DOCUMENT RESUME

ED 037 344 SE 008 007

AUTHOR Doran, Rodney L.; Pella, Milton O.
TITLE Development of Test Items Related to Selected Concepts Within the Scheme the Particle Nature of Matter.
INSTITUTION State Univ. of New York, Buffalo. Dept. of Instruction.; Wisconsin Univ., Madison. School of Education.
PUB DATE [70]
NOTE 12p.; Paper presented at Annual Meeting of the National Association for Research in Science Teaching (43rd, Minneapolis, Minne., March 5-8, 1970)

EDRS PRICE EDRS Price MF-$0.25 HC-$0.70
DESCRIPTORS *Concept Formation, *Elementary School Science, *Evaluation, Evaluation Criteria, Evaluation Methods, *Physical Sciences, *Test Construction

ABSTRACT

The purpose of this study was to develop tests items with a minimum reading demand for use with pupils at grade levels two through six. An item was judged to be acceptable if the item satisfied at least four of six criteria. Approximately 250 students in grades 2-6 participated in the study. Half of the students were given instruction to develop the ability to use concepts from the particle nature of matter. The test items developed were designed to assess this ability. Test items pertaining to concepts within a unit were administered to each of the instructed classes on the day following completion of instruction on the last concept. The same test items were administered to each of the noninstructed classes a day or so later. Although only two items met all six criteria, 104 items met at least four of the six. Items related to each of the 25 concepts that met the established requirements were developed. (BR)

DEVELOPMENT OF TEST ITEMS RELATED TO SELECTED CONCEPTS WITHIN THE SCHEME THE PARTICLE NATURE OF MATTER

Rodney L. Doran
Department of Instruction
State University of New York
at Buffalo
Buffalo, New York

Milton O. Pella
School of Education
University of Wisconsin
Madison, Wisconsin

THE PROBLEM

Introduction

Concept development is presently a popular goal in science instruction at all levels. Knowledge of and ability to use science concepts has become one of the essential outcomes of science instruction at the elementary school level. The determination of how well elementary school pupils learn these concepts in and of science is consequently an essential part of good science teaching. These techniques have not been widely available up to the present. "Quite candidly, it must be admitted that our present procedures of evaluation are still in the trial and error stage" (Lee, 1967, p.69). "The improvement of evaluative tools and techniques in elementary science instruction is one of the major priorities of science education" (Kuslan & Stone, 1968, p. 228).

Statement of the Problem

The purpose of this study was to develop test items with a minimum reading demand related to selected concepts from the scheme, the particle nature of matter, for use with pupils at grade levels two through six. An item was judged to be acceptable if the item satisfied at least four of the following criteria:

1. The proportion of the instructed group choosing the accepted response to each item, when administered as part of a posttest, is greater

than 0.50, the level attributable to random guessing.

2. The proportion of the instructed group choosing the accepted response to each item, when administered as part of a posttest, is greater than when administered as part of a pretest.

3. The proportion of the instructed group choosing the accepted response to each item, when administered as part of a posttest, is greater than the proportion of the noninstructed group choosing the accepted response to the same item when administered as part of a posttest.

4. The proportion of each instructed class choosing the accepted response to each item, when administered as part of a posttest, increases progressively with grade level.

5. The item when administered as part of a posttest to the instructed group as of nonextreme difficulty. ($-2\sigma \leq X_{50} \leq -2\sigma$)

6. The item when administered as part of a posttest to the instructed group is a positive discriminator. ($\beta \geq +0.30$)

Criteria for Test Items

The constructs underlying the first four criteria are self-evident. The last two criteria are based on the fact that both the difficulty and discrimination of test items affect their usefulness. An item so easy that all answer it correctly or so difficult that none answer it correctly yields no information about relative levels of achievement. Items of low discrimination are not useful in differentiating between those students who possess the tested attribute and those who do not. The difficulty and discrimination values used in this study were X_{50} and β respectively.

PROCEDURE

Definition of Educational Objectives

The objective of the instruction related to this study was to aid students in developing the ability to use selected concepts from the scheme, the particle nature of matter, in explaining natural phenomena. The test items developed in this study were designed to assess this ability with students at grade levels two through six.

An assumption underlying this study was that the prepared instructional sequence, as presented by the regular classroom teacher is effective in increasing the ability of students at grade levels two through six to use the selected concepts in explaining natural phenomena.

Construction of the Test Items

The item used in this study was of the alternate response pictorial item type presented via motion picture film. Alternate response items were chosen because they were adequate for the purpose and feasible for the motion picture medium. Pictorial items were chosen so as to (1) reduce the demands on the reading and verbal ability of the students, and (2) assess the ability to use the concepts rather than reading or verbal ability. As motion was a relevant factor in some items, motion picture film was used to present the questions and model choices.

Five separate test items were developed for each of the 25 concepts chosen for the study,

> If only one item per objective is used, the test will generally be less reliable than if more items are used. On the other hand, it may be assumed that, if a pupil can perform well on five or 10 items of similar content, he can perform well on 20 or 50 such items (Lindvall & Cox, 1969, p. 175).

Because testing using motion pictures was unfamiliar to most of the students, extensive directions were prepared on how to mark the test booklet.

Selection of the Population

The criteria for population selection was that the school have two or more classes per grade and that the students had been randomly assigned to their class. The first criterion was established because the design of the study included an instructed and noninstructed class at each grade level.

The second criterion was used to insure that the classes would be as similar as possible. Student date, such as IQ and scores on standardized

tests were obtained from the school files to be used in verifying the equality of the instructed and noninstructed classes. The IQs were obtained from the Kuhlman-Anderson Test and the level of academic achievement by use of the Stanford Achievement Test (SAT). The Mathematical Concept, Paragraph Meaning, and science subtests of the SAT were chosen as being representative of abilities in mathematics, reading, and science.

A Wisconsin school was selected to participate in the study as it satisfied the above criteria and the faculty was willing to participate in the program. The principal selected the teachers who would be in the instructed and noninstructed classes.

Collection of Data

Instruction on the concepts was initiated with the instructed classes after all the test items had been administered to all classes (both instructed and noninstructed). Each concept was the subject of a 30 minute lesson for one or two days of instruction by the regular classroom teachers. All of the instructed classes were taught the same lesson on the same day.

On the day following completion of instruction on the last concept in a given unit, test items pertaining to concepts within that unit were administered to each of the instructed classes. The same test items were administered to each of the noninstructed classes a day or so later. The test items were administered by the researcher to each class in its regular classroom. The verbal part of each question was read aloud by the researcher as the students viewed the written questions projected on the screen and printed on the test booklet.

RESULTS

Introduction

The results of this study are presented in the following order: (1) selected concepts, (2) population, (3) item results for a representative concept, and (4) questions and model choices for the concept in (3).

Selected Concepts

The concepts from the scheme, the particle nature of matter, selected for this study are listed below. The order indicated was the order in which the concepts were taught.

Unit I

1. All matter is made up of particles.
2. Particles of matter have spaces between them.
3. Particles of matter are very small.
4. Particles of matter are in motion.
5. Particles of matter move faster when the matter is heated.
6. Particles of matter move farther apart when the matter is heated.

Unit II

7. In the solid state, the particles of matter are packed together in a pattern and move within a small space.
8. In the liquid state, the particles of matter are loosely clustered together and move about.
9. In the gas state, the particles of matter are far apart and move freely.
10. The state of matter can be changed from solid to liquid and from liquid to solid.
11. The state of matter can be changed from liquid to gas and from gas to liquid.

Unit III

12. The push against the surface by a gas depends upon the number and rate of motion of the particles of the gas.
13. Some particles of matter (molecules) are made up of simpler particles (atoms).
14. Some molecules are made up of only one kind of atom (element).
15. Some molecules are made up of two or more kinds of atoms (compounds).
16. Some samples of matter contain more than one kind of molecule (mixtures).

Unit IV

17. Each type of molecule is formed from definite numbers and kinds of atoms.
18. Particles of matter have mass and occupy space.
19. The average size and mass of the atoms of each element do not vary.
20. Atoms are made up of particles; protons, electrons, and neutrons.
21. Electric charges are associated with the particles of matter.

Unit V

22. The particles of matter attract each other.
23. The mass of an atom is determined by the number and kind of particles that it contains.
24. All atoms of a given element are made up of the same number of electrons and protons.
25. Molecules are made up of atoms that are held together by electrical forces.

Population

The numbers of students in the instructed and noninstructed classes completing all the test items, when administered as part of a pretest and as part of a posttest, are given in Table 1.

TABLE 1

NUMBERS OF STUDENTS IN THE INSTRUCTED AND NONINSTRUCTED CLASSES COMPLETING THE TEST ITEMS BY GRADE LEVEL

Grade Level	Instructed		Noninstructed	
	Pre	Post	Pre	Post
2	26	26	23	23
3	23	22	24	24
4	27	26	27	27
5	26	25	27	26
6	26	25	25	24
	128	124	126	124

The instructed and noninstructed classes at a given grade level were assumed to be comparable because the students were assigned to classes irrespective of their academic ability. The credibility of this assumption was testable within the second and fifth grades as current standardized test data were available for these grades. The assumption was reasonable as none of the pairs of classes were significantly different (per t-tests, $\alpha = 0.01$) with respect to IQ and mathematical, reading and science abilities as measured by the subtests of the Stanford Achievement Test.

Item Results

Due to space limitations, one concept (No. 11) was chosen as an example. The results related to the items that tested Concept No. 11 are shown in Tables 2, 3, and 4. Table 2 includes the proportion of each of the classes choosing the accepted response to each item when administered as part of a pretest and as part of a posttest. Presented in Table 3 are the item statistics (X_{50} and β) for each item when it was administered as part of a posttest to the instructed group. Table 4 indicates which items satisfied the several item criteria. It is noted from Table 4 that each item related to Concept No. 11 satisfied five criteria and failed to meet Criterion #4.

Questions and Model Choices

The questions and model choices for the items related to Concept No. 11 are illustrated on Page 9. The models for items II-11, II-18, and II-20 are dynamic; the particles move and change their configurations (i.e. particle configurations change from a gas configuration to a liquid configuration).

SUMMARY

Examination of the results of the data, from the frame of reference of "which items meet four of the six criteria" and hence make up a pool of acceptable items, reveals that although only two items meet all six criteria, there are 104 items that meet at least four of the six (Table 5).

TABLE 2

PROPORTION OF THE INSTRUCTED AND NONINSTRUCTED GROUPS CHOOSING THE ACCEPTED RESPONSE TO THE TEST ITEMES RELATED TO CONCEPT NO. 11

Item Number		Noninstructed	Instructed					
				Grade Level				
		All Grades	All Grades	2	3	4	5	6
II-3	Pre	.40	.44	.31	.35	.63	.58	.31
	Post	.47	.56	.46	.82	.58	.48	.52
II-9	Pre	.68	.73	.62	.70	.70	.65	.96
	Post	.69	.85	.92	.82	.69	.92	.92
II-11	Pre	.56	.53	.46	.30	.48	.58	.81
	Post	.61	.73	.62	.59	.73	.76	.96
II-18	Pre	.83	.74	.69	.70	.67	.73	.92
	Post	.88	.94	.88	.86	.92	1.00	1.00
II-20	Pre	.51	.66	.62	.48	.63	.65	.88
	Post	.60	.77	.46	.58	.77	1.00	1.00

TABLE 3

STATISTICS OF THE TEST ITEMS RELATED TO CONCEPT NO. 11 WHEN ADMINISTERED TO THE INSTRUCTED GROUP AS PART OF A POSTTEST

Item Number	X_{50}	β
II-3	-0.26	+0.83
II-9	-1.28	+1.46
II-11	-0.72	+1.79
II-18	-1.94	+1.26
II-20	-0.75	+4.14

TABLE 4

EVALUATION OF THE TEST ITEMS RELATED TO CONCEPT NO. 11

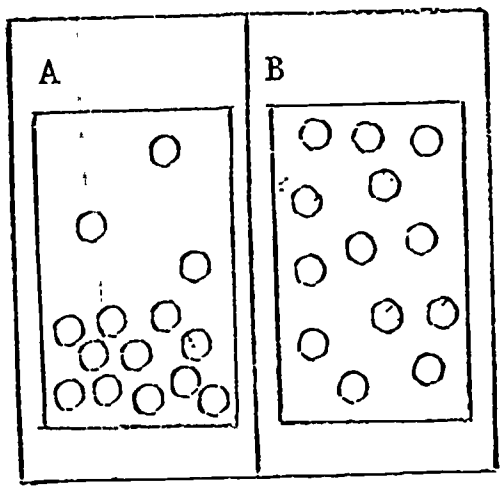
Item Number	Criterion Number					
	1	2	3	4	5	6
II-3	*	*	*		*	*
II-9	*	*	*		*	*
II-11	*	*	*		*	*
II-18	*	*	*		*	*
II-20	*	*	*		*	*

* Indicates the criterion was satisfied

Questions and Model Choices for Concept No. 11

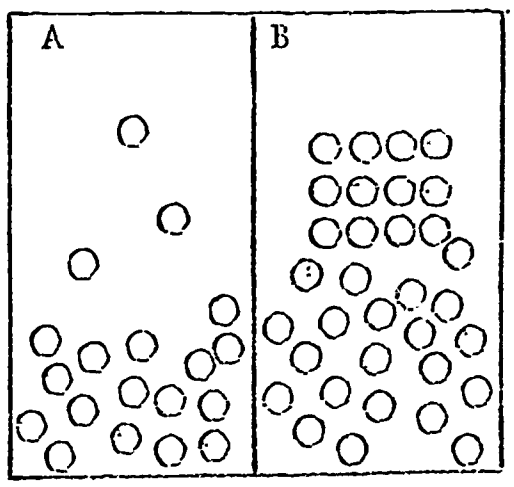
II-3

Which model is more useful in explaining boiling?



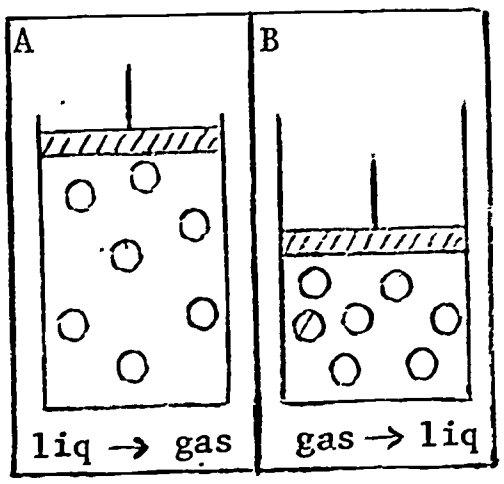
II-9

Which model is more useful in explaining a liquid changing to a gas?



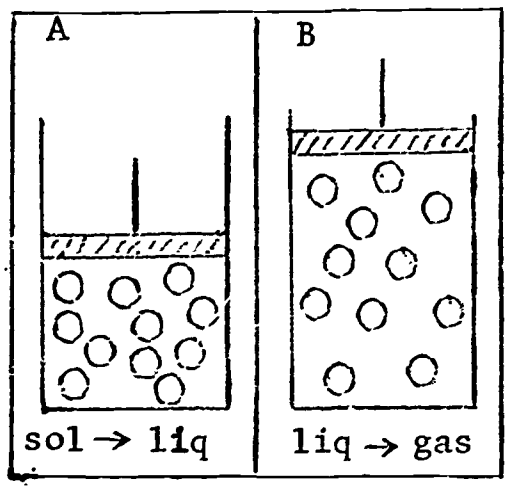
II-11

Which model is more useful in explaining a gas changing to a liquid?



II-18

Which model is more useful in explaining what happens after a liquid is heated until it becomes a gas?



II-20

Which model is more useful in explaining what happens when a gas is cooled until it becomes a liquid?

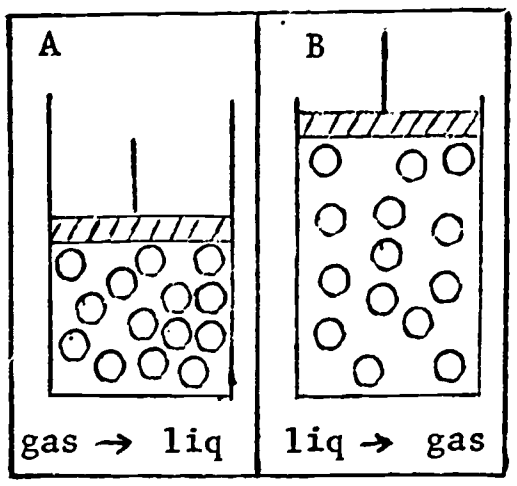


TABLE 5

FREQUENCY DISTRIBUTION OF ITEMS ACCORDING TO TOTAL NUMBER OF CRITERIA SATISFIED

Total Number of Criteria Satisfied	Number of Items
6	2
5	71
4	31
3	17
2	4
1	0
0	0

The one criterion not met by most items was #4 (Table 6). It seems that the items either do not differentiate between levels of understanding of the scheme, the particle nature of matter, or the instructional sequence has not extended beyond the levels possible for children in the lower grades.

TABLE 6

FREQUENCY DISTRIBUTION OF ITEMS THAT SATISFIED EACH CRITERIA ARRANGED ACCORDING TO CRITERION NUMBER

Criterion Number	Number of Items
1	108
2	104
3	106
4	2
5	111
6	119

It is noted from Table 7 that the mode of the item X_{50} distribution is in the range -1.00σ to -0.50σ and the items are distributed in an approximately "normal" manner. A range of X_{50} values is often stated as a desired goal for the items within a test or an item pool, when one wishes to maximally discriminate among all students.

TABLE 7

FREQUENCY DISTRIBUTION OF ITEMS ACCORDING TO X_{50} RANGE

Range of X_{50} (σ units)	Number of Items
$-\infty$ -- -2.00	8
-2.00 -- -1.50	16
-1.50 -- -1.00	22
-1.00 -- -0.50	33
-0.50 -- -0.00	22
0.00 -- +0.50	12
+0.50 -- +1.00	4
+1.00 -- +1.50	2
+1.50 -- +2.00	0
+2.00 -- $+\infty$	0

It is noted from Table 8 that that the frequency distribution of items according to the range of β values, an index of item discriminating power, is positively skewed and its mode is the range of +1.00 to +1.25. The high β values indicate that the items are effective in discriminating between the students who possess the tested ability and those who do not.

TABLE 8

FREQUENCY DISTRIBUTION OF ITEMS ACCORDING TO β RANGE

Range of β	Number of Items
2.50 -- ∞	5
2.25 -- 2.50	3
2.00 -- 2.25	5
1.75 -- 2.00	1
1.50 -- 1.75	16
1.25 -- 1.50	20
1.00 -- 1.25	29
0.75 -- 1.00	25
0.50 -- 0.75	14
0.25 -- 0.50	1
0.00 -- 0.25	0

When using the item pool developed here, care must be exercized to select items according to concept if the item is to determine levels of achievement within the scheme, the particle nature of matter. A total of 104 acceptable items (those that satisfied four at the six criteria) were developed in this study. Items related to each of the 25 concepts that meet the established requirements were developed and in most cases more than one item is available. It is recommended that more than one item be selected for each concept to increase the precision of measurement

REFERENCES

Kuslan, Louis & Stone, A. Harris Teaching Children Science: An Inquiry Approach. Belmont, California: Wadsworth Publishing Company, 1968.

Lee, Eugene New Developments in Science Teaching. Belmont, California: Wadsworth Publishing Company, 1967

Lindvall, C. M. & Cox, Richard C. The role of evaluation in programs for individualized instruction. In Ralph W. Tyler (Ed.) Education Evaluation: New Roles, New Means, the sixty-eighth yearbook of the National Society for the Study of Education. Chicago, Illinois: National Society for the Study of Education.