(12) SC LEVEL III

This research was sponsored by the
Defense Advanced Research Projects
Agency under ARPA Order no.: 3699
Contract No.: ~~MDA903-70-C-0912~~ 79-C-0192
Monitored by Dr. J. D. Fletcher

(15) MDA903-79-C-0192, ARPA Order-3699

(12) 60

(9) Technical rept.,

(6) A STRUCTURAL APPROACH TO THE VALIDATION OF HIERARCHICAL TRAINING SEQUENCES.

TASK 2 TECHNICAL REPORT

(14) TR-2

(11) 1979

(10)

John R. Bergan, Principal Investigator
Anthony A. Cancelli, Co-Principal Investigator
Olga Towstopiat, Research Associate

033800

12

Source: Arizona Univ., Tucson. Jim Cundiff per telecom to Ms. Alma Springs, DARPA. slt

5-9-80

Technical Report Number 2

Report Summary

This report ~~is aimed at~~ describes developing the technology necessary to conduct cost effective and efficient validations of the sequencing of instruction used in the training of military occupational specialties. The specific objective ~~covered by this technical report~~ was to validate the hierarchical ordering of task domains. A total of 317 subjects were tested on two algebra skill domains, constructed from the Precision Measuring Equipment Curriculum of the Air Force Advanced Instructional System.

Latent structure techniques recently ~~developed by Leo Goodman at the University~~ of Chicago were used to validate the hypothesized ordering between domains. The first step in the analysis was to construct a set of models representing hypotheses about the tasks under examination. The models developed for use in the present analysis assumed three basic classes of individuals for tasks in an hypothesized domain. These classes included masters of the skill represented in the domain, non-masters, and individuals in transition between non-mastery and mastery. Non-masters were characterized as failing all items in the domain, and masters as passing all items. Transitional individuals were assumed to respond inconsistently in a manner congruent with the assumption that they were still in the process of acquiring the concept or rule underlying mastery of the tasks in the domains under examination. Models asserting that tasks were in the same domain were compared to models asserting that the tasks were hierarchically ordered.

A Texas Instrument 745 terminal purchased for the project was used in testing the extent to which the hypothesized models accurately represented the observed performance of the subjects. The analysis revealed two hierarchically ordered domains.

The finding of hierarchically ordered domains and the discovery that tasks within a domain may vary in difficulty level raise questions about generalization during the course of learning to master domain tasks. These questions may have far-reaching implications for training. More specifically, it may be possible to use information about difficulty level within a domain to determine where to begin instruction for the domain, and how to advance from one domain to the next. This possibility has significant implications for training efficiency.

Narrative Table of Contents

Rationale for and Objectives of the Proposed Research

Since the time that Robert Gagné (1962) introduced his learning-hierarchy model in the early 1960's, there has been a growing recognition of the usefulness of empirically validated hierarchical learning sequences in teacher based, computer assisted, and computer managed training programs aimed at promoting the acquisition of basic math and science skills or at the development of performance capabilities related to various technical specialties pursued in military and industrial settings (Glaser, 1976; Glaser & Nitko, 1971; Glaser & Resnick, 1972; Nitko & Hsu, 1974; Resnick, Wang & Kaplan, 1973; White, 1973, 1974). However, despite the recognized usefulness of hierarchies, validated hierarchical sequences that can be applied in training are lacking. Moreover, there is at present, no adequate, practical technology for conducting hierarchy validations. Unless such a technology is developed, the contribution that validated sequences could make to training will not be realized.

The validation of a learning hierarchy requires the testing of three hypotheses. One is that the specific trainee responses measured in the validation process represent response classes defining skills capable of being applied under a range of different stimulus conditions (Gagné, 1977). The second is that subordinate skills in a hierarchy are prerequisite or necessary to superordinate skills (Gagné, 1977), and the third is that prerequisite skills mediate transfer for superordinate skills (Gagné, 1977). The present project is designed to investigate research questions related to the testing of these hypotheses for the purpose of establishing guidelines that can be used in the development of a technology for hierarchy validation.

The Need for Validated Hierarchies

The need for validated hierarchies stems from their recognized potential value in training and from the fact that there are no adequately validated hierarchies in use in training programs today. Validated hierarchies could make two kinds of contributions in training. One of these relates to issues in instructional design, the other to assessment.

The Potential Role of Hierarchies in Instructional Design

The central advantage claimed for hierarchies in the area of instructional design has to do with the development of instructional sequences to facilitate transfer of learning. In numerous places in the literature, Gagné has advanced the view that lower level subordinate skills which are prerequisite to superordinate skills at higher levels in a hierarchy mediate transfer for the superordinate skills to which they are related (e.g., Gagné, 1962, 1968, 1973, 1977). The implication for instructional design is that instructional sequences should be arranged so that prerequisite skills are available to the trainee at the time that superordinate skills are to be mastered (Gagné, 1973).

Advocates of the learning-hierarchy view have pointed out that instructional sequences which ensure that prerequisite skills are available at the time of learning may produce highly beneficial results (e.g., Gagne, 1973; Glaser & Resnick, 1972). A sequence which takes into account prerequisite skills maximizes the likelihood that trainees will have appropriate prerequisite competencies at the time they are needed for superordinate-skill learning. On the other hand, a sequence developed without consideration for prerequisite relations leaves the question of whether or not trainees possess needed prerequisite competencies to chance. The result may be that some trainees will fail to master superordinate skills because they lack the prerequisites to superordinate skill mastery.

The Potential Role of Hierarchies in Assessment

The main advantage of empirically validated hierarchies with respect to assessment relates to the problem of adapting instruction to the needs of individual trainees. Given validated hierarchies, tests may be developed to individualize the placement of trainees in an instructional sequence (Glaser & Nitko, 1971; Nitko & Hsu, 1974; Resnick, Wang, & Kaplan, 1973). Placement tests based on validated hierarchies may be used in the initial phases of instruction to determine the point in an instructional sequence which will enable a trainee to encounter readily attainable goals and at the same time to avoid activities related to objectives that have already been mastered. In addition, placement tests may be used at the end of a sequence to determine what has been learned and thereby to establish what should be taught next (Nitko & Hsu, 1974).

The Current Lack of Validated Hierarchies

White and Gagné (1974) have noted that although the learning-hierarchy model has had some influence on the development of instructional materials it has not yet had the wide application that might have been expected. One apparent reason for the failure of the learning-hierarchy model to have a greater impact on training than it has had is that there are currently no adequately validated hierarchies that could be used in training programs.

During the period since Gagné (1962) introduced the learning-hierarchy model, there have been several studies attempting to validate isolated hierarchical sequences (White & Gagné, 1974). However, early investigations on hierarchies were marred by serious methodological flaws (White, 1973). White (1973, 1974) suggested modifications in hierarchy validation procedures which eventuated in marked improvements in validation techniques. Despite these advances, adequate hierarchy validation has not yet been achieved.

As indicated in the initial paragraphs of the proposal, adequate hierarchy validation requires the examination of three hypotheses. Two of these three hypotheses have never been effectively tested in hierarchical research.

The hypothesis that skills in a hierarchy represent definable response classes has never been tested in hierarchy investigations. A few attempts have been made to assess the assumption that prerequisite skills mediate transfer for superordinate skills, but research in this area has had methodological flaws. Cotton, Gallagher, and Marshall (1977) have recently reviewed the literature on the transfer hypothesis and have concluded that Gangé's transfer assumption has never been tested. Gagné's third hypothesis, the prerequisite-skills assumption, has recently been subjected to effective study (White, 1974). However, the validation procedures used to examine the prerequisite-skills assumption are extremely time consuming and may not be suitable for broad scale application.

<u>Advances in Statistics that Make a Practical Technology for Hierarchy Validation Possible</u>

A major reason for the lack of progress in hierarchy validation described above is that until recently appropriate statistical procedures have not been available to test hypotheses germain to the development of effective, practical procedures for validating hierarchies. A number of procedures have recently become available which should make it possible to conduct hierarchy validations in a practical and effective way.

<u>New Techniques for Validating Prerequisite Relations</u>. During recent years Gagné's prerequisite-skills assumption has served as a focal point for efforts to develop statistical procedures for use in hierarchy validation. White (1973) has shown that techniques used to assess prerequisite relations by Gagne and his colleagues in early hierarchy research were

inadequate in that they failed to provide a statistical test for prerequisite associations which took into account errors in measurement. More recent research on prerequisite relations using a variety of scaling techniques including scalogram analysis (Guttman, 1944), multiple scalogram analysis (Lingoes, 1963), and the ordering theoretic method (Bart & Airasian, 1974; Bart & Krus, 1973) has been faulted on similar grounds. None of these procedures provides a suitable statistical test for prerequisite relations (Airasian, Madaus, & Woods, 1975; Dayton & Macready, 1976; White, 1974).

During recent years a number of attempts have been made to develop procedures to test Gagne's prerequisite-skills hypothesis statistically (Emrick & Adams, Note 2; Murray, Note 3; Proctor, 1970; White & Clark, 1973). Dayton and Macready (1976) have shown that each of these procedures represents a special case of a general latent-structure model which has the advantage of being capable of testing for prerequisite relations in both linear and nonlinear hierarchies. Goodman (1974, 1975) has also developed a latent-structure approach and a related model for scaling response patterns, both of which can be used to test for prerequisite associations in linear and nonlinear hierarchies.

New Techniques for Validating Positive Transfer. Although attempts to establish statistical techniques for use in hierarchy validation have focused mainly on Gagné's prerequisite-skills hypotheses, the need for procedures to examine Gagné's second major hypothesis, the positive-transfer assumptions are equally great. A recent review by Cotton, Gallagher, and Marshall (1977) attests to this fact. As indicated above, these investigators failed to find a single published study which provided a suitable test of Gagne's positive transfer assumption. Bergan (in press) has shown that structural equation models based on Sewall Wright's (1921, 1960) pioneering work in path analysis can be used to assess positive transfer in a learning hierarchy.

Structural-equation procedures based on regression analysis (Kerlinger & Pedhazur, 1973) are available for use with interval scale dependent measures (Duncan, 1975; Heise, 1975). In addition, Goodman (1972, 1973a, 1973b) has developed structural-equation techniques involving the use of log-linear models (Bishop, Fienberg, & Holland, 1975) that can be applied with dichotomous and polytomous scores of the types typically used in hierarchy validation.

New Techniques for Domain Validation. As indicated above, Gagné (1977) assumes that the skills in a learning hierarchy represent response classes rather than discrete behavioral capabilities. For example, within the learning-hierarchy viewpoint, it is assumed that a trainee who possesses a skill such as multiplying two mixed numbers will be able to use that skill to solve a broad range of similar problems.

One of the major problems in hierarchy validation is to determine whether or not the items on a test of skill performance measure the trainee's ability to perform the full range of behaviors included in the response class assumed to be represented in the skill under examination. Hively, Patterson, and Page (1968) used the term item domain to refer to the response class associated with a given skill. In addition, Hively and his colleagues developed a set of rules for generating test items falling within various domains. Since the early work of Hively and his associates, other investigators have elaborated on the concept of item domain and have attempted to develop item generating procedures for various types of domains (Shoemaker, 1975).

Although awareness of the need to determine empirically the extent to which specific test items represent an item domain has existed for some time, statistical procedures for empirically validating item domains associated with different skills have been lacking. For example, White (1974), in an article on hierarchy validation, discussed the need for determining statistically the extent to which different items assessed the same skill, but was forced to conclude that there were no available statistical procedures for making such a determination.

The Goodman (1975) response scaling technique and the Dayton and Macready (1976) latent-structure model are both suitable for use in empirically validating an item domain. For instance, to test the hypothesis that a set of items belong within the same domain using the Goodman scaling technique, one would hypothesize a scaling model composed of two scale types. One of these would represent those learners who had acquired the skill being assessed by the items in the domain under investigation. Trainees in this group would be expected to pass all domain items presented to them. The second scale type would represent learners who had not acquired the skill in question. Trainees in this group would be expected to fail all daomin items which they encountered. Either the chi-square goodness-of-fit or likelihood-ratio statistic can be used to test the fit of a model of this type to a set of data collected on item performance in the domain targeted for study.

A Structural Approach to Hierarchy Validation. The present research combines use of the Goodman (1974) latent structure techniques with structural equation procedures in which may be termed a structural approach to hierarchy validation. The research examines the validity of item domains in a hierarchy and addresses both Gagne's prerequisite-skills and positive-transfer hypotheses as these asusmptions relate to the task of developing practical procedures that can be applied in hierarchy validation in domain-referenced assessment and training design. The hierarchical relations selected for examination involve basic algebra skills included in military training. The specific skills targeted for study have been selected from the Precision Measuring Equipment Curriculum of the Advanced

Instructional System (AIS), an individualized training program operated by the Airforce at Lowrey Airforce Base. Analysis of these skills in the present project not only affords general guidelines for the validation of military training sequences, but also provides direct information that could be used to improve the efficiency and effectiveness of the precision measurement instructional unit.

Hierarchy Research Needs

Although adequate statistical procedures for examining hierarchical relations are now available, information is lacking on how to go about the validation process. Three kinds of research needs must be met before it will be possible to determine the most efficacious procedures for validating hierarchical associations. One of these involves the issue of how skills should be measured in validating the prerequisite-skills hypothesis. The second has to do with skill measurement in validating the positive-transfer hypothesis, and the third deals with domain validation in hierarchical sequences.

Needs Related to Prerequisite-Skills Validation. One of the initial steps in hierarchy validation is to test for hypothesized prerequisite relations in the hierarchy under examination. Two strategies have been suggested for accomplishing this task. Research is needed to determine whether or not these two procedures yield different results.

One of the strategies used in prerequisite-skills validation is the psychometric approach (Resnick, 1973; Wang, 1973). In this approach, trainees are tested on skills under examination in a hierarchy, and a statistical procedure is applied to determine the existance of prerequisite dependencies. Some years ago White (1973) criticized the psychometric approach on the grounds that it does not control for random forgetting. White took the position that skills in a hierarchy may be forgotten in a different order than the order in which they are learned. In accordance with this position, White (1974) argues that validation of the prerequisite-skills hypothesis requires a validation procedure in which learners who do not initially possess the skills in a hierarchy are taught the skills. He further suggested that testing for skill acquisition should be conducted during the course of learning rather than when instruction has been completed.

In support of the assumption of random forgetting, White cited only one study, an early investigation by Gagné and Bassler (1963). There are a number of reasons why the Gagné and Bassler study does not provide convincing evidence for the random forgetting assumption. First, adequate statistical procedures for testing the prerequisite skills hypothesis were unavailable at the time of the Gagné and Bassler investigation. Thus, it is not certain that all of the prerequisite relations that were assumed to be shown by the data actually did exist (White, 1976). Second, at the time of the investigation, there were no statistical techniques to assess the extent to which observed differences between learning and retention reflected measurement error as opposed to forgetting. Finally, the retention test which Gagné and Bassler used involved items which were different from the tiems used to assess learning. Thus, what Gagne and Bassler called a retention test could also be described as a test of generalization.

Recognition of the lack of convincing evidence provided by the Gagne and Bassler study has recently led White (1976) to suggest that the psychometric procedure ought to be reconsidered for use in hierarchy validation. The widespread application of hierarchical sequences in military training will require the validation of vast numbers of hierarchies. The psychometric approach to testing the prerequisite-skills hypothesis is much more efficient than the instructional strategy advocated by White. If it were possible to use the psychometric approach in the validation process and attain accurate results, a huge savings in time and personnel would be realized . In view of the superior efficiency of the psychometric approach and the lack of convincing evidence contra-indicating the use of the approach, research to assess the efficacy of the psychometric technique is clearly warranted. In this regard, there is a need to determine the extent to which hierarchical models validated under White's instructional strategy match models validated psychometrically. The present project is designed to meet this research need.

As indicated in the discussion of the Gagné and Bassler study, the extent to which skills are retained in the order in which they are learned has implications with respect to the utility of the psychometric approach. Skill retention may be affected not only be forgetting processes, but also by the kinds of experiences the learner has after training has been completed. For example, the extent to which an individual uses skills on the job after a training program has been terminated may influence skill retention. In order to establish fully the utility of the psychometric validation strategy there is a need for additional research on the question of whether or not skills are forgotten in a different order than the order in which they are learned. Such research should include not only the examination of retention shortly after the completion of training, but also the study of retention in the post-training work environment. The present project addresses this research need.

Needs Related to Positive-Transfer Validation. As indicated above published studies assessing Gagné's positive-transfer hypothesis are lacking. One possible reason for this lack is that procedures advocated for testing positive transfer are difficult and time consuming to implement. Many investigators, particularly those studying complex hierarchies involving many connections have dealt with the issue of transfer by ignoring it and focusing instead on the validation of prerequisite relations (White & Gagné, 1974).

Validation of Gagné's positive-transfer hypothesis has generally been conceptualized within a transfer-of-training paradigm. White and Gagné (1974) suggest a validation strategy which illustrates this fact. The White and Gagné approach involves the following steps: First, choose as many prerequisite relations in the hierarchy under consideration as can be examined within existing constraints on time and resources. Second, for each connection to be studied, identify groups of learners who possess all relevant prerequisite skills, but who lack the specific prerequisite and superordinate skills targeted for study. Third, conduct a standard transfer-of-training experiment in which half of the learners receive training on the superordinate skill. Positive transfer is indicated if learners receiving prerequisite skill training perform significantly better on the superordinate-skill training task than learners who do not receive prerequisite skill instruction.

As indicated above, Bergan (in press) has shown that Gagné's positive-transfer hypothesis can be tested using structural equation models. Within a structural-equation approach, direct and indirect effects among a set of variables can be examined in the absence of an experiment involving random assignment of individuals to treatment conditions (Duncan, 1975; Goodman, 1972; Heise, 1975). For example, in the case of interval scale data, the direct effects of one variable on another can be assessed using ordinary least squares regression techniques (Duncan, 1975). The magnitude of the direct effect of the first variable on the second is given by a structural coefficient which in ordinary least squares regression analysis is the regression coefficient in the regression equation.

A structural approach to testing Gagné's positive-transfer hypothesis is potentially more efficient than the procedure suggested by White and Gagné. The increased efficiency derives from the fact that structural equations can be used with the same data-collectionprocedures as those employed in prerequisite-skills validation. Thus, for example, structural equations can be used to examine positive transfer using White's (1974) instructional procedure for prerequisite-skills validation. White's instructional procedure requires less time and is more practical to implement than the White and Gagné (1974) transfer paradigm in that it necessitates only one group of learners who are taught all skills in a linear sequence whereas many groups learning different skills are needed to implement the White and Gagne transfer procedure.

Structural equations can be used to achieve an even greater gain in efficiency than that associated wtih the use of the White instructional technique if they are coupled in positive-transfer validation with the psychometric validation procedure. The psychometric procedure is, of course, much more efficient than the White and Gagne approach in that all that is required to implement the technique is to test a group of trainees.

To apply structural equations to test the assumption that prerequisite skills mediate transfer for superordinate skills, prerequisite and superordinate skills must first be identified. This can be accomplished using prerequisite-skills validation procedures discussed above. After prerequisite and superordinate skills have been determined, a structural model comprised of equations expressing hypothesized effects of previously validated prerequisite skills on superordinate skills can be constructed. Data from either the White instructional procedure of the psychometric procedure can then be used in testing model-data fit.

It is possible that structural equations used either with White's instructional technique or with the psychometric procedure would not yield the same results as would be attained using the White and Gagne experimental paradigm. If this were to occur, it could be argued that the White and

Gagné approach provided a more valid demonstration of transfer than a structural equation approach using prerequisite-skills validation procedures in that the White and Gagné paradigm is experimental whereas the structural-equation approach is not. However, if structural-equation procedures used with prerequisite-skills validation procedures could be assumed to yield the same transfer relations as identified through the White and Gagne paradigm, then a substantial gain in efficiency could be attained in the validation process.

Research is needed to determine the extent to which structural-equation techniques coupled with instructional or psychometric validation procedures reveal the same transfer relations as those established through the use of the White and Gagne experimental paradigm. The present project is designed to meet this research need.

A corollary of the positive-transfer hypothesis that has appeared in the literature from time to time (e.g., Cotton, Gallagher, & Marshall, 1977; Resnick, 1973; Uprichard, 1970) is that not only will prerequisite skills mediate transfer for superordinate skills, but also that instruction given in the order suggested by the hierarchy will produce superior transfer to that attained through the use of any other order. This hypothesis can be investigated by teaching the skills under examination in all possible orders (Cotton, Gallagher, & Marshall, 1977; Uprichard, 1970). The number of connections that can be examined in this way is limited since the number of possible orders becomes quite large when more than a few skills are subjected to study. Thus, to validate hypothesized order effects in a large hierarchy, it is necessary to conduct several studies on subsets of skills in a manner analogous to the White and Gagné (1974) approach described above.

Cotton, Gallagher, and Marshall (1977) point out that the assumption that hierarchical sequencing is maximally effective is important in determining the usefulness of hierarchies in designing instructional sequences. They indicate further that the order assumption has never been adequately tested. One of the aims of the present project is to test the hypothesis that hierarchical sequencing produces optimal learning.

Needs Relating to Domain Validation. The validation of item domains is an essential precursor to adequate examination of the other major hypothesis involved in hierarchy validation. Without domain validation, it is impossible to determine the extent to which test items reflect the response classes that they are assumed to represent (Gagné, 1977). In the absence of domain validation, failure to confirm either prerequisite-skills or positive-transfer hypotheses could be attributed to the possibility that the specific items used in validation did not adequately represent hypothesized response classes for the skills under investigation.

The empirical determination of relations among tasks within domains requires the construction of models to represent homogeneous domains. A number of models assume some kind of equivalence relation among tasks in a homogeneous domain. That is, they all assume that tasks will tend to be responded to in the same way by at least some groups of individuals. For example, Dayton and Macready (1976) have conceptualized homogeneous domains in terms of models that assume a mastery class composed of individuals who tend to perform all domain tasks correctly and a non-mastery class comprised of individuals who tend to fail all tasks in the domain. By contrast, Bergan, Cancelli, and Luiten (in press) have described models based on Goodman's (1975) work in response scaling that assume three classes of individuals in a homogeneous domain, non-masters, masters, and what Goodman (1975) calls unscalable individuals. Masters are assumed to perform all tasks in the domain correctly while non-masters are assumed to fail all tasks. Individuals in the unscalable category tend to manifest responses inconsistent with non-mastery or mastery and may be thought of as being in a transition state between non-mastery and mastery.

Varying assumptions may be made about task difficulty (i.e., the probability of accurate performance) within mastery, non-mastery, or transitional classes. More specifically, it may be assumed that task difficulty varies within classes or that it is equal across tasks within classes. For example, consider two algebra tasks shown empirically to belong in a domain characterized by problems in which a common term, say x, has to be factored from an expression such as (xa + xb). Suppose that the tasks were similar in all significant respects except that one necessitated three steps to achieve a solution and the other required only two steps. Suppose further that a model including masters, non-masters, and transitional individuals were used to describe relations among the tasks in this domain. Under this kind of model, masters would be assumed to perform all tasks correctly. For masters, the two tasks would be equally difficult in that the same proportion of individuals (i.e., all individuals) would display mastery of each task. Since non-masters would be assumed to fail all tasks, the tasks would also be equally difficult for them. By contrast, the tasks could vary in difficulty for transitional individuals. It would be reasonable to assume that the problem requiring three steps for solution would be more difficult than the problem requiring two steps for transitional individuals.

The possibility of within domain variations in task difficulty suggests that in a certain sense there may be sequential ordering within domains as well as between hierarchically related domains. As already indicated, the tasks within a domain are assumed to be equivalent, but equivalence may not always imply complete symmetry. Tasks that vary in difficulty for a given class such as that of transitional individuals may be thought of as being asymmetrically equivalent. Sets of asymmetrically equivalent tasks may be ordered by difficulty to form a sequence within a domain. Nothing is known about the conditions that may produce asymmetrical equivalence relations within a domain. The present technical report examines the hypothesis that tasks within domains comprised of algebra problems will form asymmetrical equivalence relations congruent with variations in the number of steps required to achieve problem solution.

The presence of an ordered relation between tasks provides one criterion that can be used to establish boundaries between domains. The concept of domain boundaries is, of course, essential in delimiting the content of a domain. Nonetheless, it is not necessary to think of boundaries as impermeable walls. Domains may include large numbers of tasks, and it is quite possible that some inter-domain task comparisons may suggest boundary permeability. For example, suppose that a group of item sets were used to assess performance on three academic tasks, A, B, and C. Assume that task A was shown to be asymmetrically equivalent to task B and that task B was found to be asymmetrically equivalent to task C. In addition, suppose that an ordered relation were observed in which A was found to be subordinate to C. In a case such as this, A and C would be in separate domains, but B would be in both the A domain and the C domain. Thus, the boundary between the A domain and the C domain would be permeable. The present technical report examines the possibility of permeability in domain boundaries. In this connection it is hypothesized that if permeability does exist, it will occur between tasks at the higher levels of a subordinate domain and the lower levels of the related superordinate domain.

Project Objectives

Objectives for the present technical report focus on the attainment of Task 1 objectives. These include both outcome and enabling objectives.

Outcome Objective for Task 1. To validate psychometrically the ordering of item domains for algebra tests selected from an examination of the Precision Measuring Equipment Curriculum.

Enabling Objectives.

a. To task analyze algebra skills from psychometrically validated domains selected from the Precision Measuring Equipment Curriculum.

b. To construct and write item domains for each hypothesized domain.

c. To construct a domain referenced test of items randomly selected from each domain.

d. To administer the test to approximately 200 subjects.

e. To score responses.

f. To construct and test latent class models to determine the extent to which hypothesized models fit (i.e., accurately represent) observed test performance.

Method

Subjects

The subjects were 317 volunteers from a high school and university in the Southwest selected to represent a wide range of skill levels in solving algebra problems. Subjects ranged from high school freshmen taking a first course in basic mathematics to university students a number of whom had had college math courses. There were approximately equal numbers of males and females representing a broad spectrum of ethnic backgrounds. Approximately 88% were Anglo, 8% were Mexican-American, and 4% were divided among Blacks, native American Indians, and Asians. More subjects were used than the 200 originally intended for the study so that the full range of algebra skills likely to be present in military trainees would be represented.

Tasks

A group of algebra tasks hypothesized to form an ordered set of behavioral domains was selected for use in conducting domain structure analysis. Algebra was chosen because it is a highly structured content area. The structured nature of the discipline facilitated the formulation of hypothesized domains and domain orderings.

An adaptation of facet analysis (Berk, 1978; Millman, 1974) was used in formulating hypothesized domains and domain orderings. Facets were defined as classes of behavioral operations involved in performing algebra tasks. Three facets were identified for this study: transposition of terms, application of the distributive property, and factoring. Each facet was hypothesized to represent a homogeneous item domain.

Problems within each domain varied in terms of the number of steps required to achieve problem solution. For example, some problems could be solved in a single step such as multiplying both sides of an equation by one term or expression. Other problems required as many as five steps for solution. It was assumed that item sets within each domain would form asymmetrical equivalence relations sequenced in accordance with the number of steps necessary for problem solution.

The hypothesized domains identified in the study do not represent independent dimensions. For example, it is impossible to solve factoring problems without transposing terms. The inclusion of operations defining one domain in problems reflecting another domain suggested an ordering of the domains congruent with Gagné's (1962, 1977) view that component tasks form an ordered sequence. An examination of the hypothesized domains to identify components suggested that the term-transposition domain would be subordinate to both the distributive property and factoring domains.

Problems illustrating the hypothesized domains are shown in Table 1 in Appendix A.

The first domain included problems requiring the transposition of terms from one side of an algebra equation to the other. Transposition was effected by one or more arithmetic operations (e.g., multiplication or subtraction). For instance, the first problem shown in Table 1 for this domain required transpositing the term A to the right side of the equation by multiplying both sides of the equation by A. The second domain involved applcations of the distributive property in which a single term had to be multiplied with each of two terms in an expression. The third domain required factoring a common term from an expression. For example, in the problems in Table 1, X must be factored from expressions including the terms N and R. Factoring is regarded in algebra texts as an application of the distributive property. This application involves a reversal of the multiplication operations carried out in using the distributive property.

Each of the three hypothesized domains involved problems representing an ordered set of elements. Ordering was based on the number of steps required for problem solution. For example, the first problem shown in Table 1 for the term transposition domain required only one step to achieve problem solution. By contrast, the second problem required two steps.

Variations in number of required steps were by necessity different for different domains. For example, the simplest factoring problem required two steps for solution. First a common term X had to be factored from an expression. Then the expression had to be moved to the right side of the equation. The term transposition domain contained two step categories: one-step problems and two-step problems. The distributive property domain contained three step classes: three-step problems, four-step problems, and five-step problems. The factoring domain contained the largest number of step categories. Factoring problems ranged from two steps to five steps.

Test Construction and Scoring

Following the facet analysis, item forms and item form shells (Hively, Maxwell, Rabell, Sension & Lundin, 1973) representing each of the domains and step categories within domains. The item forms provided descriptions of the classes of problems to be solved, stimulus and response characteristics of those classes, and cell matrices indicating class variations. The item form shells indicated materials, directions, scoring specifications, and replacement rules for generating items. The item form approach was used because it makes it possible to represent the population of problems in a domain in a precise fashion.

Test items were constructed to correspond to item form specifications. Two items representing identical problems were prepared for each type of algebra task included in the study. These items varied only in the specific letters used to represent equation terms. This made it possible to reflect variations in response consistency in the models used to assess domain structure.

Each pair of items representing a task was scored 1, 2, or 3. A 1 indicated that neither of the two items was answered correctly. A 2 indicated that one of the two item pairs was answered correctly, and a 3 indicated that both items were responded to correctly.

Procedures

Testing was carried out in groups of about thirty. The participants were told that the purpose of the study was to determine how people solved algebra problems. After the test booklets were passed out, the experimenter gave instructions for responding to the test. Trainees were instructed to solve the algebra problems presented and to write their solutions in the test booklets provided. Trainees were instructed to attempt all problems and to provide solutions even in cases in which they were unsure of the answers. Following the instructions the trainees were told to begin the test and were assured that they would have as much time as necessary to complete the problems. During the course of the testing, the experimenter and an assistant monitored each subject's performance to insure that the task was understood. The vast majority of the subjects comprehended what they were to do on the basis of the initial instruction. However, in one or two cases there were some questions. When this happened, the experimenter simply repeated the instructions for the individual having difficulty. In all cases the repeated instruction was sufficient to enable the individual to respond to the questions. Latent class models (Goodman, 1974) were used to assess equivalence and ordered relations among the algebra tasks examined in the study. Latent class models explain association in a contingency table in terms of a latent (i.e., unobserved) variable or set of latent variables each of which includes a set of latent classes. For example, in the present research latent class models were constructed to reflect variations in task mastery. The latent variable in this case was mastery variations. This variable included different latent classes, such as a mastery class and a non-mastery class. A latent class model can be used to generate maximum likelihood estimates of expected cell frequencies which indicate expected response patterns under the assumption that the model being examined is true.

Testing Latent Class Models

Latent class models are tested by assessing the correspondence between observed cell frequencies and estimates of expected cell frequencies using the chi-squared statistic. When the correspondence between observed and expected frequencies is close, the value of X^2 will be low and the model being tested can be said to provide an adequate fit for the data. Clifford Clogg (Note 1) has developed a computer program that carries out the iterative process used to generate maximum likelihood estimates of expected cell frequencies and that computes the X^2 value to test the fit of a model to a data set. Clogg's program was used in the present investigation.

Models Tested

The latent class models designed for the present project were intended to distinguish between ordered and equivalence relations among algebra tasks. To understand why the models were designed as they were, it is necessary to understand model distinctions involving the ordering and equivalence of tasks. Consider Table 2 in Appendix A, cross-classifying performance on two items. Thus, a subject's score for each task may fall into one of three categories, zero right, one right, or two right. These categories can be designated by the numbers 1, 2, and 3 respectively.

In a table of this kind, a score of 1 on each task would suggest non-mastery. This response pattern would be reflected in the 11 cell in the table. A score of 3 on each task would suggest mastery. This pattern is reflected in the 33 cell. A score of 3 on task A and 1 on task B would indicate mastery of task A without evidence of mastery of task B. Scores of 2 would reveal inconsistent performance characteristic of transition between non-mastery and mastery. Since the items for each task are identical, scores of 2 should reflect errors which ought to occur at a relatively low frequency.

Given an ordered relation between tasks A and B, the number of responses in the 31 cell should be significantly greater than the number in the 32 cell. Under the assumption of ordering, a build-up would be expected in the

31 cell indicating that a significant number of subjects had mastered A without having begun to master B. The 32 cell would be expected to have relatively few responses because the 2 category represents response inconsistency for task B.

If the tasks were equivalent, the number of individuals in both the 31 and 32 cells would be small since both these cells would reflect response inconsistency. The relation between the 31 and 32 cells would not be crucial so long as the probability for the 31 cell was not larger than the probability for the 32 cell. Two relations between the 31 and 32 cells could occur without contraindicating the equivalence assumption. Either the cells could be equiprobable or there might be a significantly greater number of individuals in the 32 cell than in the 31 cell.

As this discussion shows, a critical issue in determining whether two tasks form an ordered or equivalence relation is that of determining whether the hypothesis that the occurrence of responses in the 31 and 32 cells is equiprobable is supported by the data. If this hypothesis is rejected, it is necessary to determine whether the probability of a response in the 31 cell is greater than the probability of a response in the 32 cell for masters of task A. If this turns out to be the case, a model describing an ordered relation between the tasks may be considered. If the probability for the 31 cell is not greater than the probability for the 32 cell, an equivalence model may be suggested to represent the data.

Eight latent class models were examined in the study. The models are described in the following paragraphs and are displayed visually in Table 3 (App.A). The E's and curved lines in the visual display indicate cells constrained to be equiprobable under a given model. The I's indicate cells for which the assumption is made that the probability of a given response level on task A is independent of the probability of any particular response level on task B. The X's indicate response patterns associated with specific latent classes. For example, the X in the 11 cell of H_1 indicates the association of the 11 response pattern with the non-mastery latent class.

The Independence-Equiprobability Model. The first model, designated H_0, asserts independence between task pairs and equiprobability between categories 1 and 2 for the task assumed to be the least difficult in the task pair. This model served as a standard against which to compare the other models tested. The equiprobability provision was included to make the model congruent with models being examined. As mentioned earlier, the central criterion for distinguishing between ordered and equivalence relations is one asserting equiprobability between certain task categories. The equiprobability provision was included in model H_0, as well as some of the other models examined, to provide a basis for distinguishing between ordered and equivalence relations. If there had been any instances in which model H_0 provided an adequate description of tasks in the domain under examination, the hypothesis that the tasks were not related would have been supported.

The Model of Symmetry. Model H_1 asserted symmetrical equivalence between tasks. Model H_1 included 6 latent classes: a non-mastery class, a partial mastery class, a mastery class, and 3 transition classes reflecting symmetrical inaccuracies in responding. The 3 classes assuming inaccurate responding each asserted equiprobability for one pair of cells in the table cross-classifying the tasks under examination. For example, one of these classes asserted that the probability of the 12 cell would be equal to the probability of the 21 cell. The second asserted that the probability of the 13 cell would be equal to the probability of the 31 cell, and the third assumed that the probability of the 23 cell would be equal to the probability of the 32 cell. Because of the symmetrical nature of its equiprobability restrictions, this model has been described in the literature as the model of symmetry (Bishop, Fienberg, & Holland, 1975). The model of symmetry implies equal item difficulty for the tasks under examination. Tasks for which this model provided an adequate fit for the data were described as being symmetrically equivalent.

Asymmetrical Equivalence Models. Model H_2 included 3 latent classes, a mastery class, a non-mastery class, and an unscalable class composed of transitional individuals. Model H_2 assumed that masters would respond correctly to all problems presented to them. Thus, in the mastery class the probability of the 33 response pattern was restricted to be 1. Similarly, the model assumed that non-masters would fail all problems. Thus, in the non-mastery class the probability of the 11 category was restricted to be 1. It was presumed that in the unscalable category, the probability of a particular level of performance on one task would be independent of a given level of performance on the other tasks, and that the 1 and 2 categories would be equiprobable for one of the tasks. The equiprobability restriction was included as a criterion for distinguishing between equivalence and ordered relations for reasons already discussed.

Model H_2' is a special case of model H_2. It is like model H_2 in all respects except that is does not include the equiprobability restriction imposed under H_2. Model H_2' was included to reflect the fact that two tasks may be equivalent even though the 1 and 2 categories of the more difficult task are not equiprobable. It may happen that the probability of a response in the 32 cell is greater than the probability of a response in the 31 cell. This is exactly the opposite of what is to be expected under the hypothesis of an ordered relation between tasks. When the hypothesis of equiprobability is rejected, but the probability of the 32 cell is greater than the probability of the 31 cell, it is appropriate to test models which assert equivalence, but which do not include equiprobability restrictions. Model H_2' is one such model.

Model H_3 included 4 latent classes, a non-mastery class, a partial mastery class, a mastery class, and an unscalable class. The partial mastery class was similar to the unscalable class in that both reflected less than completely accurate responding on the part of examinees. However, model H_3 asserted that individuals in the partial mastery class consistently performed 1 out of 2 problems correctly on both tasks under examination for a given task pair. More specifically, the partial mastery class asserted that for members of that class the probability of getting 1 out of 2 items correct for both tasks would be 1. The unscalable class did not assume this kind of consistency in partially accurate responding.

Model H_3' assumed four latent classes, a non-mastery class, a partial mastery class, a mastery class, and an unscalable class. The restrictions for non-mastery, partial mastery, and mastery classes were the same as those given for H_3. Moreover, similar restrictions were imposed for partial mastery.

Model H_3' differed from H_3 because it did not impose an equiprobability restriction in the unscalable category. The concept of partial mastery implies a significant number of individuals who get 1 problem right. Given this state of affairs, not only should a build-up of individuals in the 22 category be expected, but also it would not be unreasonable for the probability of occurrence of the 32 category to be greater than the probability for the 31 category. Model H_3' reflects the fact that equiprobability need not always occur in a model asserting equivalence between tasks.

Model H_4 is very similar to H_2. The difference between the two is related to the equiprobability restriction in the unscalable class. In asserting both independence and equiprobability, model H_4 necessarily makes the 21 and 22 cells as well as the 31 and 32 cells equiprobable in the unscalable latent class. Equiprobability does not obtain for the 11 and 12 cells because the 11 cell represents a separate latent class, i.e., the non-mastery class. Model H_4 restricts equiprobability in the unscalable class to the 31 and 32 cells. This is accomplished by making the 21 cell represent a separate latent class. The probability of the 21 response pattern in this class is restricted to be 1. The effect of this is to make the observed and expected cell frequencies for the 21 pattern equal. Thus the pattern contributes nothing to the value of X_2. With the exception of the restriction on the 21 cell, model H_4 is exactly the same as H_2. Like H_4, it contains mastery, non-mastery and unscalable latent classes. Moreover, the restrictions on the mastery and non-mastery classes are the same as those for H_2. The unscalable category assumes independence between tasks with the 21 pattern ruled out of consideration. In addition, it asserts equiprobability for the 31 and 32 cells.

An Ordered Relation Model. Model H_5 asserted an ordered relation between task pairs. This model contained four latent classes, a non-mastery class, an unscalable class, a mastery class and a subordinate task mastery class. The restrictions in the non-mastery and mastery classes were identical to those used in the equivalence models. Independence was assumed in the unscalable class. In the subordinate task mastery class the probability of passing both subordinate task items was assumed to be 1. The probability of passing both superordinate task items was assumed to be zero and the probabilities of getting no correct responses and 1 correct response on the superordinate task were set equal to the observed proportions of responses in those two categories. The last of these restrictions was imposed so that all individuals who had mastered the subordinate task including those in transition toward superordinate task mastery would be included in the latent class reflecting subordinate task mastery.

Results

Within Domain Results

Results of the model testing within-domains revealed two domains instead of the three hypothesized. The factoring and distributive property problems turned out to be in one domain. Tables 4 and 5, in Appendix A, present the observed responses for the cross-classification of every possible task pair for each of the two domains. Table 4 shows the cross-classification for the term transposition domain while Table 5 displays the cross-classification for the Distributive Property-factoring domain. In Table 4 the letters indicate the addition-subtraction (A) and multiplication-division (M) dimensions. In Table 5 they stand for factoring (F) and distributive property (D) problems. Numbers in both tables represent the number of steps required for problem solution.

The response patterns in the tables indicate various combinations of the number of correct responses for each task pair examined. For example, the 11 pattern indicates no correct responses on either task while the 33 pattern represents 2 correct responses for each task. Note the large number of responses falling in the 11 and 33 categories in the tables. These patterns represent the critical cells for establishing equivalence relations. Notice further that most task sets have about the same number of individuals in the 31 and 32 cells. The 31 cell represents individuals who have mastered one task, but have not begun to acquire the second task. As already indicated, given an ordered relation between tasks, the number of individuals in the 31 cell would be expected to be larger than the number of individuals in the 32 cell. On the other hand, given an equivalence relation between the tasks, the number of individuals in both the 31 and 32 cells would be expected to be small.

Tables 6 and 7, in Appendix A, present the results of model testing for the hypothesized domains. In the model testing process, all possible pairs of tasks within a given domain were compared. Table 6 shows the chi-squared tests for all possible task pairs in the term transposition domain. The letters designating tasks refer to the addition-subtraction (A) and multiplication-division(M) dimensions for this domain. The numbers refer to the number of steps required for problem solution. For example, 1 refers to a problem requiring only one step for solution.

The model testing process required the selection of a preferred model based on statistical comparisons among various models examined. To illustrate the comparison process, consider the results for H_0 and H_2 for the A1-M1 task pair given in Table 6. The X^2 value for model H_0 is 200.65 with 5 degrees of freedom, which is significant well beyond the .01 level. The X^2 value for model H_2 is 1.18 with 3 degrees of freedom which has a p value of about .90. Model H_0 and H_2 are hierarchical. That is, H_2 contains all of the characteristics of H_0 plus 2 additional characteristics. These additional characteristics reflect the inclusion of a mastery and non-mastery latent class under H_2. Model H_2 has 3 degrees of freedom, whereas H_0 has 5. The loss of 2 degrees of freedom reflects the inclusion of the non-mastery and mastery latent classes. Because H_0 and H_2 are hierarchical, they can be compared statistically (Goodman, 1974). The X^2 for H_2 can be subtracted from the X^2 for H_0. The result will be an X^2 with 2 degrees of freedom. In the case of the A1-M1 task pair, the subtraction of H_2 from H_0 yields an X^2 of 198.47 with 2 degrees of freedom, which is significant far beyond the .01 level. Model H_2 provides an excellent fit for the data. Moreover, none of the models improve over H_2. Consequently, H_2 was selected as the preferred model for the A1-M1 task pair. Not all of the models in Table are hierarchically related. For example, H_1, the symmetry model, is not hierarchically related to either H_0 or H_2. Consequently, it is not possible to compare H_1 directly with H_0 or H_2.

The results on Table 6 show that in no case did model H_0 or H_1 provide an acceptable fit for the data. Consequently, the hypothesis that the task pairs under examination were unrelated and the hypothesis that they were symmetrically equivalent could be rejected for all of the tasks investigated.

In all cases except one, one of the asymmetrical equivalence models provided an acceptable fit for the data. In some instances, the model including an equiprobability restriction provided an adequate fit. In other cases, for example in the case of task pair A1-A2, the equiprobability assumption was rejected. However, the probability of being in the 32 cell was found to be higher than the probability of being in the 31 cell. Consequently, it could safely be concluded that the tasks for this pair were not ordered.

The one instance in which the hypothesis of equivalence relations was rejected was that involving the A1-M2 task pair. The two tasks involved in this comparison represented marked differences in difficulty level within the item domain. The one-step addition problem was the simplest task in the domain, whereas the two-step multiplication problem was among the most difficult. Model H_5 provided an acceptable fit for these two tasks indicating an ordered relation between them. The ordered relation for the A1-M2 task pair suggests permeability in domain boundaries. Tasks A1 and A2 are in the same domain. A2 and M2 are not in the same domain. The fact that A1 and M2 are found to be in separate domains suggests that the boundaries between domains may not be rigid.

The results for the term transposition domain reflect a highly consistent pattern. As already indicated, the hypothesis of asymmetrical equivalence was supported in every instance except one. The asymmetrical equivalence observed in the domain reveals a structured arrangement of tasks. The tasks requiring two steps for problem solution are more difficult than those requiring only a single step.

Table 7 shows the results of model testing for the combined distributive property-factoring domain. The letters in Table 7 refer to factoring problems (F) and distributive property problems (D), and the numbers indicate the number of steps required for problem solution.

As in the case of the transposition domain, the results for the combined distributive property-factoring domain reveal a highly consistent pattern. In most instances, one of the asymmetrical equivalence models provides a suitable fit for the data. However, in some cases, the model of symmetry fit the data to an acceptable degree. This suggests that at the higher levels of algebra skill, problems are more likely to be equivalent for all groups of individuals, including those in transition. This is understandable since those in transition with respect to higher level skills bring a broad background of subordinate skills to the task of solving higher level factoring and distributive property problems.

In only one case did a task not form an equivalence relation with other tasks. This was the case for the most difficult factoring task. Model H_5 provided an acceptable fit for comparisons involving this task. Model testing revealed that this task was superordinate to all of the other distributive property and factoring tasks. Analysis of the characteristics of the task revealed that it required not only factoring, but also application of the multiplication operations used in distributive property problems. This suggests the existence of a superordinate domain hierarchically related to the factoring-distributive property domain. Further research is needed to investigate this possibility.

Between Domain Results

Table 8, in Appendix A, presents observed response patterns for the cross-classification of tasks representing the term transposition and factoring-distributive property domains. Note the large number of individuals attaining the <u>31</u> response pattern and the relatively small number of individuals for the <u>32</u> pattern. This is what is to be expected under the hypothesis of an ordered relation between task pairs.

Table 9 in Appendix A, shows the chi-squared tests for the cross-classifications in Table 8. In all cases model H_5 afforded an acceptable fit for the data, and in all cases except four model H_5 was preferred over the other models tested. Two equivalence models were preferred over H_5 in these four cases. Model H_4 was preferred for the comparison involving two-step addition and three-step application of the distributive property and the comparison of two-step addition with the five-step distributive property problem. Model H_3 was preferred for the comparison of two-step multiplication and two-step factoring. These cases provide additional evidence of boundary permeability.

Figure 1 summarizes both the within-domain and between-domain findings. The circles indicate domains. Ordering of tasks within domains and between domains is indicated by position in the vertical dimension. The long tube penetrating the two circles represents permeability in domain boundaries.

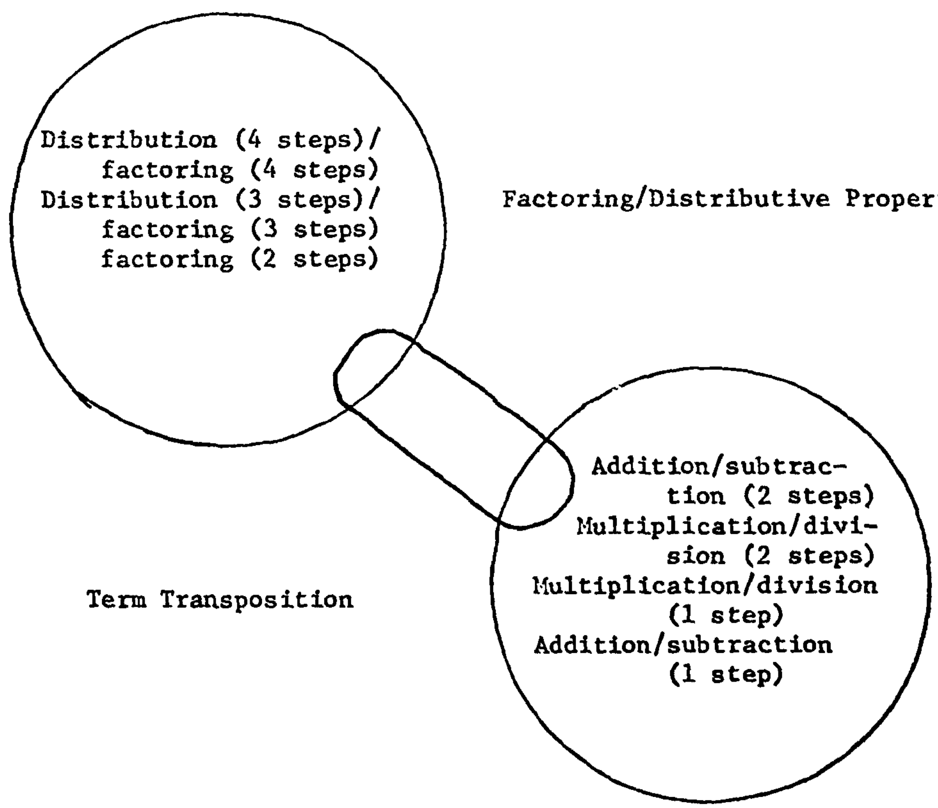

Figure 1.

Discussion

The results for task comparisons both within and between domains supported the major hypotheses advanced in the study. The within-domain findings are congruent with the view that algebra tasks representing a class of mathematical operations may be organized into homogeneous domains that involve asymmetrical equivalence relations. Moreover, as hypothesized asymmetrical equivalence is related to the number of steps required to achieve problem solution. The discovery of asymmetrical ordering raises questions about generalization and transfer within domains that may be important for instruction. For example, it is possible that instruction in a high difficulty task but also in generalization to low difficulty tasks. By contrast, instruction in a low difficulty task might not generalize directly to high difficulty problems. However, mastery of a low difficulty task could mediate positive transfer facilitating high difficulty task learning. Possibilities such as these call for research relating domain structure to generalization and transfer issues.

The unexpected finding that factoring and distributive property problems involving term and expression multiplication were in the same domain suggests that homogeneous domains may encompass rather broad classes of tasks. While it is true that factoring and multiplying an expression by a term are both regarded by mathematicians as applications of the distributive property, these tasks are nonetheless quite different in terms of the specific operations that they require. The fact that they were found to be in the same domain suggests that generalization of algebra skills may be very broad indeed. Research is needed to determine the breadth of generalization within domains.

The results for the between domain comparison support the hypothesis that ordered relations may exist between pairs of tasks in which one task is a component of the other. This finding linked to the within-domain results raises additional generalization and transfer questions with potentially important instructional implications. All of these relate to the question of how a student can best advance from a subordinate domain to a superordinate domain. For example, it would be of interest to know whether positive transfer would be significantly greater from a high difficulty subordinate domain task to a low difficulty superodinate domain task than from a high difficulty subordinate task to a high difficulty superordinate task.

The results with respect to boundary permeability raise additional questions regarding advancement from a subordinate to a superordinate domain. The findings suggest that permeability may exist and thereby raise the possibility of direct generalization between subordinate and superordinate domains. However, ambiguity in the permeability findings indicate the need for further research on the permeability phenomenon before conducting generalization studies. Permeability did not always occur in the manner hypothesized. In some cases it did take place as expected between the top level of the subordinate domain and the bottom level of the superordinate domain. However, in other instances it involved problems not adjacent to the domain boundary. This can be explained by the fact that to some extent permeability may be a function of unreliable responding. For example, a large number of individuals performing inconsistently on an hypothesized superordinate task would produce a buildup in the 32 cell that could mask the presence of an ordered relation. Examination of the observed response patterns in Table 9 suggests that some instances of apparent permeability may have resulted from high levels of inconsistent superordinate task responding. However, it is also true that the numbers in the 31 cell were generally smaller for task pairs close to the boundary between domains than for pairs far from the boundary. This suggests permeability. In order to resolve the permeability question, constant low levels of inconsistent superordinate task responding would be required. Further research is needed to study the relation of permeability to response inconsistency.

Reference Notes

1. Bernard, M.E. Task analysis in instructional program development. Theoretical Paper No. 52. Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, Wisconsin, May, 1975.

2. Emrick, J.A. & Adams, E.N. An evaluation model for individualized instruction (Research Report RC 2674) Yorktown Hts., New York: IBM Thomas J. Watson Research Center, 1969.

3. Murray, J.R. Statistical models for qualitative data with classification errors. Unpublished doctoral dissertation, University of Chicago, 1971.

4. Traub, R.E. Stirring muddy water: Another perspective on criterion-referenced measurement. Ontario Institute for Studies in Education, mimeo, 1975.

References

Airasian, P. W., Madaus, G. F., & Woods, E. M. Scaling attitude items: A comparison of scalogram analysis and ordering theory. Educational and Psychological Measurement, 1975, 35, 809-819.

Bart, W. M., & Airasian, P. W. Determination of the ordering among seven Piagetian tasks by an ordering-theoretic method. Journal of Educational Psychology, 1974, 66, 277-284.

Bart, W. M., Krus, D. J. An ordering-theoretic method to determine hierarchies among items. Educational and Psychology Measurement, 1973, 33, 291-300.

Bergan, J.R. The structural analysis of behavior: An alternative to the learning hierarchy model. Review of Educational Research, in press.

Bergan, J. R., Cancelli, A. A. & Luiten, J. Mastery assessment with latent class and quasi-independence models representing homogeneous item domains. Journal of Educational Statistics, in press.

Berk, R. A. The application of structural facet theory to achievement test construction. Educational Research Quarterly, 1978, 3, 62-72.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. Discrete multivariate analysis. Cambridge: MIT Press, 1975.

Cotton, J.W., Gallagher, J.P., & Marshall, S. P. The identification and decomposition of hierarchical tasks. American Educational Research Journal, 1977, 14, 189-212.

Dayton, C. M., & Macready, G. B. A probabilistic model for validation of behavioral hierarchies. Psychometrika, 1976, 41, 189-204.

Duncan, O. D. Introduction to structural equation models. New York: Academic Press, 1975.

Gagné, R.M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R.M. Contributions of learning to human development. Psychological Review, 1968, 75, 177-191.

Gagné, R.M. Learning and instructional sequence. In F. Kerlinger (Ed.), Review of Research in Education (Vol. 1). Itasca, Ill.: Peacock, 1973.

Gagné, R.M. The conditions of learning (3rd ed.). New York: Holt, Rinehart & Winston, 1977.

Gagné, R.M., Bassler, O.C. Study of retention of some topics of elementary nonmetric geometry. Journal of Educational Psychology, 1963, 54, 123-131.

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

Glaser, R., & Nitko, A.J. Measurement in learning and instruction. In R.L. Thorndike (Ed.), Educational Measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Glaser, R., & Resnick, L.B. Instructional psychology. Annual Review of Psychology, 1972, 23, 207-276.

Goodman, L.A. A general model for the analysis of surveys. American Journal of Sociology, 1972, 77, 1035-1085.

Goodman, L.A. Causal analysis of data from panel studies and other kinds of surveys. American Journal of Sociology, 1973, 78, 1135-1191. (a)

Goodman, L.A. The analysis of multidimensional contingency tables when some variables are posterior to others: A modified path analysis approach. Biometrika, 1973, 60, 179-192. (b)

Goodman, L.A. The analysis of systems of quantitative variables when some of the variables are unobservable. Part I--A modified latent structure approach. American Journal of Sociology, 1974, 79, 1179-1259.

Goodman, L.A. A new model for scaling response patterns: An application of the quasi-independence concept. Journal of the American Statistical Association, 1975, 70, 755-768.

Guttman, L. A basis for scaling qualitative data. American Sociological Review, 1944, 9, 139-150.

Heiss, D.R. Causal analysis. New York: Wiley, 1975.

Hively, W., Maxwell, G., Rabehl, G., Sension, D., & Lunden, S. Domain referenced curriculum evaluation: A technical handbook and a case study from the Minnemast Project (CSE Monograph Series in Evaluation No. 1). Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation, 1973.

Hively, W., Patterson, H.L., & Page, S.H. A "universe-defined" system of arithmetic achievement tests. Journal of Educational Measurement, 1968, 5, 275-290.

Kerlinger, F.N., & Pedhazur, E.J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, 1973.

Nitko, A.J. & Hsu, T. Using domain-referenced tests for student placement, diagnosis and attainment in a system of adaptive, individualized instruction. In W. Hively (Ed.), Domain-referenced testing. Englewood Cliffs, New Jersey: Educational Technology Publications, 1974.

Popham, W.J. Criterion-referenced measurement. Englewood Cliffs, N.J.: Prentice-Hall, 1978.

Popham, W.J. An approaching peril: Cloud-referenced tests. Phi Delta Kappan, 1974, 56, 614-615.

Proctor, C.H. A probabilistic formulation and statistical analysis for Guttman scaling. Psychometrika, 1970, 35, 73-78.

Resnick, L.B. (Ed.). Hierarchies in children's learning: A symposium. International Science, 1973, 2, 311-362.

Resnick, L.B., Wang, M.C., & Kaplan, J. Task analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Journal of Applied Behavioral Analysis, 1973, 7, 679-710.

Shoemaker, D.M. Toward a framework for achievement testing. Review of Educational Research, 1975, 45, 127-147.

Uprichard, A.E. The effects of sequence in the acquisition of three set relations: An experiment with preschoolers. Arithmetic Teacher, 1970, 17, 597-604.

Wang, M.C. Psychometric studies in the validation of an early learning curriculum. Child Development, 1973, 44, 54-60.

White, R.T. Learning hierarchies. Review of Educational Research, 1973, 43, 361-375.

White, R.T. The validation of a learning hierarchy. American Educational Research Journal, 1974, 11, 121-136.

White, R.T. Effects of guidance, sequence, and attribute-treatment interactions on learning, retention, and transfer of hierarchically ordered skills. Instructional Science, 1976, 5, 133-152.

White, R.T., & Clark, R.M. A test of inclusion which allows for errors of measurement. Psychometrika, 1973, 38, 77-86.

White, R.E., & Gagne, R.M. Past and future research on learning hierarchies. Educational Psychologist, 1974, 11, 19-28.

Wright, S. Correlation and causation. Journal of Agricultural Research, 1921, 20, 557-585.

Wright, S. Path coefficients and path regressions: Alternative or complimentary concepts? Biometrics, 1960, 16, 189-202.

APPENDIX A

Table 1

Sample Problems From Hypothesized Domains[1]

Domain	Problems
Term Transposition	$X/A = B$
	$X + A + B = C$
Distributive Property	$N(X+R) = Z$
	$\frac{A(X+B)}{C} = D$
Factoring	$NX + RX = Y$
	$(NX + RX)Y = Z$

[1] In each case the task was to solve for X.

Table 2

Cross-Classification of Two

Hypothesized Tasks

	Task B		
Task A	1	2	3
1			
2			
3			

Table 3

Models Used in Establishing Item Domains[1]

H_o

B

		1	2	3
	1	E	E	I
A	2	E	E	I
	3	E	E	I

H_1

B

		1	2	3
	1	X	E	E
A	2	E	X	E
	3	E	E	X

H_2

B

		1	2	3
	1	X	I	I
A	2	E	E	I
	3	E	E	X

H_2'

B

		1	2	3
	1	X	I	I
A	2	I	I	I
	3	I	I	X

H_3

B

		1	2	3
	1	X	I	I
A	2	I	X	I
	3	E	E	X

H_3'

B

		1	2	3
	1	X	I	I
A	2	I	X	I
	3	I	I	X

H_4

B

		1	2	3
	1	X	I	I
A	2	X	I	I
	3	E	E	X

H_5

B

		1	2	3
	1	X	I	I
A	2	I	I	I
	3	X	X	X

1. The E's connected by curved lines indicate cells constrained to be equiprobable. The I's indicate cells for which the hypothesis of independence prevails. The X's indicate cells reflecting response patterns associated with specific latent classes.

Table 4

Observed Cross-Classifications for the Term-Transposition Domain[1]

Response Patterns		Cross-Classifications					
Tasks							
A	B	A1 - M1	A1 - A2	A1 - M2	M1 - A2	M1 - M2	A2 - M2
1	1	65	72	69	82	99	97
1	2	4	2	2	12	2	22
1	3	6	1	4	9	2	16
2	1	14	22	28	14	16	6
2	2	12	12	4	13	10	10
2	3	12	4	6	8	9	10
3	1	24	19	38	17	20	10
3	2	19	40	20	29	14	22
3	3	161	145	146	133	145	124

1. The letters in the letter-number combinations labeling the columns below the cross-classifications heading indicate addition-subtraction (A) or multiplication-divison (M) problems. The numbers refer to the number of steps required for problem solution.

Table 5

Observed Cross-Classifications for the Factoring-Distributive Property Domain[1]

Response Patterns: Tasks A	B	Cross-Classifications: F2-F3	F2-F4	F2-F5	F2-D3	F2-D4	F2-D5	F3-F4	F3-F5	F3-D3	F3-D4	F3-D5	F4-F5	F4-D3	F4-D4	F4-D5	F5-D3	F5-D4	F5-D5	D3-D4	D3-D5	D4-D5
1	1	160	157	159	144	151	154	169	169	150	161	166	175	156	168	175	159	172	185	152	161	162
1	2	2	6	1	11	18	16	11	2	8	11	12	1	3	6	10	3	4	3	17	8	6
1	3	1	0	3	15	25	41	0	11	12	22	33	7	11	20	26	8	18	23	1	1	0
2	1	11	13	15	10	8	5	9	14	12	8	6	15	16	9	8	19	12	13	8	19	26
2	2	10	6	5	9	5	6	6	2	5	3	5	3	4	3	5	1	1	5	13	4	13
2	3	10	12	11	15	9	32	14	13	17	11	32	14	14	10	30	14	9	25	13	11	17
3	1	9	13	28	9	4	4	5	21	14	7	4	12	19	14	8	24	18	4	8	31	23
3	2	17	20	5	11	8	9	15	7	6	5	2	7	14	12	6	7	6	3	26	31	24
3	3	97	90	90	93	89	50	88	80	93	89	54	83	80	75	49	82	77	56	79	51	46

1. Letters in the cross-classification columns indicate factoring (F) and distributive property (D) problems. Numbers refer to the number of steps required for problem solution.

Table 6

Chi-squared Tests for the Term-transposition Domain[1]

Tasks	H_o		H_1		H_2		H_2'		H_3		H_3'		H_4		H_5	
A B	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p
A1-M1	200.65	<.01	19.04	<.01	1.18*	<.90	.29	<.90	1.19	<.50	.16	<.90	.87	<.75	.28	<.75
A1-A2	251.78	<.01	73.48	<.01	10.92	<.025	9.71	<.01	10.94	<.01	1.38*	<.25	7.73	<.025	.09	<.90
A1-M2	237.56	<.01	66.65	<.01	29.12	<.01	6.73	<.05	29.14	<.01	2.63*	<.25	5.75	<.10	.07	<.90
M1-A2	197.55	<.01	15.32	<.01	3.36*	<.05	1.93	<.50	3.37	<.25	1.35	<.25	3.27	<.25	.10	<.90
M1-M2	328.54	<.01	30.58	<.01	2.58*	<.50	.06	<.175	2.59	<.50	.04	<.90	1.07	<.75	.01	<.90
A2-M2	203.97	<.01	15.73	<.01	6.69	<.10	.41*	<.90	6.50	<.05	0	<.99	6.70	<.05	.19	<.75

[1]The letters in the letter-number combinations used in designating task pairs indicate the addition-subtractions (A) and multiplication-division(M) dimensions. Numbers refer to the number of steps required to achieve problem solution.

The degrees of freedom for H_o through H_5 are as follows:

H_o	5	H_3	2
H_1	3	H_3'	1
H_2	3	H_4	2
H_2'	2	H_5	1

Asterisks indicate preferred models.

Table 7

Chi-squared Tests for the Factoring-distributive Property Domain[1]

Tasks	H_o		H_1		H_2		H_2'		H_3		H_3'		H_4		H_5	
A B	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p
D3-F2	305.52	<.01	1.6086*	<.75	.33	<.975	.28	<.90	.34	<.90	.02	<.90	.32	<.90	.12	<.75
D3-F3	345.09	<.01	6.4418*	<.10	6.51	<.10	1.66	<.50	6.51	<.05	1.07	<.50	4.76	<.10	1.47	<.25
D3-F4	337.67	<.01	11.924	<.01	10.35	<.025	3.52	<.25	10.39	<.01	1.99	<.25	.67*	<.75	1.00	1.00
D3-F5	382.18	<.01	29.4	<.01	30.40	<.01	3.63	<.25	31.02	<.01	0	1	11.97	<.01	2.10	<.10
D3-D4	323.14	<.01	13.93	<.01	25.33	<.01	11.71	<.01	25.34	<.01	4.27*	<.05	25.34	<.01	11.28	<.01
D3-D5	285.98	<.01	49.99	<.01	15.28	<.01	13.55	<.01	15.35	<.01	2.25*	<.25	9.60	<.01	9.59	<.01
D4-F2	292.64	<.01	20.939	<.01	2.06*	<.75	2.02	<.50	2.07	<.50	1.18	<.50	1.53	<.25	.08	<.90
D4-F3	340.53	<.01	10.929	<.025	2.69*	<.50	.86	<.75	2.69	<.50	0	1	1.03	<.75	.69	<.50
D4-F4	320.66	<.01	1.85072*	<.75	4.54*	<.25	2.08	<.50	4.56	<.25	1.55	<.25	.15	<.95	4.00	1.00
D4-F5	363.86	<.01	4.79*	<.25	20.33	<.01	1.92	<.50	20.88	<.01	.67	<.50	6.66	<.05	.38	<.75
D4-D5	325.15	<.01	27.57	<.01	12.39	<.01	11.21	<.01	12.43	<.01	.86*	<.50	11.41	<.01	8.74	<.01
D5-F2	234.24	<.01	55.13	<.01	5.32*	<.25	2.10	<.50	5.32	<.10	.02	<.90	5.32	<.10	2.00	<.25
D5-F3	310.60	<.01	59.904	<.01	3.14*	<.50	3.14	<.25	3.14	<.25	1.83	<.25	3.15	<.25	2.20	<.25
D5-F4	305.29	<.01	12.4176	<.01	2.29*	<.75	2.06	<.50	2.29	<.50	.73	<.50	2.26	<.50	1.97	<.25
D5-F5	387.10	<.01	41.264	<.01	6.35*	<.10	1.12	<.75	6.35	<.05	1.12	<.50	.45	<.90	.30	<.75

1. The letters in the task descriptions refer to factoring (F) and distributive property (D) problems. The numbers indicate number of steps to problem solution.

* Asterisk indicates preferred model.

Table 7 (continued)

Tasks	H_o		H_1		H_2		H_2'		H_3		H_3'		H_4		H_5	
A B	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p	X^2	p
F2-F3	434.24	<.01	16.06	<.01	2.81*	<.50	1.66	<.50	2.82	<.25	0	< 1	2.80	<.25	.30	<.75
F2-F4	392.54	<.01	22.68	<.01	12.63	<.01	12.50	<.01	12.63	<.01	4.58*	<.05	11.85	<.01	10.36	<.01
F2-F5	420.47	<.01	40.26	<.01	23.67	<.01	.77	<.75	23.68	<.01	.06	<.90	17.24	<.01	.06	<.90
F3-F4	427.12	<.01	7.17*	<.01	23.69	<.01	20.09	<.01	23.70	<.01	5.92*	<.025	23.70	<.01	18.25	<.01
F3-F5	384.90	<.01	15.13	<.01	19.84	<.01	1.17	<.75	20.18	<.01	.52	<.50	7.38	<.025	.02	<.90
F4-F5	441.56	<.01	18.41	<.01	12.40	<.01	2.63	<.50	12.42	<.01	1.72	<.25	1.44*	<.50	.11	<.75

The degrees of freedom for the models are:

H_o	5
H_1	3
H_2	3
H_2'	2
H_3	2
H_3'	1
H_4	2
H_5	1

Table 8

Observed Cross-classifications for the Term-transposition and Factoring-distributive Property Domains

Response Patterns		Cross-Classifications						
Tasks								
A	B	A_1-D_3	A_1-D_4	A_1-D_5	A_1-F_2	A_1-F_3	A_1-F_4	A_1-F_5
1	1	74	74	75	74	75	74	75
1	2	0	1	0	0	0	1	0
1	3	1	0	0	1	0	0	0
2	1	32	36	37	32	33	35	33
2	2	4	1	0	3	2	1	2
2	3	2	1	1	3	3	2	3
3	1	64	84	99	57	68	82	94
3	2	30	20	43	28	17	19	9
3	3	110	100	62	119	119	103	101
Tasks								
A	B	M_1-D_3	M_1-D_4	M_1-D_5	M_1-F_2	M_1-F_3	M_1-F_4	M_1-F_5
1	1	97	102	103	97	100	101	103
1	2	2	1	0	3	1	1	0
1	3	4	0	0	3	2	1	0
2	1	26	30	33	28	29	32	32
2	2	5	3	1	4	4	1	1
2	3	4	2	1	3	2	2	2
3	1	47	62	75	38	47	58	67
3	2	27	18	42	24	14	19	10
3	3	105	99	62	117	118	102	102

Table 8 (continued)

Response Patterns		Cross-classifications						
Tasks								
A	B	A_2-D_3	A_2-D_4	A_2-D_5	A_2-F_2	A_2-F_3	A_2-F_4	A_2-F_5
1	1	107	111	112	101	106	106	107
1	2	3	1	0	5	1	1	2
1	3	3	1	1	7	6	6	4
2	1	32	40	44	33	36	41	42
2	2	11	9	5	12	7	4	2
2	3	11	5	5	9	11	9	10
3	1	31	43	55	29	34	44	53
3	2	20	12	38	14	11	16	7
3	3	99	95	57	107	105	90	90
Tasks								
A	B	M_2-D_3	M_2-D_4	M_2-D_5	M_2-F_2	M_2-F_3	M_2-F_4	M_2-F_5
1	1	128	134	133	126	131	133	134
1	2	2	1	2	5	2	1	1
1	3	5	0	0	4	2	1	0
2	1	15	19	21	14	15	15	17
2	2	2	3	2	4	3	2	2
2	3	9	4	3	8	8	9	7
3	1	27	41	57	23	30	43	51
3	2	30	18	39	22	14	18	8
3	3	99	97	60	111	112	95	97

Table 9

Chi-squared Tests for the Cross-classification of the Term-transposition and Factoring-distributive Property Domains

	H_o			H_1			H_2			H_2'			H_3			H_3'			H_4			H_5		
	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p
A1-D3	245.08	5	<.01	153.54	3	<.01	41.61	3	<.01	9.74	2	<.01	41.64	2	<.01	3.90	1	<.05	14.50	2	<.01	1.92*	1	<.25
A1-D4	282.59	5	<.01	179.62	3	<.01	84.55	3	<.01	7.97	2	<.025	84.63	2	<.01	.21	1	<.75	43.40	2	<.01	1.03*	1	<.50
A1-D5	234.28	5	<.01	239.99	3	<.01	73.99	3	<.01	23.22	2	<.01	22.70	2	<.01	0.00	0	-	30.03	2	<.01	0.00*	0	-
A1-F2	255.42	5	<.01	137.93	3	<.01	41.75	3	<.01	11.49	2	<.01	42.54	2	<.01	3.40	1	<.10	11.34	2	<.01	.32*	1	<.75
A1-F3	294.32	5	<.01	150.83	3	<.01	65.96	3	<.01	4.45	2	<.25	68.62	2	<.01	0.00	0	-	34.69	2	<.01	0.00*	0	-
A1-F4	280.64	5	<.01	170.35	3	<.01	83.33	3	<.01	7.70	2	<.025	87.23	2	<.01	.42	1	<.75	44.08	2	<.01	1.73*	1	<.25
A1-F5	316.31	5	<.01	179.20	3	<.01	114.91	3	<.01	.35	2	<.90	117.96	2	<.01	0.00	0	-	84.37	2	<.01	0.00*	0	-
M1-D3	250.24	5	<.01	86.20	3	<.01	25.51	3	<.01	7.09	2	<.05	25.52	2	<.01	4.01	1	<.05	6.20	2	<.05	.72*	1	<.50
M1-D4	318.08	5	<.01	134.81	3	<.01	51.46	3	<.01	3.63	2	<.25	51.48	2	<.01	.39	1	<.75	26.51	2	<.01	.91*	1	<.50
M1-D5	266.43	5	<.01	199.83	3	<.01	47.55	3	<.01	18.63	2	<.01	49.95	2	<.01	0.00	1	1	13.85	2	<.01	0.00*	1	1
M1-F2	279.96	5	<.01	77.23	3	<.01	26.18	3	<.01	8.91	2	<.025	26.18	2	<.01	2.88*	1	<.10	3.25*	2	<.25	.07	1	<.90
M1-F3	326.77	5	<.01	94.16	3	<.01	44.45	3	<.01	3.51	2	<.25	44.45	2	<.01	2.39	1	<.25	19.75	2	<.01	.91*	1	<.50
M1-F4	319.03	5	<.01	124.34	3	<.01	59.12	3	<.01	9.77	2	<.01	60.16	2	<.01	1.02	1	<.50	22.60	2	<.01	.14*	1	<.75
M1-F5	368.08	5	<.01	143.06	3	<.01	84.07	3	<.01	3.09	1	<.25	86.57	2	<.01	0.00	0	-	49.61	2	<.01	0.00*	1	1
A2-D3	276.43	5	<.01	57.54	3	<.01	13.67	3	<.01	2.07*	2	<.50	13.67	2	<.01	.45	1	<.75	2.39*	2	<.25	0.00	1	1
A2-D4	338.43	5	<.01	101.86	3	<.01	40.79	3	<.01	.40	2	<.90	40.79	2	<.01	.27	1	<.75	18.69	2	<.01	.15*	1	<.75
A2-D5	266.18	5	<.01	155.94	3	<.01	42.17	3	<.01	18.70	2	<.01	7.93	2	<.05	3.73	1	<.10	4.45*	2	<.25	1.30	1	<.50

Table 9 (continued)

	H_o			H_1			H_2			H_2'			H_3			H_3'			H_4			H_5		
Tasks	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p	X^2	d.f.	p
A2-F2	264.70	5	<.01	38.62	3	<.01	18.94	3	<.01	1.44	2	<.50	18.96	2	<.01	1.42	1	<.25	6.08	2	<.05	.73*	1	<.50
A2-F3	308.47	5	<.01	63.73	3	<.01	40.61	3	<.01	3.29	2	<.25	40.61	2	<.01	3.11	1	<.10	13.87	2	<.01	1.54*	1	<.25
A2-F4	291.12	5	<.01	83.38	3	<.01	57.29	3	<.01	8.63	2	<.025	57.70	2	<.01	5.14	1	<.025	17.03	2	<.01	.70*	1	<.50
A2-F5	328.79	5	<.01	95.31	3	<.01	87.38	3	<.01	1.82	2	<.50	87.75	2	<.01	.01	1	<.95	43.90	2	<.01	.62*	1	<.50
M2-D3	302.80	5	<.01	39.80	3	<.01	12.32	3	<.01	10.66	2	<.01	13.07	2	<.01	2.70*	1	<.25	3.36	2	<.25	.25	1	<.75
M2-D4	365.06	5	<.01	86.26	3	<.01	22.78	3	<.01	3.61	2	<.25	22.81	2	<.01	.75	1	<.50	10.74	2	<.01	1.53*	1	<.25
M2-D5	277.69	5	<.01	133.92	3	<.01	22.56	3	<.01	11.17	2	<.01	4.12*	2	<.25	.71	1	<.50	6.23	2	<.05	2.83	1	<.10
M2-F2	312.90	5	<.01	26.01	3	<.01	5.92*	3	<.25	4.11	2	<.25	5.93	2	<.01	.12	1	<.75	1.06	2	<.75	1.04	1	<.50
M2-F3	362.28	5	<.01	42.31	3	<.01	14.70	3	<.01	1.76	2	<.50	14.75	2	<.01	.01	1	<.95	6.61	2	<.05	.66*	1	<.50
M2-F4	344.26	5	<.01	69.21	3	<.01	21.81	3	<.01	2.71	2	<.50	10.62	2	<.01	.06	1	<.90	14.16	2	<.01	.84*	1	<.50
M2-F5	387.86	5	<.01	88.00	3	<.01	49.58	3	<.01	2.68	2	<.50	37.37	2	<.01	2.41	1	<.25	37.64	2	<.025	2.68*	1	<.10