DOCUMENT RESUME

ED 108 897 SE 019 085

AUTHOR Holliday, William G.; Partridge, Louise A.
TITLE Evaluation of a Practical Procedure for Sequencing Pictorial Classification Tasks in Science.
PUB DATE Mar 75
NOTE 19p.; Paper presented at the annual meeting of the National Asso tion for Research in Science Teaching (48th, Los Angeles, California, March 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Educational Research; Elementary Education; *Elementary School Science; Instruction; Learning Processes; Science Education; *Sequential Learning; *Serial Learning; Symbolic Learning; *Teaching Techniques; *Visual Aids
IDENTIFIERS Research Reports

ABSTRACT

The purpose of this study was to evaluate a practical procedure for determining the optimum sequencing of pictorial classification tasks in science using mean difficulty indices generated by a standard sample of grade two students. Acquisition of such skills can affect the learner's performance on subsequent material and the teacher's efforts in diagnosing specific learner problems. It was predicted that: (1) an ascending order of difficulty in the sequence would be more effective than a random order and (2) a random order of task difficulty would be a more effective sequence than a descending order. Fifty-one grade two subjects constituted the standard sample. The experimental sample was composed of another group of 134 second grade subjects from the same school and from a nearby school. Subsequently, the experimental subjects were randomly assigned to three treatment groups of ascending, randomized and descending orders of task difficulty. The effects of three presentation sequences of the science tasks were experimentally evaluated in terms of total student performance. The results indicated that a highly consistent series of a pictorial classification tasks sequenced in an ascending order of difficulty was more effective than a random order and that a random order was more effective than a descending order of difficulty. (Author/BR)

Evaluation of a Practical Procedure for Sequencing Pictorial Classification Tasks in Science

William G. Holliday
Louise A. Partridge

Univeristy of Calgary
Calgary, Alberta, Canada

A paper presented at the National Association for Research in Science Teaching annual meeting, Los Angeles, March 1975.

Introduction

The sequencing of science materials under instructive and testing conditions represents an area commonly investigated by educational researchers. The practical purpose behind such investigations in science education is twofold. First, selected science materials that are properly sequenced can facilitate the student's efforts in acquiring intellectual skills (Gagne, 1973). Acquisition of such skills can affect the learner's performance on subsequent material. In theory, easier tasks that are followed by more difficult tasks of a similar nature allow the learner to draw upon his memory and also be reinforced by his successes as he proceeds through a sequence of tasks. Secondly, science materials that are properly sequenced can facilitate the teacher's efforts in diagnosing specific learner problems. The purpose of this study is to evaluate a rather practical procedure for determining the optimum sequencing of pictorial classification tasks using mean difficulty indices generated by a standard sample of grade two students. The effects of three presentation sequences of the science tasks were experimentally evaluated in terms of total student performance.

It was hypothesized that multiple classification tasks such as those generally pictured in modern elementary school science programs are most effectively presented in an ascending order of difficulty, as defined by statistically obtained mean item difficulty values from a sample of grade two subjects. Random ordering was predicted to be less than best, while a descending order was expected to constitute the poorest order. Tasks generally associated with these multiple classification skills require the child to attend to two or more stimulus dimensions (e.g., color, shape) in a simultaneous fashion. Children's capacity to perform these tasks have been identified as a key ability directly related to the Piagetian stage of concrete operations (Inhedler and Piaget, 1964), to "non-Piagetian" cognitive development theory (Glaser and Resnick, 1972), and to general non-verbal intelligence (Raven, 1956). It appears logical that learners who encounter an ascending order of difficulty rather than a random or a descending sequence should acquire appropriate problem-solving techniques, subordinate skills and positive anticipations of success that can be expected to facilitate performance on subsequent classification tasks of a highly similar nature. This explanation could be related to work done in set induction and direct transfer effects within learning hierarchies. Interestly, psychometric research on test item sequencing does not currently support predictions of significant differences among the three orders of items sequencing mentioned above.

Method

Subjects

Fifty-one grade two subjects from a Calgary (Alberta, Canada) elementary school constituted the standard sample. The experimental sample was composed of another 134 second grade subjects from the same school and from a nearby school. Subjects ranged in age from seven to eight years and came from middle class, English speaking families.

Materials

Fifty-six multiple classification tasks were developed for this study. Each task required subjects to sort pictures of domestic animals on the bases of two or more stimulus properties (i.e., position, size, shading and size). The selection of the animal classifications were based on familiar variations of typical animal groupings found throughout North America.

The stimulus material for each task consisted of three different positive instances of a class of animals pictorially displayed in the top half of an 8½ x 11 inch sheet of paper. These positive instances varied in terms of two or more stimulus dimensions as defined in Table 1. In addition, three different negative instances and a new positive instance were displayed in the bottom half of the same paper as illustrated in Figure 1. The subject's task was to select the new positive instance from among the four instances. The successful subject selected the new positive instance from among these four choices in the bottom row based on the relative importance of the animal characteristics illustrated in the top row of each paper. The logical hierarchical order of animal characteristics was: shape, shading, size and portion. For example, the stimulus class shown in Figure 1 represents a typical specimen and consists of three furry dogs with shaded ears and spots of varying size. The relevant stimulus dimensions are shape (furry dog) and shading (shaded ears and spots). The irrelevant stimulus dimensions are size and position because more than one value of each is displayed among the positive instances in the top row of this illustrated example. More generally, subjects always had to attend to the shading and position dimensions in some manner throughout the task series.

[TABLE 1 AND FIGURE 1 HERE]

Procedure

Each subject from the standard sample received 56 randomized tasks in an individually sequenced order. Subsequent item analysis of the results yielded task difficulty values (i.e., the percentage of subjects who correctly performed a task) which were used to rank each task. Preliminary observations in the present study and in a highly similar experiment (Friebergs and Tulving, 1961) suggested that children generally begin to perform selected classification tasks with little problem after about the twentieth item. Therefore, the 56 ranked tasks were placed in one of 20-five percentile intervals of mean item difficulty. Subsequently, one task was random selected from each cell which was used in the experimental phase of this study.

The inclusion of the experimental subject's response in the data analysis was contingent upon his ability to perform two preliminary tasks. These plant classification tasks were judged very easy by the experimenters and equivalent in nature to the subsequent animal classification tasks except for the content material. It was assumed that subjects who were unable to perform these two preceding tasks would be unable to perform the subsequent animal classification tasks. The presentation of initially easier tasks is the requirement prescribed by Ebel (1972) in his recommended procedure for sequencing test items.

The experimental subjects were randomly distributed to three treatment groups. The ascending treatment consisted of a task order from the easiest to the most difficult item based on the standard subjects' performance on individual tasks. The randomized treatment consisted of a random ordering of tasks for each subject. The descending treatment consisted of an order opposite to the ascending treatment sequence. The standard and experimental subjects were instructed to pretend that they were selecting different kinds of animals for a children's zoo. They were told to print an "x" on the pictured animal in the bottom row of each paper which was the same kind of animal as those three in the top row of the same paper. A sample task using an analogous geometric stimuli task was displayed on a 20 by 26 inch display. Subjects were told the correct choice by an experimenter. Subsequently, subjects were instructed to perform the two plant classification tasks and then the animal tasks. The experimenters did not encourage subjects to hurry through the task series.

Results

Seventeen items (the first two and last percentile interval of the mean item difficulty range resulted in empty cells) of the 56 normative tasks were used in the three experimental treatments. Four standardized and six experimental subjects failed to perform one or both preliminary tasks and were eliminated from the analysis. Two other experimental subjects were unable to complete the 17 tasks in a reasonable amount of time. Item analysis of collected data from this study is shown in Table 2. The internal consistency reliability (Cronbach's Alpha) was .96 for the standard sample and .72 for the randomized experimental sample of the study. Two orthogonal comparisons of the means using analysis of variance supported the two *a priori* hypotheses as shown in Table 3. As predicted, the ascending treatment scores were significantly higher ($F=14.28$, $df=1;88$, $p<.001$) than the randomized treatment scores which, in turn, were significantly higher ($F=4.68$, $df=1;87$, $p<.05$) than the descending treatment scores.

[Tables 2 and 3 Here]

Discussion

These results suggest that a highly consistent series of a pictorial classification tasks sequenced in an ascending order of difficulty is more effective than a random order and that a random order is more effective than a descending order of difficulty. Consequently, this procedure of sequencing science materials represents a more promising practical approach than many other commonly suggested methods. Subsequent behavioral analyses of the individual tasks and construct validation of the task series should increase its utility in terms of future science education experiments which require multiple classification of consistent, concrete and familiar classroom material. The experimental tasks incorporate more concrete, familiar concepts than the abstract, colored matrices used in the Raven Coloured Progressive Matrices Tests (Raven, 1956). The Raven test is the instrument typically used in laboratories and schools to assess the multiple classification ability and the non-verbal intelligence of young children. The concrete tasks developed for the present study required children to sort pictures of domestic animals on the basis of two or more stimulus properties (i.e., position, shape, size, and shading). The teaching of such concrete intellectual skills has direct extra-school relevance and represents the type of material that should be taught in the elementary schools according to Rohwer (1971). Instructive tasks highly similar to the ones explored in the present study, can be found in Science - A Process Approach (See A.A.A.S., 1967), Science Curriculum Improvement Study (See S.C.I.S., 1968) and Elementary Science Study (See E.S.S., 1968).

Previous classification research in science education has been largely based in Piagetian theory and has examined such topics as the differential effects of varying the response made between the instructive and evaluative components (Popp and Raven, 1972), relative performance by middle and lower socioeconomic class children (Raven, 1968), and practice effects on simple and complex classification rules (Raven, 1970). Multiple classification research more directly relevant to the present study has been based on Piagetian and behavioral analysis theory. Information gained from these studies have mainly dealt with investigation of hierarchical positions and importance of this skill (Kofsky, 1966), modification of children's ability to perform the skill (Klaismeier and Hoover, 1974) and training programs used to facilitate transfer to generalized standardized tasks such as the Raven Colored Progressive Matrices Test (Jacobs and Vandevender, 1971a, 1971b). One of the most interesting studies in this area was reported by Resnick, Siegal and Kresh (1971) in which two double classification tasks of a similar nature were carefully analyzed and compared. Kindergarten subjects learning the easier task first needed fewer trials on the first task than those subjects who learned the tasks in the reverse order. However, subjects who learned the more complex task first were better able to learn the less complex task second. These results were consistent with a science education experiment performed by Raven (1970) in which 12 classification rules were taught to third grade children. Further, discussion will concentrate on the experimental results in light of work done in psychometry, set induction and learning hierarchies.

The effects of varying the order (i.e., ascending, random and occasionally descending) of classroom materials has been explored, in numerous psychometric studies of achievement power (nonspeeded) tests with a consistent result of no significant differences. According to Ebel's (1972) review of this topic, the order of material such as that found in school tests does not make a significant difference in the student's scores as long as the first one or two test items are relatively easy. More specifically, Marso, (1970) in two experiments, found no differences among college subjects using over 100 vocabulary test items with a wide difficulty range. Furthermore, Sax and Cormack (1966) and Munz and Smoure (1968) found no differences between sequences. Sirotnik and Wellington (1974) investigated a similar question using content as an index of order with grade 8 subjects responding on mathematics, science, social science, reading and language arts achievement tests. This was the only study in which reliability coefficients were reported. This statistic is of particular importance when evaluating the potential effects items or tasks have on one another. No ordering effects nor significant differences in estimates of examining percentage correct score, variance estimates of item difficulty, or K-R 20 estimates of internal consistences were found. Furthermore, such similar effects of pictures as stimulus material rather than objects or verbal symbols is unexplored in the psychometric literature. Pictorial stimuli have been used in numerous other science education experiments (Holliday, 1973). However, most of the work has involved instructional variables with secondary school science subjects (See: Holliday, 1975 in press; Holliday and Harvey, in press; Koran, 1973; Wilson, 1973). Therefore, the extent to which the present classification experiment is generalizable to other stimulus modes should also be investigated at all levels of cognitive development.

The differential student performance from the three task orders were likely the result of cognitive and affective set and direct transfer effects from a possible learning hierarchy. The specific contribution of each of these elements to the success of the differential order presentations is currently a matter of conjecture and not a direct concern of this study.

Set induction occurs when a learner becomes predisposed to view and approach tasks in a manner affected by previous experiences. The influence of a cognitive set can have either a positive effect (See Wittlock, 1963; Schuck, 1970, and 1971; and Torrance and Harman, 1961) or a negative effect (See Luchins, 1942, and Birch and Rabinowitz, 1951). Experiments in set induction suggest that subjects, performing a series of internally consistant tasks, are better able to assimilate and organize the stimulus materials (MacDonald, 1965). It is also a well established phenomena that previous successful experiences generate increased levels of expectation and self-confidence which, in turn, can lead to additional success with simiiliar problems (DeCecco, 1968). Children who anticipate success as a result of previous experience seem to increase their chances of success on subsequent similar tasks. Two critical ingredients consistent with cognitive and affective set theory can be found in this study. First, set cannot be induced without consistency among problems. Consistency among the intellectual tasks in this study was clearly established by the extremely high reliability coefficients. Second, different degrees of success were generally experienced by the experimental subjects depending upon their assigned treatment. These two characteristics of the present study support the importance of set theory under conditions of varied sequencing of pictorial classification tasks.

Findings in this study suggest that mean difficulty values obtained from a standard sample population constituted an effective procedure for determining optimum order of pictorial classification tasks in science and might be partly explained in terms of direct transfer effect. The practical consequences of such correlational methods relative to the more common procedure delineated by Gagne (1970) have "suggestive utility for instructional purposes" (Glaser et al, 1972), but their future importance is "unclear" at this time (Gagne, 1973). Such research has mainly dealt with rigorous validation procedures for determining specific relationships among individual tasks (White, 1974a). White (1974b) has outlined a complex model for such a validation procedure. In addition, White (1974c) discussed his method in terms of other indexes evaluated in science education, such as the classical work of Capie and Jones (1971). Of course, the objective in the present study was to evaluate a more practical and less rigorous method of sequencing a given set of tasks and not to evaluate a procedure for validating a learning hierarchy as prescribed by Gagne (1970) and later by White (1974b).

Finally, learning hierarchy research (Gagne,1973) suggests that the sequence of classroom materials is more likely to have practical consequences under three conditions. First, the materials should be associated with intellectual skills or cognitive strategies (as called by some theorists). Second, small units of materials should be evaluated rather than entire school programs or curriculums. Third, learners who are unsophisticated readers or relatively young are more likely to be influenced by differential sequencing of classroom materials. These three elements were contained within the experimental conditions of the present study, which probably increased the change of direct transfer effects. In conclusion, the three presentation orders of difficulty were experimentally shown to have a differential effect on total student performance. These results indicated that the way a teacher sequences a large number of pictorial classification tasks in science can make a real difference.

REFERENCES

American Association for the Advancement of Science, Commission on Science Education, Science A - A Process Approach, Parts A - Lesson 0 and V. New York: Xerox Corporation, 1967.

Birch, H.G. and H.S. Rabinowitz, "The Negative of Previous Experience on Productive Thinking", Journal of Experimental Psychology, 41, 121-125, 1951.

Capie, W. and H.L. Jones, "An Assessment of Hierarchy Validation Techniques", Journal of Research in Science Teaching, 8, 139-147, 1971.

De Cecco, J.P., The Psychology of Learning and Instruction. Educational Psychology, Englewood Cliffs, N.J.: Prentice-Hall Inc., 1968.

Ebel, R.L., Essentials of Educational Measurement, Englewood Cliffs, N.J.: Prentice-Hall Inc., 1972.

Elementary Science Study, Attributes, Games and Problems, Teachers Guide, New York: McGraw Hill, 1968.

Friebergs, V. and E. Tulving, "The Effect of Practice on Utilization of Information from Positive and Negative Instances in Concept Identification", Canadian Journal of Psychology, 15, 101-106, 1961.

Gagne, R.M., The Conditions of Learning , (2nd Edition), New York: Holt, Rinehart and Winston, 1970.

Gagne, R.M., "Learning and Instructional Sequence", in F.N. Kerlinger (Ed.), Review of Research in Education, No. 1, Itasca, Ill: Peacock, 1973.

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

Holliday, W.G., "Critical Analysis of Pictorial Research Related to Science Education", Science Education, 57, 201-214, 1973.

Holliday, W.G., "The Effect of Verbal and Adjunct Pictorial-Verbal Information in Science Instruction, Journal of Research in Science Teaching, Vol. 12, 1975.

Holliday, W.G., "Conceptualizing and Evaluating Learner Aptitudes Related to Instructional Stimuli in Science Instruction", Journal of Research in Science Teaching, in press.

Holliday, W.G. and D.A. Harvey, "Adjunct Labeled Drawings in Teaching Physics to Junior High School Students", Journal of Research in Science Teaching, in press.

Inhelder, B. and J. Piaget, The Early Growth of Logic in the Child, New York: Norton, 1964.

Jacobs, P.I. and M. Vandeventer, "The Learning and Transfer of Double-Classification Skills by First-Graders", Child Development, 42, 149-159, 1971a.

Jacobs, P.I. and M. Vandeventer, "The Learning and Transfer of Double Classification Skills: A Replication and Extension", Journal of Experimental Child Psychology, 12, 240 - 257, 1971b.

Klausmeier, H.J. & F.H. Hooper, "Conceptual Development and Instruction", in F.N. Kerlinger (Ed.), Review of Research in Education, No. 2, Itasca, Ill: Peacock, 1974.

Kofsky, E., "A Scalogram Study of Classificatory Development", Child Development, 37, 191-204, 1966.

Koran, J.J., Jr., "Process Analysis of Aptitude and Treatment Interaction in Learning from Visual Stimuli in Science Education", a paper presented at the annual meeting of the National Association for Research in Science Teaching, Detroit, Michigan, March, 1973.

Luchin, A.S., "Mechanization in Problem-Solving: The Effect of 'Einstellung'", Psychological Monograph, No. 248, 1942.

MacDonald, F.J., Educational Psychology (2nd Ed), Belmont, California: Wadeworth, 1965.

Marso, R.N., "Test Item Arrangement, Testing Time and Performance," Journal of Educational Measurement, 7, 113-118, 1970.

Munz, D.C. & A.D. Smouse, "Interaction Effects of Item-Difficulty Sequence and Achievement-Anxiety Reaction on Academic Performance", Journal of Educational Psychology, 59, 370-374, 1968.

Popp, L. and R. Raven, "The Effects of Response Format of a Structured Learning Sequence on Third Grade Children's Classification Achievement", Journal of Research in Science Teaching, 9, 177-184, 1972.

Raven, J.C., Guide to Using the Coloured Progressive Matrices, Sets A, Ab, B (Revised order, 1956), London: H.J. Lewis & Co. Ltd., 1956.

Raven, R., "The Development of Classification Abilities in Culturally Disadvantaged Children", Journal of Research in Science Teaching, 5, 224-229, 1968.

Raven, R.J., "The Effects of a Structured Learning Sequence on Second and Third Grade Children's Classification Achievement", Journal of Research in Science Teaching, 7, 153-160, 1970.

Resnick, L.B., A.W. Siegel and E. Kresh, "Transfer and Sequence in Learning Double Classification Skills", Journal of Experimental Child Psychology, 11, 139-149, 1971.

Sax, G. and T.R. Cormack, "The Effects of Various Forms of Item Arrangements on Test Performance", Journal of Educational Measurement, 3, 309-311, 1966.

Science Curriculum Improvement Study, Organisms, Teachers Guide, Boston: D.C. Heath, 1968.

Schuck, R.F., "The Influence of Set Induction Upon Pupil Perception of Effective Teaching, Achievement, and Retention in Units on Respiration and Circulation in the BSCS Curricula", Journal of Research in Science Teaching, 8, 51-56, 1971.

Schuck, R.F., "The Influence of Set Induction Upon Student Achievement and Perception of Effective Thinking", Journal of Research in Science Teaching, 7, 35-40, 1970.

Sirotnik, K. and R. Wellington, "Scrambling Content in Achievement Testing: An Application of Multiple Matrix Sampling in Experimental Design", Journal of Educational Measurement, 179-188, 1974.

Torrance, E.O. and J. Harmon, "Effects of Memory, Evaluative, and Creative Reading Sets on Test Performance", Journal of Educational Psychology, 52, 207-214, 1961.

White, R.T., "The Validation of Learning Hierarchy", American Educational Research Journal, 11, 121-136, 1974a.

White, R.T., "A Model for Validation of Learning Hierarchies", Journal of Research in Science Teaching, 11, 1-4, 1974b.

White, R.T., "Indexes Used in Testing the Validity of Learning Hierarchies", Journal of Research in Science Teaching, 11, 61-66, 1974c.

Wilson, J.T., "The Effects of Individual Differences on Learning from Written Materials: The Control of Inspection Behavior by Test-Like Events", unpublished doctoral dissertation, University of Florida, Gainesville, Florida, 1973.

Wittlock, M. "Effect of Certain Sets Upon Complex Verbal Learning", Journal of Educational Psychology, 54, 85-88, 1963.

TABLE 1

A DESCRIPTION OF THE FOUR DIMENSIONS ALONG WHICH A STIMULUS COULD VARY AND THE PROPERTIES OF EACH DIMENSION

Position	Size	Shading	Shape
The possible positions were: 1. Sitting or standing and a. facing left b. facing right c. facing forward d. facing backward	The possible sizes were: 1. The original size. 2. 20% smaller than the original size. 3. 40% smaller than the original size.	A stimulus could have: 1. No shading. 2. Complete shading with: a. Letratone #42 which is very dark. b. Letratone #137 which is lighter shading with a lined appearance. c. Letratone #98 which is very light shading. 3. Partial shading with any of the above. 4. Spots - either several small ones or a few large ones. 5. Irregular stripes.	The possible shapes were: 1. Catfish. 2. Plain fish. 3. Furry dog. 4. Plain dog. 5. Frog. 6. Cat. 7. Cow. 8. Horse. 9. Duck. 10. Chicken.

TABLE 2

ITEM DIFFICULTY LEVELS, POINT BISERIAL CORRELATIONS, AND RELIABILITY COEFFICIENTS FOR THE SEVENTEEN SELECTED ITEMS ON ALL FOUR FORMS OF THE INSTRUCTIONAL MATERIAL

(Column A: Point Biserial Correlations)
(Column B: Item Difficulty Levels)

Item Identification Number	Standardized Sample		Experimental Treatments: Ascending		Randomized		Descending	
	A	B	A	B	A	B	A	B
56	.32	.94	1.00	1.00	.27	.93	.28	.86
3	.28	.90	1.00	1.00	.55	.93	.10	.95
25	.33	.84	.30	.93	.42	.84	.68	.68
18	.52	.76	.05	.96	.23	.96	.49	.80
42	.47	.73	.41	.80	.48	.73	.47	.59
26	.68	.69	1.00	1.00	.49	.82	.70	.73
48	.68	.61	.45	.91	.55	.87	.71	.61
2	.73	.57	1.00	1.00	.46	.80	.78	.50
5	.44	.55	.41	.80	.55	.76	.49	.45
21	.55	.47	.63	.80	.62	.64	.58	.48
32	.73	.45	.53	.84	.51	.53	.73	.36
16	.43	.39	.63	.76	.56	.51	.72	.45
31	.14	.31	.26	.60	.47	.49	.61	.64
53	.58	.27	.23	.40	.41	.29	.61	.27
34	.41	.22	.41	.47	.21	.22	.29	.20
40	.45	.18	.26	.11	.08	.09	.34	.14
7	.40	.14	.23	.04	.24	.07	.30	.16
Reliability	.96		.48		.72		.85	

TABLE 3

Observed Means and Standard Deviations on the

Experimental Group Scores

Treatment Groups	n	X	S.D.
Ascending	45	12.47	1.93
Randomized	45	10.51	2.89
Descending	44	8.89	4.10

FIGURE 1

A DIFFICULT ITEM FROM THE

MULTIPLICATIVE CLASSIFICATORY PROGRAM