DOCUMENT RESUME

ED 050 145 TM 000 532

AUTHOR O'Malley, J. Michael
TITLE Application of a Curriculum Hierarchy Evaluation (CHE) Model to Sequentially Arranged Tasks.
INSTITUTION Hawaii Univ., Honolulu. Div. of Special Education.
PUB DATE Feb 71
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Ability Identification, Academic Aptitude, Concept Formation, *Curriculum Evaluation, Developmental Tasks, Disadvantaged Youth, Early Childhood Education, *Learning Theories, *Models, *Preschool Children, Preschool Curriculum, Sequential Learning, Skill Analysis, *Task Performance, Transfer of Training, Visual Stimuli

ABSTRACT

A curriculum hierarchy evaluation (CHE) model was developed by combining a transfer paradigm with an aptitude-treatment-task interaction (ATTI) paradigm. Positive transfer was predicted between sequentially arranged tasks, and a programed or nonprogramed treatment was predicted to interact with aptitude and with tasks. Eighteen four and five year-old urban disadvantaged boys and girls from a Head Start class in Honolulu, randomly assigned to one of three groups, were administered multiple discrimination and concept tasks under sequences which would reveal predicted transfer and interaction effects. The CHE model successfully identified transfer on the curriculum hierarchy from the multiple discrimination to the concept task, and appeared to serve as an empirical check upon a task analysis of the concept task. A programed sequence was superior to a non-programed sequence irrespective of aptitude or task. The most severe restriction of the model is that it is limited in application to only two tasks in a linear hierarchy. It is suggested that the model be extended to include different tasks and aptitudes. (Author/LR)

EDUCATION RESEARCH AND DEVELOPMENT CENTER

COLLEGE OF EDUCATION
UNIVERSITY OF HAWAII
HONOLULU, HAWAII

APPLICATION OF A CURRICULUM HIERARCHY EVALUATION (CHE) MODEL TO SEQUENTIALLY ARRANGED TASKS[1]

J. Michael O'Malley
Education Research and Development Center
University of Hawaii[2]

1. This paper was presented at the Annual Meetings of the American Educational Research Association, New York, February, 1971.

2. The research reported herein was supported by the Education Research and Development Center (EDRAD) of the University of Hawaii. The author would like to express his gratitude for cooperation and assistance to the Susannah Wesley Preschool and in particular to Mrs. Vi Brown, the classroom teacher. J. Michael O'Malley is currently Assistant Researcher in EDRAD and Assistant Professor of Educational Psychology at the University of Hawaii, Honolulu, Hawaii 96822.

Abstract

A curriculum hierarchy evaluation (CHE) model was developed by combining a transfer paradigm with an aptitude-treatment-task interaction (ATTI) paradigm. Positive transfer was predicted between sequentially arranged tasks, and a programed or nonprogramed treatment was predicted to interact with aptitude and with tasks. Eighteen preschool children randomly assigned to one of three groups were administered multiple discrimination and concept tasks under sequences which would reveal predicted transfer and interaction effects. The CHE model successfully identified transfer on the curriculum hierarchy from the multiple discrimination to the concept task, and appeared to serve as an empirical check upon a task analysis of the concept task. A programed sequence was superior to a non-programed sequence irrespective of aptitude or task.

APPLICATION OF A CURRICULUM HIERARCHY EVALUATION (CHE) MODEL TO SEQUENTIALLY ARRANGED TASKS

J. Michael O'Malley

University of Hawaii

Theoretical analysis of complex behavioral repertoires have been the foundation for curricula in a variety of subject areas (Resnick, 1967). The theoretical analysis reveals a basic set of skills which are conceived as components or prerequisites of performance on the complex repertoire; it also suggests a method for hierarchically arranging these skills to produce efficient learner acquisition. Resnick and Wang (1969) have commented on the need for models to evaluate the efficiency of the learning produced by a theoretical hierarchy and have proposed a model which predicts positive transfer between sequentially arranged tasks. A more flexible paradigm will be produced by coordinating the transfer model with the model for aptitude-treatment-task interactions (Tobias, 1969) owing to the added aptitude and task dimensions.

The blend of the transfer and ATTI paradigms, the curriculum hierarchy evaluator (CHE) can be applied by selecting a theoretical hierarchy from tasks in a typical early school curriculum and determining whether the model successfully verifies the proposed sequence. A theoretical hierarchy exists, according to Gagné (1965, 1970), between a task involving acquisition of a multiple discrimination (MD) and one involving concept learning (CL). The two tasks are related since a task analysis will identify the MD process as a component of the CL task.

Verification of a theoretical sequence with the CHE model is accomplished by equating the treatment dimension with the paths through the curriculum from a lower level to a higher level task. The treatment dimension may be comprised of the following three different paths through the MD task leading to the CL task: no experience on the MD task, experience with a serial order of item presentation, and experience with a programed order of item presentation. A programed order of item presentation designed to reduce intralist interference has been shown by O'Malley (1970) to be more effective than a serial order in producing learning on a MD task. The difficulty of the CL task thus should be reduced substantially for children who receive a programed order on the MD task and less so for children who receive a serial order or no experience at all. Furthermore, the children who perform better on the CL task, those who received the programed MD task, should also perform better on a second CL task since they will have acquired the first CL task to a greater extent than the other two groups.

One alternative in verifying the theoretical hierarchy in the CHE model is to reverse the predicted sequence of tasks. If a MD task is a component of a CL task, as indicated in the theoretical analysis, acquisition of the CL task should enable a child to completely perform the MD task.

The task dimension of the CHE model consists of the two tasks in the theoretical hierarchy, tasks on which a treatment may be differentially effective. The programed order used by O'Malley (1970) to reduce intralist interference on a MD task has yet to be applied to a CL task. The CL task is predicted to be easier under a programed order since it is presented in the present study in the form of a list of items, much as the MD task.

The research on ATTIs has often failed to produce the predicted interactions with aptitude for a variety of suspected reasons, one of which may be that the tasks investigated do not contain similar processes. The use of tasks in which a hierarchical relationship has grown out of a theoretical analysis of a complex repertoire may resolve some of the difficulty in attaining the desired interactions. The aptitude dimension of the CHE model in this study will consist of performance on a MD task which parallels the MD component of the concept task, but which involves labels which are unrelated to the concept. This aptitude dimension is predicted to interact on both MD and CL tasks with the treatment, a programed or nonprogramed order of item presentation. The advantages of a programed order which reduces intralist interference may accrue particularly to lower ability children since children of higher ability may not experience as much interference.

Hypotheses

1. Children with a programed order of item presentation obtain significantly more correct responses on a given number of trials on a multiple discrimination (MD) or a concept learning (CL) task than children with a serial order of item presentation. The programed order is more effective for children of lower abilities as measured by an unrelated MD task. Support for the main effects on the MD task, but not necessarily the aptitude-treatment interaction, is necessary before the second hypothesis can be evaluated since the rationale for the second hypothesis is based upon these main effects.

2. Children with a programed order of item presentation on a MD task obtain significantly more correct responses on a given number of trials on a hierarchically related concept task than children with a

nonprogramed order or children with no MD task experiences. Because of their superior acquisition of the concept, which was based upon superior acquisition of the MD task component, these children will also obtain significantly more correct responses on a given number of trials on an identical concept which uses different labels.

3. Significantly more correct responses on a MD task are obtained by children having prior experience with these materials on a related concept task than a group with no such experience with these materials. The existence of the multiple discrimination process as part of the concept task should provide children having prior experience on the concept task with significantly more correct responses on the multiple discrimination.

Method

Subjects

The subjects in this study were 20 four- and five-year old urban disadvantaged boys and girls attending a Head Start class in Honolulu, Hawaii. The parents of these children all resided in a federally constructed high-rise apartment dwelling and were predominantly of mixed Polynesian, Micronesian, and Melanesian origins. Two Ss were rejected because of failure to participate in the learning tasks.

Materials

The materials in this study consisted of two sets of eight 3" x 4" cards on which were depicted pictures of objects or scenes with which the children were familiar. None of the children on a pretest, however, could designate the proper verbal label or the principle which would be associated with each picture. A principle, as Gagné (1965) uses the term, is the use of two concepts to establish a relationship, e.g., SAND IS DRY. Each set of materials (S1 and S2) was used in two tasks: the MD task, in which children learned the verbal label corresponding to the picture; and the CL task, in which children responded on a second card of a pair with its label and a familiar opposite to the characteristic verbalized by the E for the label of the first card. The concept involved in the CL task was to anticipate the opposite to the characteristic verbalized by the E. For example, if the E said THE SAND IS DRY for a picture of sand, the S was expected to say THE TREE IS WET for a picture in which rain was falling on a tree.

Procedures

All children received two MD tasks and two CL tasks with treatments consisting of either a programed sequence (PS) or a serial sequence (SS). The PS treatment was designed to permit cumulative review of prior items with a minimum number of interspersed items, whereas the SS treatment permitted no cumulative review and maximized the number of interspersed items. The two MD and CL tasks were administered in four separate 15-minute sessions on four successive days. The sequences in which the tasks-treatments were administered were designed to impose minimum training sessions but yield maximum information by virtue of the number of hypotheses which could be generated for these particular tasks and treatments. Each of the eight cards in the MD task and each of the four pairs of cards in the CL task was presented manually on five occasions irrespective of the treatment sequence or the session in which the training took place. Only social reinforcement, such as YES and RIGHT contingent upon correct responses, was used throughout all training sessions. The dependent variable was number of correct responses per S across items for the five presentations.

All children were given a MD task with familiar items as a warm-up task. Four training sessions were then introduced, as presented in Table 1. The children were all administered SS-MD-S1 in Session 1 as a test of ability level, the aptitude dimension of the CHE model. They were then randomly assigned to one of three Session 2 and Session 3 conditions: (a) Group 1 received PS-MD-S2 followed by PS-CL-S2; (b) Group 2 received SS-MD-S2 followed by PS-CL-S2; and (c) Group 3 received PS-CL-S2 followed by PS-MD-S2. Half of the Ss in each of the three groups were then randomly assigned to receive either PS-CL-S1 or SS-CL-S1 in Session 4.

Insert Table 1 about here

Results

Analytic Approach

The data of ATTI studies according to Cohen (1968) can be suitably analyzed using multiple regression techniques. These techniques are ideal for the situation where the regression slopes of ability on the dependent variable for the treatments and tasks may be different, as in the present study. Cohen (1965) elsewhere recommends reporting R^2 as well as F to support interpretations of the results.

Aptitude Treatment Interaction

The first hypothesis states that the programed and nonprogramed sequences of item presentation will be differentially effective on both multiple discrimination and concept tasks for children of different ability levels. A programed order should be of more benefit to children of lower ability levels than to children of higher ability levels.

On the multiple discrimination task, the evaluation of this aptitude-treatment-interaction hypothesis was accomplished by contrasting the predictability of a full model with the interaction against a restricted model without the interaction. In terms of the representation in Table 1, Group 1 will perform better than Group 2 at Session 2 but will do so differentially for children of different ability levels. The full regression model consisting of the aptitude-treatment interaction did not predict the multiple discrimination score significantly better than a noninteraction restricted model, with $F_{(1,8)} = .68$, and $R^2 = .74$ for the full model. However, the group with a programed multiple discrimination task performed significantly better than the group with a nonprogramed task when ability was used as a covariate. These main effects were significant at $p < .01$ with $F_{(1,9)} = 15.88$, and $R^2 = .74$ for the full model. The relationship between the covariate and the dependent variable in this analysis was $r = .52$, a value which approached being significantly different from zero at the .05 level with $r_{(11)} = .55$. Also, none of the Ss in any of the three treatment groups differed significantly in terms of the covariate, with $F_{(2, 15)} = .55$, and $R^2 = .07$ for the full model.

The evaluation of the aptitude-treatment interaction hypothesis with a concept task was limited by the fact that the relationship between the covariate, a multiple discrimination task, and the dependent variable was only $r = .07$. Thus, without a related covariate it was only appropriate to test for main effects between the programed and nonprogramed procedures in acquiring the concept task. In table 1, this is represented by contrasting Subgroup A with Subgroup B at Session 4. The test of the main effect showed the programed sequence to be significantly superior to the nonprogramed sequence at $p < .01$, with $F_{(1,16)} = 16.94$, and $R^2 = .51$ for the full model. Subgroup A did not differ from Subgroup B in terms of the Session 1 SMD1 task, however, with $F_{(1,16)} < 1.00$ and $R^2 = .004$ for the full model.

Curriculum Hierarchy

The second hypothesis states that children with a programed order or item presentation on a MD task obtain significantly more correct responses on a given number of trials on a hierarchically related CL task than children with a nonprogramed order or children with no MD task experience. This assumes that more learning took place on the MD task with a programed order, as was verified in the first hypothesis. Referring to Table 1, the second hypothesis indicates that at Session 3, Group 1 will perform better than Group 2 due to Group 1's programed experience on the MD task in Session 2; and Group 2 at Session 3 will perform better than Group 3 at Session 2, since Group 2 had some experience with the MD task at Session 2 and Group 3 had no such experience. The mean scores on the concept task for the three groups were, respectively, 11.50, 12.17, and 7.00. The differences among the three groups are significant with $F_{(2,15)} = 5.54$, and $R^2 = .42$ for the full model. Judging from the mean scores, the feature differentiating these groups was whether or not they received any experience at all on the multiple discrimination task, not the hypothesized programed or nonprogramed experience.

The second hypothesis also predicted that those children who performed better on the first CL task, the result of their exposure to a programed

MD task, will perform significantly better on a second CL task in which the same concept is involved. Although the second concept task employed different labels than were introduced on the first concept task or the multiple discrimination, performance should be superior since superior acquisition of the first CL task should facilitate learning on the second task. The programed multiple discrimination learning would produce sufficient familiarity with the labels employed in the concept task to ease the acquisition of the concept and thereby produce fewer errors on a task in which the same concept but different labels were involved. In terms of the groups used in the present study, the hypothesis indicates that superior acquisition of the concept task at Session IV will be evidenced by Subgroup 1A as contrasted with Subgroup 2A, and the latter group will perform better than Subgroup 3A. The same general trend will be evident with all the B Subgroups. The A Subgroups and B Subgroups are compared separately because they received either a nonprogramed or programed order, respectively, on the Session 4 CL task and thus are not comparable. A regression equation full model including group membership (programed, serial, or no experience) for Subgroup A did not predict the number of correct responses on the Session 4 CL task significantly better than a restricted model without group membership. The $F_{(2,6)} = 3.17$ was not significant for these data, with $R^2 = .51$ for the full model. The same analysis for Subgroup B, those with the nonprogramed order, was significant at $p < .05$ with $F_{(2,6)} = 5.18$, and $R^2 = .63$ for the full model. In this latter case, Groups 1 and 2 performed better than Group 3, much as was the case for the first concept task. As was concluded for that analysis, it was not the prior programed experience which made a difference on the concept task performance, it was the existence of any prior experience on a MD task. The fact that the second hypothesis was confirmed for the nonprogramed order but not for the programed order may be due to the greater sensivity to prior experience on the more difficult task, the nonprogramed order; however, the extremely small number of Ss used in the analysis and the fairly sizeable F make the absence of significance questionable.

The third hypothesis indicated that children will perform better on a MD task when they have had prior experience on a CL task. The rationale for this hypothesis is that prior training on a concept task in which, according to the task analysis, the multiple discrimination is involved should facilitate the subsequent acquisition of the multiple discrimination. This will be true even though from a hierarchical standpoint the task sequence is reversed. If criterion acquisition had been the goal of the CL task training, in fact, the MD task should be fully acquired as part of the concept. With respect to Table 1, the third hypothesis indicates that the performance on the CL task of Group 3 at Session 3 will be superior to the performance of Group 1 at Session 1. A regression equation full model with group membership did not predict the dependent variable significantly better than a restricted model without group membership. The $F_{(1,10)} < 1.00$ indicated no differences between the two groups.

Discussion

This study was concerned with the application of a curriculum hierarchy evaluation (CHE) model to tasks selected from an early learning curriculum. The CHE model consisted of an amalgamation of a learning transfer paradigm and an aptitude-treatment-task interaction (ATTI) paradigm such that the treatment dimension of the ATTI paradigm was the condition of experience on the lower level (MD) task and the dependent variable was a higher level (CL) task. Additionally, the aptitude dimension consisted of performance on an alternate MD task, one which contained elements (labels) which were unrelated to the CL task, and the task dimension consisted of the MD and CL task performance.

The first hypothesis was that the treatment dimension of the CHE model, a programed or nonprogramed order of item presentation on the MD task, would be differentially effective for children of different aptitudes. Although the hypothesized interaction with aptitude was not significant, the main effects for the two orders with aptitude as a covariate showed that a programed order was superior to a nonprogramed order on the MD task in terms of number of correct responses on a defined number of trials. The failure of the hypothesized interaction to appear indicates that across the range of ability level, children standardly profit on an MD task from a programed as contrasted with a serial order of item presentation. The programed sequence was also more effective than a nonprogramed sequence in producing learning on a CL task. A covariate derived from a task analysis did not in the present study correlate highly enough with the CL task to serve as an effective aptitude in the ATI. The results presented here nevertheless extend O'Malley's (1970) previous report to include a CL task and a covariate on the MD task.

A complicated programed order with eight items did not require the use of a machine to produce low error rate and effective learning for two different kinds of tasks. Rather, the items were administered by the

E after roughly one hour of distributed practice in arranging the cards in their programed order. The presentation by teachers of items in a serial sequence or of so few items that little learning takes place thus should not be construed as an unalterable condition of classroom learning. It would be interesting to determine whether the facilitation of learning via the programed sequence applies to other tasks and aptitudes, or perhaps to a situation with small group rather than individual instruction.

The second hypothesis was that children with superior performance on the MD task perform better related on a CL task since the MD component of the CL task would already be acquired; furthermore, it was predicted that both of these groups perform better on the CL task than a group which had no prior experience on the MD task. Results indicated that the existence of prior MD experience, regardless of the order of item presentation, was the determining factor producing superior performance on the concept learning task. Thus, although the programed order as contrasted with the serial order produced more learning on the MD task, the superior MD learning was not a more effective foundation for the CL task. This may have resulted from the possibility that the MD training with a programed order was sufficient to free the children from attending to the labels during the CL task, but the programed MD training was not sufficient to free the children from other aspects of the CL task. What might be required beyond the MD training is training on the "principle" involved on the first card of the CL pair. The child might be trained to recognize the label SAND for a picture, then be trained to verbalize the principle THE SAND IS DRY upon presentation of the picture, and finally be trained on recognizing the label TREES for a picture. Anticipating the correct principle in the CL task for the picture of trees, THE TREE IS WET, may under these conditions be easier, assuming that the child has actually acquired the concept of opposites, and providing that he has had experience with objects which are physically wet and dry.

The results supported the general hierarchical notion that some form of prior MD experience will facilitate performance on a CL task. Ambiguity in the verification of the hierarchy, however, seemed to be diagnostic of an improper analysis of the processes extant within the higher order task. An adequate task analysis may be prerequisite for identifying the processes operative in performance of any task, and an empirical analysis of the type presented here may be a basis for determining whether or not the task analysis has properly identified the processes involved. That is, training on the processes involved in a task, if those processes have been adequately identified, will result in facilitation of task performance. As Bunderson & Dunham (1970) have phrased it, such training reduces the process requirements of the task.

The third hypothesis was that training on a concept task enhances performance on a multiple discrimination task when the two tasks contain similar elements. The concept task--as indicated above in hypothesis two, which was partially supported--will be more difficult than if the multiple discrimination task had preceded it since the hierarchical sequence is reversed. Nevertheless, if the process of forming a multiple discrimination exists as part of the concept task, training on the concept task should ease the acquisition of the multiple discrimination task. The test of this hypothesis was not significant. There was an average of 25.17 correct responses out of 32 possible Recognition Trial responses for both groups, indicating that substantial learning of the MD task took place under both conditions. Very little CL task learning took place which would facilitate performance on the MD task, however, since there were only 7.00 mean correct responses out of 32 possible Recognition Trial responses. Training may have to be carried to criterion before concept learning could be expected to enhance multiple discrimination learning.

The results of this study indicated that the CHE model will successfully analyze a theoretically-derived curriculum hierarchy within the context of a transfer paradigm. The transfer paradigm has been recommended by Gagné (1970) as well as by Resnick and Wang (1969). The model served as an empirical check upon the task analysis performed upon the concept task insofar as it showed that the concept task used in this study may involve two processes, a multiple discrimination and a principle, rather than the hypothesized single process, a multiple discrimination alone. The ATTI component of the CHE model, as shown in prior research (Dunham & Bunderson, 1970; Tobias, 1969), has considerable potential for identifying aptitude-treatment interactions even though in the present study these interactions were not in evidence. The present study must be considered exploratory, however, in view of the small number of Ss and their unique ethnic characteristics. The most severe restriction of the present CHE model is that it is limited in application to only two tasks in a linear hierarchy. Complex human behaviors are generally analyzed into multiple levels and branched hierarchies. The present model should nevertheless be extended to include different tasks and different aptitudes in anticipation of building a classification of task performances founded upon task analysis and verified by an empirical approach.

References

Bunderson, C. V., & Dunham, J. L. Project on cognitive abilities and learning. Technical Progress Report, Advanced Research Projects Agency. Contract No. N00014-67-A-0126-0006, University of Texas at Austin, Austin, Texas 1970.

Cohen, J. Multiple regression as a general data analytic system. Psychological Bulletin, 1968, 70, 426-443.

Cohen, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), Handbook of clinical psychology. New York: McGraw-Hill, 1965. Pp. 95-121.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Gagné, R. M. Instructional variables and learning outcomes. In M. C. Witrock & D. E. Wiley (Eds.), The evaluation of instruction. New York: Holt, Rinehart & Winston, 1970. Pp. 105-126.

O'Malley, J. M. Stimulus dimension pretraining and set size in learning multiple discriminations with letters of the alphabet. Unpublished manuscript, University of Hawaii, 1970.

Resnick, L. B. Design of an early learning curriculum. Working Paper 16. Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania, 1967.

Resnick, L. B., & Wang, M. C. Approaches to the validation of learning hierarchies. Paper presented at the Eighteenth Annual Western Regional Conference on Testing Problems, San Francisco, California, 1969.

Tobias, S. Research strategy in the effect of individual differences on achievement from programed instruction. Paper presented at the annual meetings of the American Psychological Association, Washington, D. C., September, 1969.

Table 1

Treatment Conditions by Session

Group	Sub-group	Session 1	2	3	4
1	A	SS-MD-S1	PS-MD-S2	PS-CL-S2	PS-CL-S1
	B	SS-MD-S1	PS-MD-S2	PS-CL-S2	SS-CL-S1
2	A	SS-MD-S1	SS-MD-S2	PS-CL-S2	PS-CL-S1
	B	SS-MD-S1	SS-MD-S2	PS-CL-S2	SS-CL-S1
3	A	SS-MD-S1	PS-CL-S2	PS-MD-S2	PS-CL-S1
	B	SS-MD-S1	PS-CL-S2	PS-MD-S2	SS-CL-S1