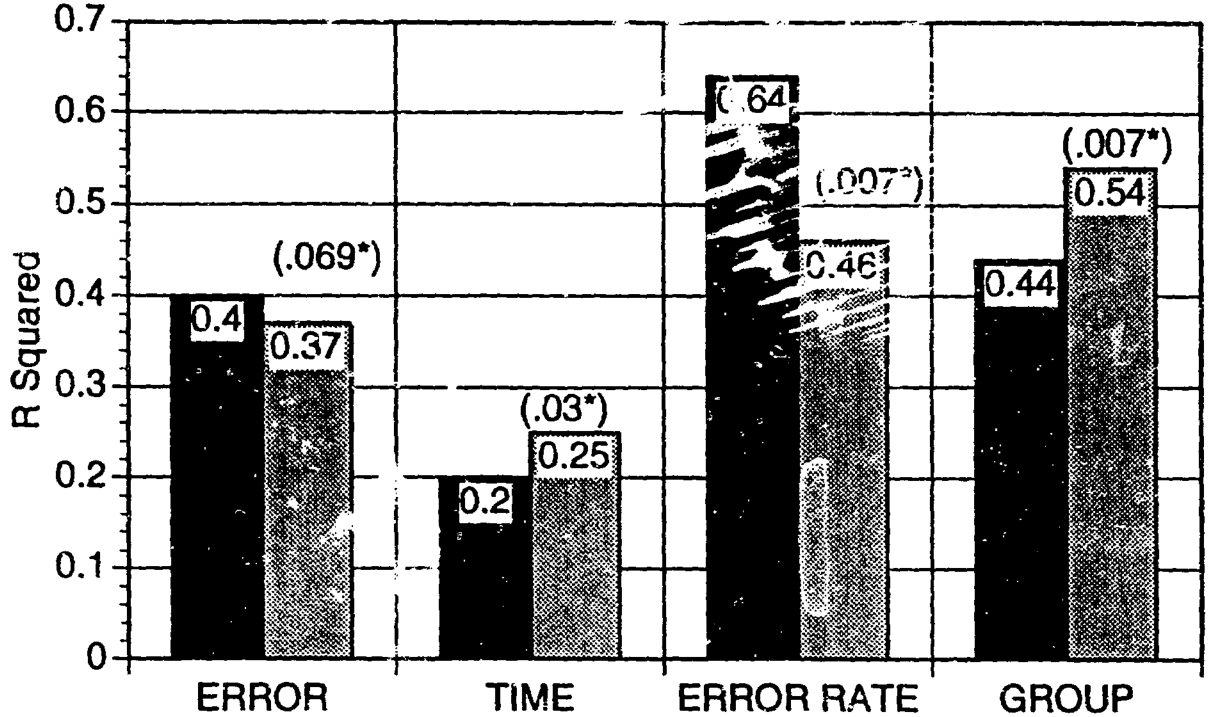


driving a car
enter car
start car
engage gear

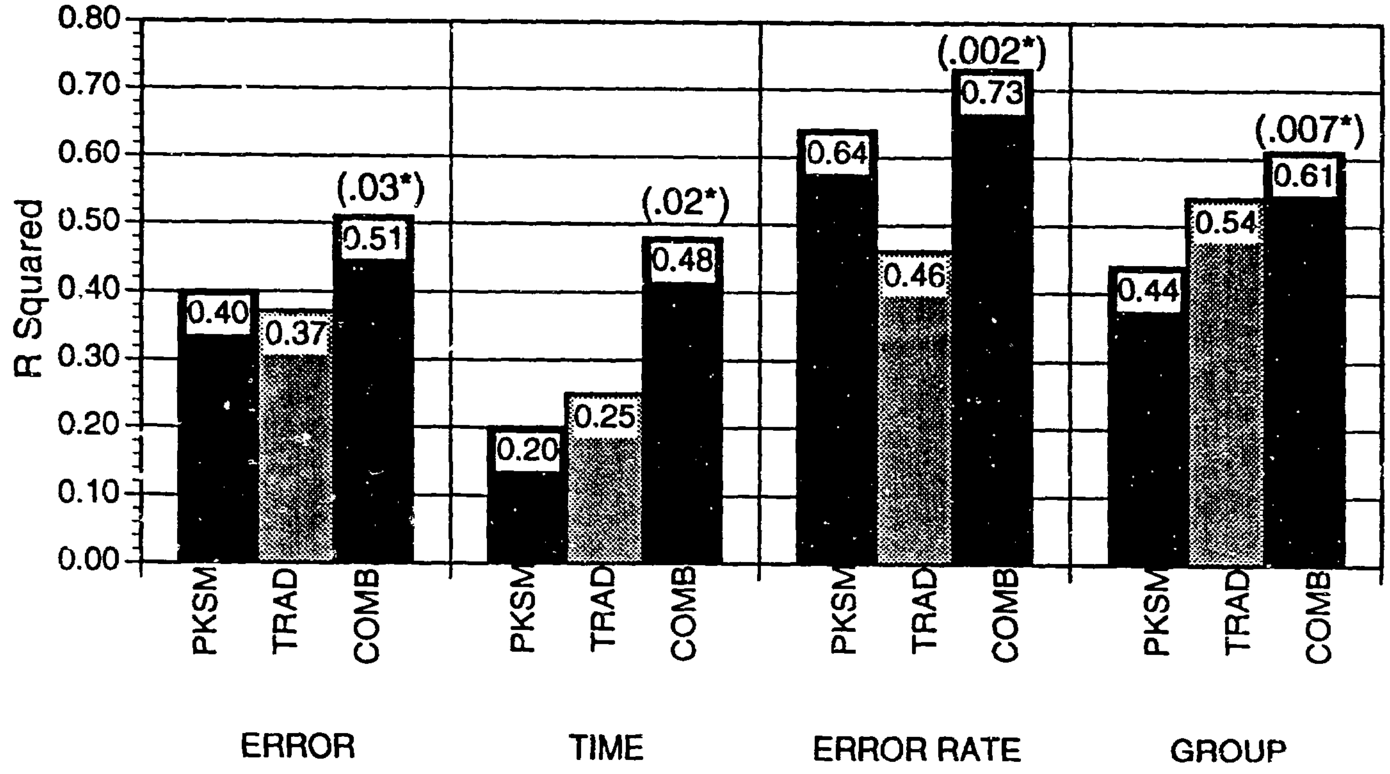

driving a car
enter car
start car
engage gear
start manual
if manual
start auto
if manual
prep. ignition
prep gear
turn key
if started
n
y
if manual
if manual
prep. ignition
turn key
if started
y
n



* denotes the level of significance



* denotes the level of significance



* denotes the level of significance

Distribution List

Dr. David L. Alderton, Code 131
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Terry Allard
Code 1142CS
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217-5000

Dr. John R. Anderson
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Nancy S. Anderson
Department of Psychology
University of Maryland
College Park, MD 20742

Dr. Thomas H. Anderson
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champaign, IL 61820

Dr. Stephen J. Andriole, Chairman
College of Information Studies
Drexel University
Philadelphia, PA 19104

Prof. John Annett
University of Warwick
Department of Psychology
Coventry CV4 7AL
ENGLAND

Dr. Gregory Anrig
Educational Testing Service
Princeton, NJ 08541

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Edward Atkins
13705 Lakewood Ct.
Rockville, MD 20850

Dr. Patricia Baggett
School of Education
610 E. University, Rm 1302D
University of Michigan
Ann Arbor, MI 48109-1259

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Leo Beltracchi
United States Nuclear
Regulatory Commission
Washington DC 20555

Dr. William O. Berry
Director of Life and
Environmental Sciences
AFOSR/NL, N1, Bldg. 410
Bolling AFB, DC 20332-6448

Dr. [illegible] G. [illegible]
Department of Psychology
University of Rochester
River Station
Rochester, NY [illegible]

Dr. Manuel Bernabeu
Educational Testing
Service
Princeton, NJ [illegible]

Dr. Werner P. Birke
Personalstammamt der Bundeswehr
Kolner Strasse 262
D-5000 Koeln 90
FEDERAL REPUBLIC OF GERMANY

Dr. John Black
Teachers College, Box 8
Columbia University
525 West 120th Street
New York, NY 10027

Dr. Deborah A. Boehm-Davis
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030-4444

Dr. Kenneth R. Boff
AL/CFH
Wright-Patterson AFB
OH 45433-6573

Dr. J. C. Boudreaux
Manufacturing Engineering Lab
National Institute of
Standards and Technology
Gaithersburg, MD 20899

Dr. Richard L. Branch
HQ, USMEPCOM/MEPCT
2500 Green Bay Road
North Chicago, IL 60064

Dr. Robert Breaux
Code 252
Naval Training Systems Center
Orlando, FL 32826-3224

Dr. Ann Brown
Graduate School of Education
University of California
EMST-4533 Tolman Hall
Berkeley, CA 94720

Dr. Pat Carpenter
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

CDR Robert Carter
HCI Laboratory
Naval Research Laboratory
Washington, DC 20390

Dr. Eduardo Cascallar
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Dr. Davida Charney
English Department
Penn State University
University Park, PA 16802

Dr. Michelene Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Susan Chipman
Cognitive Science Program
Office of Naval Research
800 North Quincy Street
Code 342CS
Arlington, VA 22217-5660

Dr. Raymond E. Christal
UES LAMP Science Advisor
AL/HRMIL
Brooks AFB, TX 78235

Dr. Deborah Claman
National Institute for Aging
Bldg. 31, Room 5C-35
9000 Rockville Pike
Bethesda, MD 20892

Dr. William J. Clancey
Institute for Research
on Learning
2550 Hanover Street
Palo Alto, CA 94304

Chief of Naval Education and
Training (N-5)
NAS Pensacola, FL 32508

Dr. Paul Cobb
Purdue University
Education Building
W. Lafayette, IN 47907

Dr. Rodney Cocking
NIMH, Basic Behavior and
Cognitive Science Research
5600 Fishers Lane, Rm 11C-10
Parklawn Building
Rockville, MD 20857

Director, Life Sciences
Office of Naval Research
Code 34 (Code 114)
Arlington, VA 22217-5000

Director, Cognitive and
Neural Sciences, Code 1142
Office of Naval Research
Arlington, VA 22217-5000

Director
Training Systems Department
Code 15A
Navy Personnel R&D Center
San Diego, CA 92152-6800

Library, Code 231
Navy Personnel R&D Center
San Diego, CA 92152-5800

Dr. Magda Colberg
Office of Personnel
Management
1900 E Street, N.W.
Washington, DC 20415-0001

Commanding Officer
Naval Research Laboratory
Code 4827
Washington, DC 20375-5000

Dr. Albert T. Corbett
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. Michael Cowen
Code 142
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Kenneth B. Cross
Anacapa Sciences, Inc.
P.O. Box 519
Santa Barbara, CA 93102

CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940-5380

Dr. Charles E. Davis
Educational Testing Service
Mail Stop 22-T
Princeton, NJ 08541

Dr. Geory Delacote
Exploratorium
3601 Lyon Street
San Francisco, CA
94123

Dr. Sharon Derry
Florida State University
Department of Psychology
Tallahassee, FL 32306

Dr. Stephanie Doane
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Michael Drillings
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Defense Technical
Information Center
DTIC/DDA-2
Cameron Station, Bldg 5
Alexandria, VA 22314
(4 Copies)

Mr. David DuBois
Personnel Decisions Research
Institutes
43 Main Street, SE
Riverplace, Suite 405
Minneapolis, MN 55414

Dr. Richard Duran
Graduate School of Education
University of California
Santa Barbara, CA 93106

Dr. Nancy Eldredge
College of Education
Division of Special Education
The University of Arizona
Tucson, AZ 85721

Dr. John Ellis
Navy Personnel R&D Center
Code 15
San Diego, CA 92152-6800

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. Susan Epstein
144 S. Mountain Avenue
Montclair, NJ 07042

ERIC Facility-Acquisitions
1301 Piccard Drive, Suite 300
Rockville, MD 20850-4305

Dr. K. Anders Ericsson
University of Colorado
Department of Psychology
Campus Box 345
Boulder, CO 80309-0345

Dr. Martha Evens
Dept. of Computer Science
Illinois Institute of Technology
10 West 31st Street
Chicago, IL 60616

Dr. Lorraine D. Eyde
US Office of Personnel Management
Office of Personnel Research and
Development
1900 E St., NW
Washington, DC 20415

Dr. Franco Faina
Direttore Generale LEVADIFE
Piazzale K. Adenauer, 3
00144 ROMA EUR
ITALY

Dr. Jean-Claude Falmagne
Irvine Research Unit in
Mathematical & Behavioral Sciences
University of California
Irvine, CA 92717

Dr. Beatrice J. Farr
Army Research Institute
PERI-IC
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Marshall J. Farr
Farr-Sight Co.
2520 North Vernon Street
Arlington, VA 22207

Dr. Pat-Anthony Federico
Code 13, NPRDC
53335 Ryne Road
San Diego, CA 92152-7250

Dr. J. D. Fletcher
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311

Dr. Lawrence T. Frase
Executive Director
Division of Cognitive and
Instructional Science
Educational Testing Service
Princeton, NJ 08541

Dr. Carl H. Frederiksen
Dept. of Educational Psychology
McGill University
3700 McTavish Street
Montreal, Quebec
CANADA H3A 1Y2

Dr. Norman Frederiksen
Educational Testing Service
(05-R)
Princeton, NJ 08541

Dr. John Tangney
AFOSR/NL Bldg 410
110 Duncan Avenue
Washington, DC 20332-0001

Dr. John D. E. Gabrieli
Department of Psychology
Jordan Hall
Stanford University
Stanford, CA 94305

Dr. Jack J. Gelfand
Department of Psychology
Princeton University
Princeton, NJ 08544-1010

Chair, Department of
Computer Science
George Mason University
Fairfax VA 22030

Dr. Alan S. Gevins
EEG Systems Laboratory
51 Federal Street, Suite 401
San Francisco, CA 94107

Dr. Helen Gigley
Naval Research Lab., Code 5530
4555 Overlook Avenue, S. W.
Washington, DC 20375-5000

Dr. Herbert Ginsburg
Box 184
Teachers College
Columbia University
525 West 121st Street
New York, NY 10027

Dr. Drew Gitomer
Educational Testing Service
Princeton, NJ 08541

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Sam Glucksberg
Department of Psychology
Princeton University
Princeton, NJ 08544-1010

Prof. Joseph Goguen
PRG, Univ. of Oxford
11 Keble Road
Oxford OX13QD
UNITED KINGDOM

Dr. Paul E. Gold
University of Virginia
Department of Psychology
Charlottesville, VA 22903

Dr. Susan R. Goldman
Peabody College, Box 45
Vanderbilt University
Nashville, TN 37203

Dr. Timothy Goldsmith
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Dr. Sherrie Gott
AFHRL/MOMJ
Brooks AFB, TX 78235-5601

Dr. T. Govindaraj
Georgia Institute of
Technology
School of Industrial
and Systems Engineering
Atlanta, GA 30332-0205

Dr. Marilyn K. Gowing
Office of Personnel R&D
1900 E St., NW, Room 6462
Office of Personnel Management
Washington, DC 20415

Jordan Grafman, Ph.D.
Chief, Cognitive Neuroscience
Section, Medical Neurology
Branch-NINDS, Bldg 10, Rm. 5C422
Bethesda, MD 20892

Dr. Wayne Gray
Graduate School of Education
Fordham University
113 West 60th Street
New York, NY 10023

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. James G. Greeno
School of Education
Stanford University
Room 311
Stanford, CA 94305

Dr. Stephen Grossberg
Center for Adaptive Systems
Room 244
111 Cummington Street
Boston University
Boston, MA 02215

Prof. Edward Haertel
School of Education
Stanford University
Stanford, CA 94305-3096

Marilyn Halpern, Librarian
Brigham Library
Educational Testing Service
Carter and Rosedale Roads
Princeton, NJ 08541

Dr. Stephen J. Hanson
Learning & Knowledge
Acquisition Research
Siemens Research Center
755 College Road East
Princeton, NJ 08540

Dr. Delwyn Harnisch
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Patrick R. Harrison
Computer Science Department
U.S. Naval Academy
Annapolis, MD 21402-5002

Dr. Wayne Harvey
Education Development Center
55 Chapel Street
Newton, MA 02160

Dr. Barbara Hayes-Roth
Knowledge Systems Laboratory
Stanford University
701 Welch Road, Bldg. C
Palo Alto, CA 94304

Dr. Per Helmersen
University of Oslo
USIT
Box 1059
0316 Oslo, NORWAY

Dr. G. Gunn Hendel Van
AFD SWO
Admiraliteit Kr. D
Van Der Burchlaan 31
Post Box 20702,2500 ES The Hague
The NETHERLANDS

Prof. Lutz F. Hornke
Institut fur Psychologie
RWTH Aachen
Jaegerstrasse 17/19
D-5100 Aachen
WEST GERMANY

Ms. Julia S. Hough
Cambridge University Press
40 West 20th Street
New York, NY 10011

Dr. William Howell
Chief Scientist
AFHRL/CA
Brooks AFB, TX 78235-5601

Dr. Eva Hudlicka
BBN Laboratories
10 Moulton Street
Cambridge, MA 02238

Dr. Earl Hunt
Dept. of Psychology, NI-25
University of Washington
Seattle, WA 98195

Dr. Jack Hunter
2122 Coolidge Street
Lansing, MI 48906

Dr. Giorgio Ingargiola
Computer Science Department
Temple University
Philadelphia, PA 19122

Dr. Martin J. Ippel
Center for the Study of
Education and Instruction
Leiden University
P. O. Box 9555
2300 RB Leiden
THE NETHERLANDS

Dr. Robert Jannarone
Elec. and Computer Eng. Dept.
University of South Carolina
Columbia, SC 29208

Dr. Claude Janvier
CIRADE, X-7120
UQAM
P. O. Box 8888, Succ. A
Montreal, Quebec H3C 3P8
CANADA

Dr. Robin Jeffries
Hewlett-Packard
Laboratories, 1U-17
P.O. Box 10490
Palo Alto, CA 94303-0969

Dr. Edgar M. Johnson
Technical Director
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Peder Johnson
Department of Psychology
University of New Mexico
Albuquerque, NM 87131

Dr. Daniel B. Jones
US Nuclear Regulatory
Commission
NRR/12 E4
Washington, DC 20555

Dr. John Jonides
Department of Psychology
University of Michigan
Ann Arbor, MI 48104

Dr. Marcel Just
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Ruth Kanfer
University of Minnesota
Department of Psychology
Elliott Hall
75 E. River Road
Minneapolis, MN 55455

Dr. Michael Kaplan
Office of Basic Research
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Wendy Kellogg
IBM T. J. Watson
Research Ctr.
P.O. Box 704
Yorktown Heights, NY 10598

Dr. J.A.S. Kelso
Center for Complex Systems
Building MT 9
Florida Atlantic University
Boca Raton, FL 33431

CDR Robert S. Kennedy
Essex Corporation
1040 Woodcock Road
Suite 227
Orlando, FL 32803

Dr. David Kieras
Technical Communication Program
TIDAL Bldg., 2360 Bonisteel Blvd.
University of Michigan
Ann Arbor, MI 48109-2108

Dr. Sung-Ho Kim
Educational Testing Service
Princeton, NJ 08541

Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80309-0345

Dr. Susan S. Kirschenbaum
Code 2212, Building 1171/1
Naval Underwater Systems Center
Newport, RI 02841

Dr. Janet L. Kolodner
Georgia Institute of Technology
College of Computing
Atlanta GA 30332-0280

Dr. Sylvan Kornblum
University of Michigan
Mental Health Research Institute
205 Washtenaw Place
Ann Arbor, MI 48109

Dr. Stephen Kosslyn
Harvard University
1236 William James Hall
33 Kirkland St.
Cambridge, MA 02138

Dr. Kenneth Kotovsky
Department of Psychology
Carnegie-Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Dr. Richard J. Koubek
School of Industrial
Engineering
Grissom Hall
Purdue University
West Lafayette, IN 47907

Dr. Art Kramer
Univ. of Illinois at U-C
Beckman Institute
405 N. Mathews Avenue
Urbana, IL 61801

Dr. Benjamin Kuipers
University of Texas at Austin
Department of Computer Sciences
Austin, Texas 78712

Dr. Michael Kuperstein
Symbus Technology
325 Harvard Street
Suite 211
Brookline, MA 02146

Dr. Patrick Kyllonen
AFHRL/MOEL
Brooks AFB, TX 78235

Dr. M. Diane Langston
ICL North America
11490 Commerce Park Drive
Reston, VA 22091

Dr. Marcy Lansman
University of North Carolina
Dept. of Computer Science
CB #3175
Chapel Hill, NC 27599

Richard Lanterman
Commandant (G-PWP)
US Coast Guard
2100 Second St., SW
Washington, DC 20593-0001

Dr. Robert W. Lawler
Matthews 118
Purdue University
West Lafayette, IN 47907

Dr. Paul E. Lehner
Department of Information
Systems & Engineering
George Mason University
4400 University Drive
Fairfax, VA 22030-4444

Dr. Richard Lesh
Educational Testing Service
Princeton, NJ 08541

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
1310 South Sixth Street
University of IL at
Urbana-Champaign
Champaign, IL 61820-6990

Logicon Inc. (Attn: Library)
Tactical and Training Systems
Division
P.O. Box 85158
San Diego, CA 92138-5158

Prof. David F. Lohman
College of Education
University of Iowa
Iowa City, IA
52242

Vern M. Malec
NPRDC, Code 142
San Diego, CA 92152-6800

Dr. Sandra P. Marshall
Dept. of Psychology
San Diego State University
San Diego, CA 92182

Dr. Clessen J. Martin
Head, DoD Coordinator,
Recruiting Plans and Programs
Branch Code: PERS 2FF/234
Navy Annex, Room 2832
Washington, DC 20350

Dr. Elizabeth Martin
AL/HRA, Stop 44
Williams AFB
AZ 85240

Dr. Nadine Martin
Department of Neurology
Center for Cognitive Neuroscience
Temple University School of Medicine
3401 North Broad Street
Philadelphia, PA 19140

Dr. Joseph McLachlan
Navy Personnel Research
and Development Center
Code 14
San Diego, CA 92152-6800

Dr. Michael McNeese
DET-1, AL/CFHI
BLDG 248
Wright-Patterson AFB, OH 45432

Dr. Alan Meyrowitz
Naval Research Laboratory
Code 5510
4555 Overlook Ave., SW
Washington, DC 20375-5000

Dr. Ryszard S. Michalski
Center for Artificial Intelligence
George Mason University
Science and Tech II, Rm.411
4400 University Drive
Fairfax, VA 22030-4444

Dr. Vittorio Midoro
CNR-Istituto Tecnologie Didattiche
Via All'Opera Pia 11
GENOVA-ITALIA 16145

Dr. Mortimer Mishkin
Laboratory of Neuropsychology
National Institute of Mental Health
9000 Rockville Pike, Bldg.9, #1N107
Bethesda, MD 20892

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. Andrew R. Molnar
Applications of Advanced
Technologies, Rm. 635
National Science Foundation
Washington, DC 20550

Dr. Ben B. Morgan, Jr.
Department of Psychology
University of Central Florida
Orlando, FL 32816-0150

Dr. Randy Mumaw
Human Sciences
Westinghouse Science
& Technology Ctr.
1310 Beulah Road
Pittsburgh, PA 15235

Dr. Allen Munro
Behavioral Technology
Laboratories - USC
250 N. Harbor Dr., Suite 309
Redondo Beach, CA 90277

Academic Progs. & Research Branch
Naval Technical Training Command
Code N-62
NAS Memphis (75)
Millington, TN 30854

Deputy Director Manpower,
Personnel and Training Div.
Naval Sea Systems Command
ATTN: Code 04MP 511
Washington, DC 20362

Mr. J. Nelissen
Twente University
Fac. Bibl. Toegepaste Onderwyskunde
P. O. Box 217
7500 AE Enschede
The NETHERLANDS

Director, Instr. Devel. and
Educ. Prog. Spt. Dept.
Naval Educ. & Tng. Prog.
Mgt. Spt. Activity (NETPMSA)
Pensacola, FL 32509

Dr. Raymond S. Nickerson
5 Gleason Road
Bedford, MA 01730

Dr. Donald A. Norman
Department of Cognitive Science
University of California
La Jolla, CA 92093-0515

Director, Fleet Liaison Office
NPRDC (Code 01F)
San Diego, CA 92152-6800

Director
Training Systems Department
NPRDC (Code 14)
San Diego, CA 92152-6800

Library, NPRDC
Code 041
San Diego, CA 92152-6800

Office of Naval Research
Resident Representative
The Ohio State University
Research Center
1960 Kenny Rd., Room 113
Columbus, OH 43210-1063

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 600
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. Stellan Ohlsson
Learning R & D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Judith Reitman Olson
Graduate School of Business
University of Michigan
Ann Arbor, MI 48109-1234

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Assistant for Training Technology
and Human Factors
Office of the DCNO(MPT) (Op-11B)
Department of the Navy
Washington, DC 20350-2000

Dr. Judith Orasanu
Mail Stop 239-1
NASA Ames Research Center
Moffett Field, CA 94035

Dr. Everett Palmer
Mail Stop 262-4
NASA-Ames Research Center
Moffett Field, CA 94035

Dr. Okchoon Park
Army Research Institute
PERI-2
5001 Eisenhower Avenue
Alexandria, VA 22333

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. Roy Pea
Institute for the
Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, IL 60201

G. Pelsmakers
Rue Fritz Toussaint 47
Gendarmerie RSP
1050 Bruxelles
BELGIUM

Dr. Ray S. Perez
ARI (PERI-II)
5001 Eisenhower Avenue
Alexandria, VA 22333

C.V. (MD) Dr. Antonio Peri
Captain ITNMC
Maripers U.D.G. 3' Sez
MINISTERO DIFESA - MARINA
00100 ROMA - ITALY

Dr. Nancy N. Perry
Naval Education and Training
Program Support Activity
Code-047
Building 2435
Pensacola, FL 32509-5000

CDR Frank C. Petho
Naval Postgraduate
School
Code OR/PE
Monterey, CA 93943

Dept. of Administrative Sciences
Code 54
Naval Postgraduate School
Monterey, CA 93943-5026

Dr. Elizabeth A. Phelps
Department of Psychology
New School for Social Research
65 Fifth Avenue
New York, NY 10003

Dr. Peter Pirolli
School of Education
University of California
Berkeley, CA 94720

Dr. Martha Polson
Department of Psychology
University of Colorado
Boulder, CO 80309-0344

Dr. Peter Polson
University of Colorado
Department of Psychology
Boulder, CO 80309-0344

Dr. Joseph Psotka
ATTN: PERI-IC
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333-5600

Psyc Info - CD and M
American Psychological Assoc.
1200 Uhle Street
Arlington, VA 22201

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. J. Wesley Regian
Armstrong Laboratory
AFHRL/IDI
Brooks AFB, TX 78235-5000

Dr. Fred Reif
CDEC, Smith Hall
Carnegie-Mellon University
Pittsburg, PA 15213

Dr. Charles M. Reigeluth
Chairman, Instructional
Systems Technology
School of Education, Rm. 210
Indiana University
Bloomington, IN 47405

Dr. Daniel Reisberg
Reed College
Department of Psychology
Portland, OR 97202

Dr. Brian Reiser
Institute for the Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, IL 60201-3142

Dr. Lauren Resnick
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. Gilbert Ricard
Mail Stop K01-14
Grumman Aircraft Systems
Bethpage, NY 11714

Mr. W. A. Rizzo
Head, Human Factors Division
Naval Training Systems Center
Code 26
12350 Research Parkway
Orlando, FL 32826-3224

Dr. Linda G. Roberts
Science, Education, and
Transportation Program
Office of Technology Assessment
Congress of the United States
Washington, DC 20510

Dr. Salim Roukos
IBM Corporation
T. J. Watson Research Center
PO Box 218
Yorktown Heights, NY 10598

Mr. Louis Roussos
University of Illinois
Department of Statistics
101 Illini Hall
725 South Wright St.
Champaign, IL 61820

Dr. Eduardo Salas
Human Factors Division (Code 262)
12350 Research Parkway
Naval Training Systems Center
Orlando, FL 32826-3224

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
310B Austin Peay Bldg.
Knoxville, TN 37966-0900

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

Capitan Jesus Bernal Santos
Ministerio de Defensa.
Unidad de Psicologia (SEGENTE)
pº de la Castellana, 109
28071, Madrid
ESPANA

Dr. Daniel L. Schacter
Department of Psychology
Harvard University
Cambridge, MA 02138

Dr. Walter Schneider
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Mary Schratz
4100 Parkside
Carlsbad, CA 92008

Dr. Myrna F. Schwartz
Director
Neuropsychology Research Lab
Moss Rehabilitation Hospital
1200 West Tabor Road
Philadelphia, PA 19141

Dr. Robert J. Seidel
US Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Terrence J. Sejnowski
Professor
The Salk Institute
P. O. Box 85800
San Diego, CA 92138-9216

Dr. Michael G. Shafto
NASA Ames Research Ctr.
Mail Stop 262-1
Moffett Field, CA 94035-1000

Dr. Valerie L. Shalin
Department of Industrial
Engineering
State University of New York
342 Lawrence D. Bell Hall
Buffalo, NY 14260

Mr. Richard J. Shavelson
Graduate School of Education
University of California
Santa Barbara, CA 93106

Dr. Tracy Shors
Dept. of Psychology
Princeton Univ.
Green Hall
Princeton, NJ 08544

Dr. Randall Shumaker
Naval Research Laboratory
Code 5500
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Edward Silver
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Zita M. Simutis
Director, Manpower & Personnel
Research Laboratory
US Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Jerome E. Singer
Department of Medical Psychology
Uniformed Services Univ. of the
Health Sciences
4301 Jones Bridge Road
Bethesda, MD 20814-4799

Dr. Jan Sinnott
Department of
Computer Science
Towson State University
Towson, MD 21204

Dr. Derek Sleeman
Computing Science Department
The University
Aberdeen AB9 2FX
Scotland
UNITED KINGDOM

Dr. Robert Smillie
Naval Ocean Systems Center
Code 443
San Diego, CA 92152-5000

Dr. Richard E. Snow
School of Education
Stanford University
Stanford, CA 94305

Dr. James J. Staszewski
Dept. of Psychology
University of South Carolina
Columbia, SC 29210

Dr. Bruce D. Steinberg
Curry College
Milton, MA 02186

Dr. Frederick Steinheiser
CIA-ORD
Ames Building
Washington, DC 20505

Dr. Ted Steinke
Dept. of Geography
University of South Carolina
Columbia, SC 29208

Dr. Kurt Steuck
AL/HRTI
Brooks AFB
San Antonio, TX 78235-5601

Dr. Michael J. Tarr
Dept. of Psychology
Yale University
PO Box 11A, Yale Station
New Haven, CT 06520

Dr. Kikumi Tatsuoka
Educational Testing Service
Mail Stop 03-T
Princeton, NJ 08541

Dr. Gary Thomasson
Defense Manpower Data
Center
99 Pacific Street
Suite 155A
Monterey, CA 93940

Dr. Sharon Tkacz
CAE-Link Corporation
209 Madison Street
Alexandria, VA 22314

Dr. Douglas Towne
Behavioral Technology Labs
University of Southern California
1228 Spring Street
St. Helena, CA 94574

Dean, College of Behavioral
and Social Sciences
University of Maryland,
Baltimore County
Baltimore, MD 21228

Chair, Department of Psychology
University of Maryland,
Baltimore County
Baltimore, MD 21228

Dr. Kurt VanLehn
Learning Research
& Development Ctr.
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Frank L. Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Jerry Vogt
Department of Psychology
St. Norbert College
De Pere, WI 54115-2099

Dr. Jacques Voneche
University of Geneva
Department of Psychology
Geneva
SWITZERLAND 1204

Dr. Douglas Wetzel
Code 15
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Barbara White
School of Education
Tolman Hall, EMST
University of California
Berkeley, CA 94720

Alexandra K. Wigdor
NRC MH-176
2101 Constitution Ave.
Washington, DC 20418

Dr. David Wiley
School of Education
and Social Policy
Northwestern University
Evanston, IL 60208

Dr. David C. Wilkins
University of Illinois
Department of Computer Science
405 North Mathews Avenue
Urbana, IL 61801

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Kent E. Williams
Virginia Polytechnic Institute
and State University
Department of Industrial and
Systems Engineering/MSL
1900 Kraft Drive
Blacksburg, VA 24060

Dr. Mark Wilson
Graduate School of
Education
University of California,
Berkeley
Berkeley, CA 94720

Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Merlin C. Wittrock
Graduate School of Education
Univ. of Calif., Los Angeles
Los Angeles, CA 90024

Dr. Kentaro Yamamoto
03-0T
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Dr. Masoud Yazdani
Dept. of Computer Science
University of Exeter
Prince of Wales Road
Exeter EX44PT
ENGLAND

Frank R. Yekovich
Dept. of Education
Catholic University
Washington, DC 20064

Dr. Joseph L. Young
National Science Foundation
Room 320
1800 G Street, N.W.
Washington, DC 20550

Mr. Anthony R. Zara
National Council of State
Boards of Nursing, Inc.
676 N. St. Clair
Suite 550
Chicago, IL 60611

Prof. Gerard de Zeeuw
Center for Innovation
and Cooperative Technology
Grote Bickersstraat 72
1013 KS Amsterdam
The NETHERLANDS