DOCUMENT RESUME

ED 182 164 SE 029 945

AUTHOR	Linn, Marcia C.; Levine, David I.
TITLE	Adolescent Reasoning: The Development of Ability to Control Variables. Advancing Education Through Science-Oriented Programs.
INSTITUTION	California Univ., Berkeley. Lawrence Hall of Science.
SPONS AGENCY	National Science Foundation, Washington, D.C.
REPORT NO	AESOP-PSc-9
PUB DATE	[76]
NOTE	40p.; For related documents, see SE 029 946-947 and SE 029 877-881
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Cognitive Development; Educational Research; Elementary School Science; Elementary Secondary Education; Individual Development; *Individual Differences; *Junior High Schools; Physics; *Problem Solving; Science Education; Science Experiments; *Science Instruction; Secondary School Science; *Thought Processes

ABSTRACT

A developmental investigation of the ability to control variables was conducted. One or two logically similar physics problems were individually administered to 120 subjects age 12 to 14. The problems involved either familiar or unfamiliar variables. Each problem was presented in three different informational formats. Success on the problems in the various formats ranged between 5% and 95%. Both familiarity of the variables and format of the question influenced success. A qualitative change in ability to control variables between 12 and 16 was observed. Results are discussed in relation to Piagetian theory and to various information processing conceptualizations. (Author)

ADVANCING EDUCATION THROUGH SCIENCE-ORIENTED PROGRAMS, Report PSc-9

Adolescent Reasoning: The Development of Ability to Control Variables

Marcia C. Linn

University of California, Berkeley

David I. Levine

University College, London

Development of Ability to Control Variables

Abstract

A developmental investigation of the ability to control variables was conducted. One or two logically similar physics problems were individually administered to 120 subjects age 12 to 14. The problems involved either familiar or unfamiliar variables. Each problem was presented in three different informational formats. Success on the problems in the various formats ranged between 5 and 95 percent. Both familiarity of the variables and format of the question influenced success. A qualitative change in ability to control variables between 12 and 16 was observed. Results are discussed in relation to Piagetian theory and to various information processing conceptualizations.

Adolescent Reasoning: The Development of Ability to Control Variables

Introduction

The ability to control variables, to keep all but one variable the same so the effect of the variable that is changed can be investigated, is relevant to many everyday decisions. For instance, one must control for container size when comparing prices. School curricula often emphasize recognition of variables (see, for example, Science Curriculum Improvement Study Energy Sources, 1971) and recent research has been carried out to investigate ways to teach children to control variables (Linn, Chen, & Thier, in press). In order to improve programs teaching the control of variables, more information is needed about how this logical ability develops. The current research was undertaken to gain a clearer understanding of the development of the ability to control variables.

Inhelder and Piaget (1958) conducted the first major investigation of adolescent thinking. Part of their research involved controlling variables, or the schema Piaget calls "all other things equal." By conducting experiments using fifteen varied tasks, they found that learners progress from what they call concrete operations to what they call formal operations during adolescence. Tasks which involved controlling variables revealed that the concrete operational learner is able to separate variables and do controlled experiments when working with the apparatus for an experiment. The formal operational learner, on the other hand, can separate and control variables even when presented with an unfamiliar problem and without apparatus. Inhelder and Piaget characterize this transition from concrete to formal as a change from exploring real results to anticipating possible results.

Since Inhelder and Piaget's comprehensive work on logical thinking was published, a number of researchers have investigated particular questions suggested by their research. Both the pervasiveness and the existence of formal operations have been questioned. Many researchers, using either experiments devised by Inhelder and Piaget or other similar experiments, have found that only a fraction (30 to 50 percent) of adolescents and adults reach Piaget's level of formal operations (see, for example, Lovell, 1961; Jackson, 1965; Karplus and Karplus, 1972).

Additionally, researchers have noted that subjects do not perform uniformly on tasks that could require formal operations. Children may perform one task at the formal level and the next at a less sophisticated level. White (1965) has postulated a performance mode of least mental effort. It is possible that both adolescents and adults could use formal operations more frequently than they actually do but rarely find it necessary. Several factors including the role of past experience and the mode of presenting the task have been shown to influence performance (see, for example, Jackson, 1965; Lunzer, 1965; Bryant, 1974; Levine & Linn, 1976 ; Linn & Thier, 1975). A clearer understanding of how children process questions about controlling variables and validation of possible theoretical explanations for changes in ability to control variables are needed.

This paper focuses on two important issues related to ability to control variables. The first is the kind of information presented in the question. Our concern in this case was whether the information presented in a question would influence the likelihood of the question being answered correctly. To investigate this aspect, we posed a problem with several different variables and asked questions which included three different kinds of information. The three questions and the information in each were:

(1) Free Response Question (Free): Subjects were simply asked to propose an investigation of the effect of a particular variable in the problem. No further information was given.

(2) Multiple Choice Question (MC): Subjects were asked to help design an experiment to investigate a particular variable in the problem. Several alternative approaches were proposed. The subject chose the one he preferred and explained his choice.

(3) Screen Question (Screen): Subjects were shown the results of an experiment about one variable while a screen concealed the exact procedure. Subjects were then asked what could be found out from the part of the experiment they had seen.

The second issue we investigated was the role of the subjects' familiarity with the variables and apparatus used in the task. We investigated this in two ways:

(1) Three problems were compared which were of varying degrees of familiarity to the subjects;

(2) For one of the problems (Ramp) we assessed subjects knowledge of the variables involved.

Methods

Subjects

The 120 subjects came from a large comprehensive school (for 11- to 17-year-olds) in a middle class suburban area of London. Forty subjects, half boys and half girls, came from each of three age groups: 11-9 to 12-8 (12-year-olds); 13-10 to 14-7 (14-year-olds); 15-9 to 16-8 (16-year-olds).

The school has ten ability groups at each age level. At each age, two subjects were selected from each ability group. Only students who were taking a science course were selected. Science is a required course for all 12- and 14-year-olds and about ninety percent of the 16-year-olds in each ability group elect to study science. Teachers in the school report that 16-year-olds who study science do not differ greatly from 16-year-olds who omit science from their program. Subjects were selected randomly from the approximately 90% who volunteered to participate.

Summary of Procedure

Subjects were randomly assigned to one of two interviewers who randomly selected one of two 20-minute interviews: (1) the Ramp Problem (Ramp) or (2) the Circuit Problem(Circuit) and the Seed Problem (Seeds). Each problem required the subject to experiment with three variables. The variables for each problem will be described below.

Insert Figure 1

about here

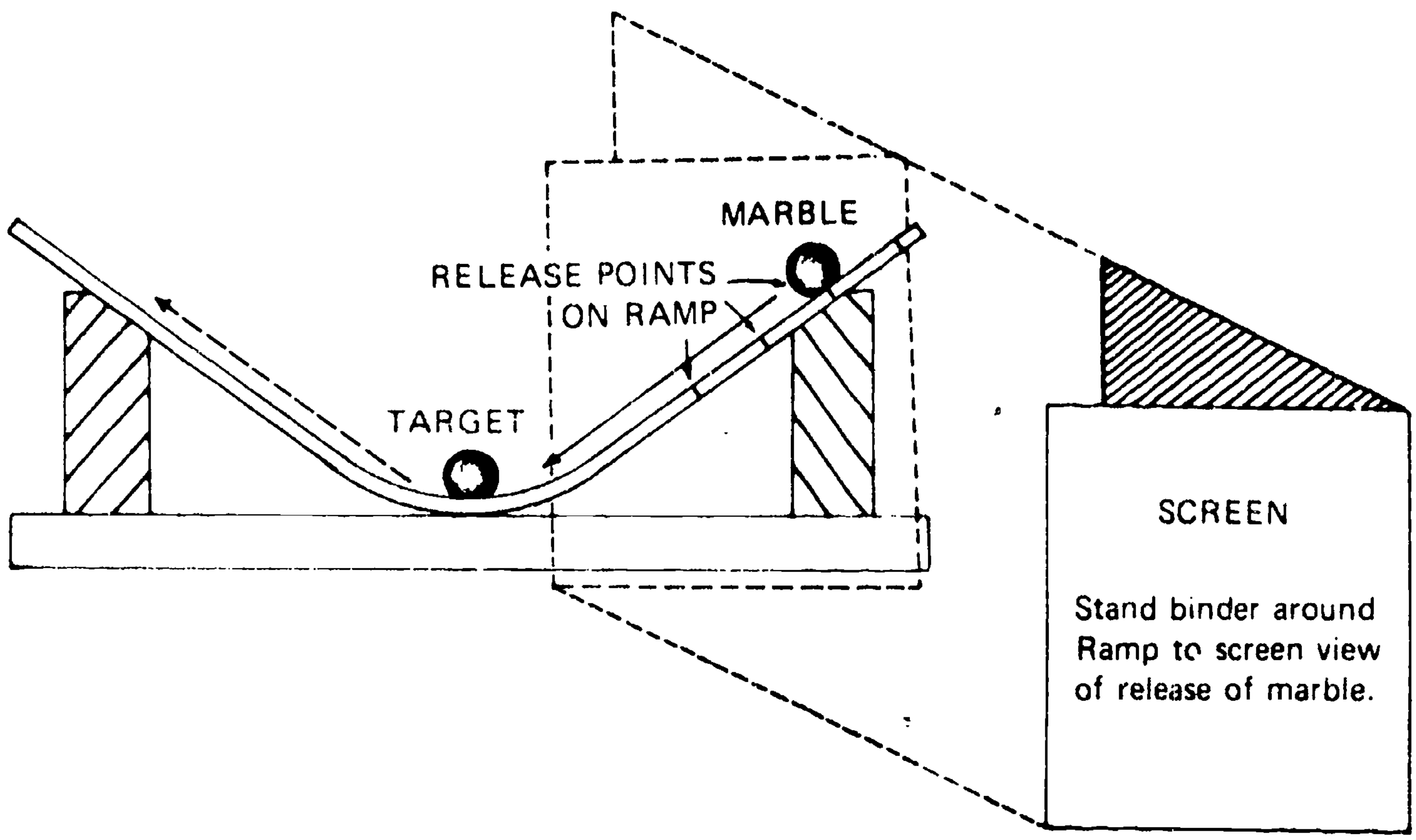
Ramp Problem

The apparatus consisted of a curved ramp made of 1 cm. wide aluminum channelling, and glass or steel marbles of different sizes and colors (Figure 1). For Ramp, the variables were height of the release point, weight of the marble, and weight of the target.

Ramp-Free Response Question (Free). The interviewer explained the apparatus and then introduced the Ramp Problem and Free Question by saying: "Suppose you have been challenged to a marble rolling contest with a boy/girl in your class. He/she will choose the position that both of the marbles in this box will be released from. [Take out closed box with steel and blue marbles inside.]

"You can decide which of the two marbles he/she will use and which one you will use. [Sex was chosen to agree with that of the student.]

Figure 1 Apparatus for Ramp Problem



Before the contest, you can try each marble once to find out which marble is best. For these trials, you can use whatever positions you want, but remember, for the contest, your friend is going to choose one position for both marbles. What would you do to decide which marble to choose?

Which position would you use for the first marble? And for the second?

[Repeat solution.] Why did you choose these positions?

After the subject suggested a way to test the marbles, the interviewer showed the marbles to the subjects and used one of two probes: (1) If the subject said to test the marbles from the same position, the experimenter asked if it would be just as good to test from unequal positions, (2) If the subject wanted to test the marbles from unequal positions, the experimenter asked about using the same position.

Ramp - Variables. Subsequently, subjects were asked eight questions to measure their knowledge about the three basic variables involved: weight of the marble, weight of the target, and height of the release point.

Ramp - Multiple Choice Question (MC). A picture of the large marble ready to be released from position 2 to hit a blue target and four choice pictures of possible experiments with the ramp, target, and marble, drawn on index cards, were placed directly in front of the subject. The choice pictures showed large or small marbles ready to be released from position 2 or 4 to hit either the blue target or another target. Each subject was then told: "Another child in your class has also been challenged to a marble rolling contest, but the marbles will be different. He tried the large marble from position 2 with the blue target. He doesn't know which of these other experiments to do to find the best marble for the contest. Can you tell which would be the test to do to find out which marble to use in the contest?" After the subject chose a response, he was asked why he chose that one and why he did not choose each of the others.

Ramp - Screen. A binder was placed over half of the ramp as in Figure 1. The subject could see how far up the ramp the target went when it was hit by each marble in turn but not where the marbles were released from. (In fact, two different release points were used such that the smaller marble hit the target farther.) The subject was then asked: "Which marble hit the target the farthest? From what you just saw, do you think each marble was released from the same position?" The screen was removed and the interviewer said: "Now I'll show you what I did. I released the clear marble from here (high position) and the other from here (low position). What can you find out from this? Do you know which marble makes the target go farther all the time? Why? Will the clear one still win when they are released from the same position?"

- - - - - - - - -

Insert Figure 2

about here

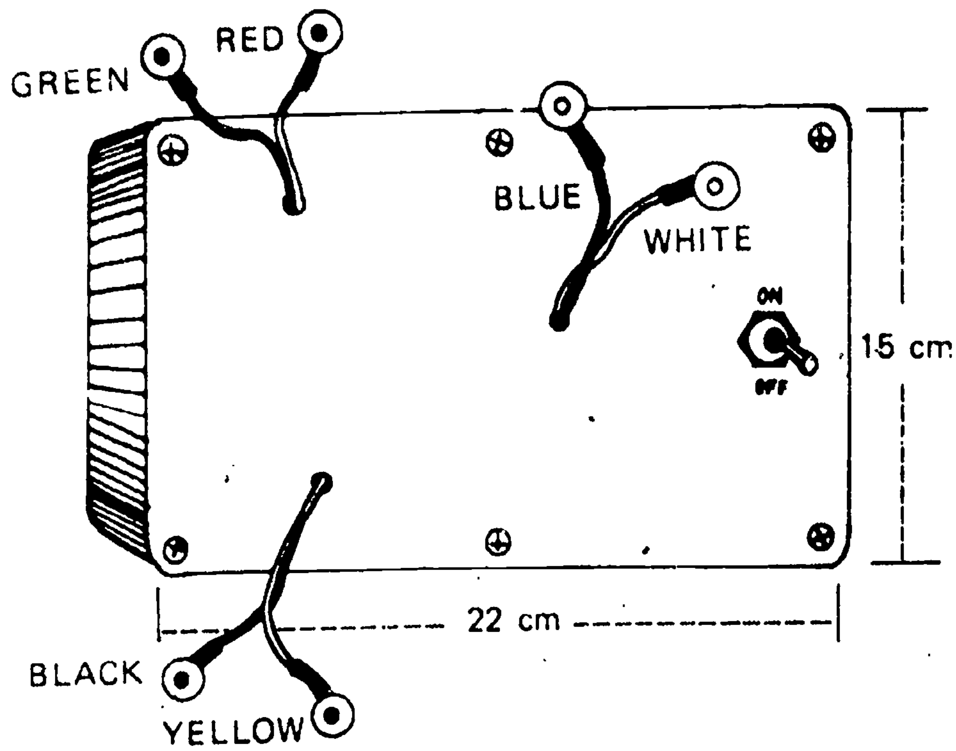
- - - - - - - - -

Circuit Problem

The apparatus for the circuit problem consisted of a metal box (Figure 2) with a labelled on-off switch, and three pairs (or "trees") of wires on the top. Each pair came out of its own hole. For Circuit the three variables were the trees: the blue-white tree, the green-red tree, and the black-yellow tree. The circuit was wired so that it could not buzz unless the one critical wire was selected as part of a three wire group. The interviewer introduced the apparatus. To explain the problem he said:

"Here I have a box that buzzes--sometimes. Now when I turn it on it doesn't buzz but sometimes it does. Just one of these six wires is connected to the buzzer. The other wires are not connected and they are all the same. However, the box never buzzes unless one wire from each of these three trees of wires is connected. One of these six wires actually makes the box buzz, but we have to take one wire from each of these trees to connect the electricity. Some other children are trying to find out if one of the wires on this tree or this tree (point) are connected to the buzzer. You are going to find out if one of the wires in

Figure 2 Apparatus for Circuit Problem



this tree is connected to the buzzer. I want you to find out if either the blue wire or the white wire is connected to the buzzer. Since the box never buzzes unless all three trees are connected, you will have to try these wires with wires from the other trees."

Circuit - Free. After the subject suggested a way to test the blue and white wires, the interviewer used one of two probes. If the subject wanted to use all six wires in discrete groups of three (for example, blue-black-red and then white-yellow-green), he was asked whether other combinations of all six wires would be an acceptable way to test the blue and white wires. Then he was asked if it would be better to test only four wires in overlapping groups of three (blue-black-red and then white-black-red). If he wanted to test four wires, he was asked if it would be better to use all six wires. The subject was always asked to explain what could be found out from his solution. Subjects were not given the opportunity to try out their solutions.

Circuit--MC and Screen. The interviewer then asked the MC and Screen questions as in the Ramp Problem. (See page 7)

Seed Problem

Only the Free question was used for Seeds. The three variables were: the amount of soil, amount of water, and amount of fertilizer. At the end of the Circuit Problem, subjects were asked the Seed Problem as follows:

"Suppose you had two kinds of seeds, white radish seeds and red radish seeds. You want to see which kind will grow the tallest in a week. You have two flower pots that are just the same, some soil, water, and fertilizer pellets. What would you do? How much soil would you put in the first pot? How much in the second?" (Similar questions were asked about the water and fertilizer.) Finally they were asked:

"Would it be just as good to use two spoons of water in the first and six spoons of water in the second pot?" Unlike Ramp or Box, Seeds was a purely verbal problem; no materials were presented to the subjects. Also unlike Box and Ramp, the Seeds question involved a concurrent comparison rather than a sequential comparison.

Sequence of Questions

We could not readily present the questions in all the possible sequences with the available sample size. For the Ramp Problem, the sequence was: Free, Variables, MC, Screen. For the Circuit Problem: Free, MC, Screen. For the Seeds Problem: Free. We presented the Free question first since we wanted to know how the child would handle the information without any suggestions from the experimenter. If we had presented MC before Free, we undoubtedly would have found that many subjects repeated their answer for MC in the Free situation instead of treating Free as a new question. Screen would also probably have influenced performance on Free, possibly by suggesting an incorrect solution. Thus we chose this question sequence on the basis of the amount of information presented in each question. It should be noted that the easiest question was MC which occurred in the middle of the sequence.

Scoring

The two authors read a sample of the interviews to establish the categories. Then all interviews were scored independently. Any discrepancies were resolved by discussion. Interrater reliability for each question ranged from 90 to 95 percent. Scores for subjects assigned to M. C. L. did not differ significantly from scores for subjects assigned to D. I. L. The categories will not be explained in detail here.[1] Table 1 briefly provides the criterion for failure and success on each question.

Insert Table 1

about here

Relationship to formal and concrete operations. Relevant to our Free question, Piaget has described the development of the "all other things equal schema" as proceeding from a preference for deliberately varying two factors at a time to a stage where the child performs "one by one multiplications," that is, has a limited ability to control variables, to a final "structured" whole where the child realizes all variables not under investigation must be controlled. We found that

Table 1

Criterion for Success and Failure on Each Question
(Numbers in parentheses refer to corresponding numerical score)

Free

	Ramp	Circuit	Seeds
Failure	(1-3) proposes to test the two marbles from different positions (4) proposes to test both from high but accepts unequal positions	(1-4) proposes a six wire solution, such as White-Yellow-Green and then Blue-Black-Red	proposes that the amount of either water, fertilizer, or soil should be different for the two seeds
Success	(5-6) realizes that the marbles must be placed at the same position	(5-6) proposes a four* wire solution, such as White-Yellow-Green and then Blue-Yellow-Green	proposes that the amounts should be the same for the two seeds

Screen

	Ramp	Circuit
Failure	(1-2) with screen in place, subject chooses the marble that makes the target go farther	(1-2) with screen in place, subject chooses the wire mentioned by the experimenter
Success	(3-4) with screen in place, subject will not make a decision based on the partial information available	(3-4) with screen in place, subject will not make a decision based on the partial information available

	MC - Ramp or Circuit	Variables - Ramp
Failure	(1-3) incorrect choice	0-6 questions answered correctly
Success	(4-5) correct choice	7 or 8 questions answered correctly

*A five wire solution such as Blue-Black-Red and Blue-Yellow-Green was also correct, but did not occur in this sample.

for Free many subjects preferred to vary two factors at a time (scores of 1, 2, or 3 for the Ramp and 1, 2, 3, or 4 for the Circuit). Some had a limited understanding of controlling variables (score of 5), and others understood the general principle of controlling variables (score of 6). For Free only scores of 5 and 6 were counted as successes.

Our category 4 for the Ramp Free question includes subjects who preferred to use the highest position on the ramp. These subjects appeared to be controlling variables, but when probed, had chosen high because it would "hit hardest" and were just as happy to have one trial at the high position, and the next trial at the low position. This preference for high as "best" is consistent with concrete thought, but is not a characteristic reported by Inhelder and Piaget.

A characteristic of concrete thought, noted by Inhelder and Piaget, that is relevant to our MC question, is that subjects prefer to test the same object twice, rather than to vary the factor under consideration. We prevented this from occurring in the format of our Free question, but not in the MC question. Many subjects who scored 1, 2, or 3 on MC wanted to test the same marble or wire twice. The remainder of those scoring 1, 2, or 3 wished to vary two factors at once. Thus, using Piaget's system we could classify subjects who score 4 and below on Free or 3 and below on MC as concrete. In our system all of these scores were counted as "failures".

Scoring of the Screen is more difficult to reconcile with Piaget's system, since he has not used this approach, but it appears to be related to his concept of the "real" and the "possible." Inhelder and Piaget (1958) say that the concrete child is limited to noting the relationships that seem to "thrust themselves on him" (p. 256) while the formal child, "in order to avoid inconsistencies as new facts emerge" (p. 256), considers relationships he observes as possible and seeks to determine the actual relationships that prevail. Thus we could say that subjects who accept the evidence from the screen condition and score 1 or 2 ("failures") are considering the real information only. Those who succeed recognize that the evidence from the screen is incomplete because they consider the possible relationships (scores of 3 or 4). By this formulation, subjects who are successful on each of our tasks are capable of thinking beyond the level of concrete operations.

Results

Each subject received either the Ramp Problem or the Circuit Problem and the Seed Problem. The Ramp and Circuit Problems each had three questions (Free, MC, Screen). The Seeds Problem only had the Free question. The number of subjects assigned to each category for each question (except Seeds) is given in Table 2.

\- - - - - - - - -

Insert Table 2

about here

\- - - - - - - - -

Table 2

Frequency of Each Category for Free, Multiple Choice, and Screen on Ramp and Circuit[1]

Ramp N = 60

Free				Multiple Choice				Screen			
Score	Age 12	Age 14	Age 16	Score	Age 12	Age 14	Age 16	Score	Age 12	Age 14	Age 16
6	2	4	10	5	4	7	8	4	1	3	12
5	6	2	0	4	5	2	7	3	0	3	1
- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
4	1	5	4	3	2	0	0	2	13	9	3
3	1	2	2	2	4	7	4	1	6	5	4
2	2	3	1	1	5	4	1				
1	8	4	3								
Total	20	20	20		20	20	20		20	20	20

Circuit N = 60

Score	Age 12	Age 14	Age 16	Score	Age 12	Age 14	Age 16	Score	Age 12	Age 14	Age 16
6	0	1	4	5	0	0	2	4	0	6	10
5	1	0	1	4	11	13	14	3	3	2	3
- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -	- - -
4	5	10	10	3	2	2	0	2	10	6	5
3	7	6	3	2	4	1	4	1	7	6	2
2	2	1	0	1	3	4	0				
1	5	2	2								
Total	20	20	20		20	20	20		20	20	20

[1] Dotted line indicates cut off for success score.

Because most measures were ordinal rather than interval, non-parametric statistics were used for data analysis. Differences between groups were analyzed using a Z score determined from Kendall's as a measure of correlation or by using the Sign Test to compare "success" on each question. Success meant that the subject understood that all but one variable must be controlled before the effects of that variable could be assessed and thus scored 5 or 6 for the Free question, 4 or 5 for MC, and 3 or 4 for Screen. All significance levels are reported for two-tailed tests.

In this section we will report (1) results for each task, (2) age differences, (3) sex differences, (4) comparisons between tasks for each of the three questions; and (5) the relationship between the three questions.

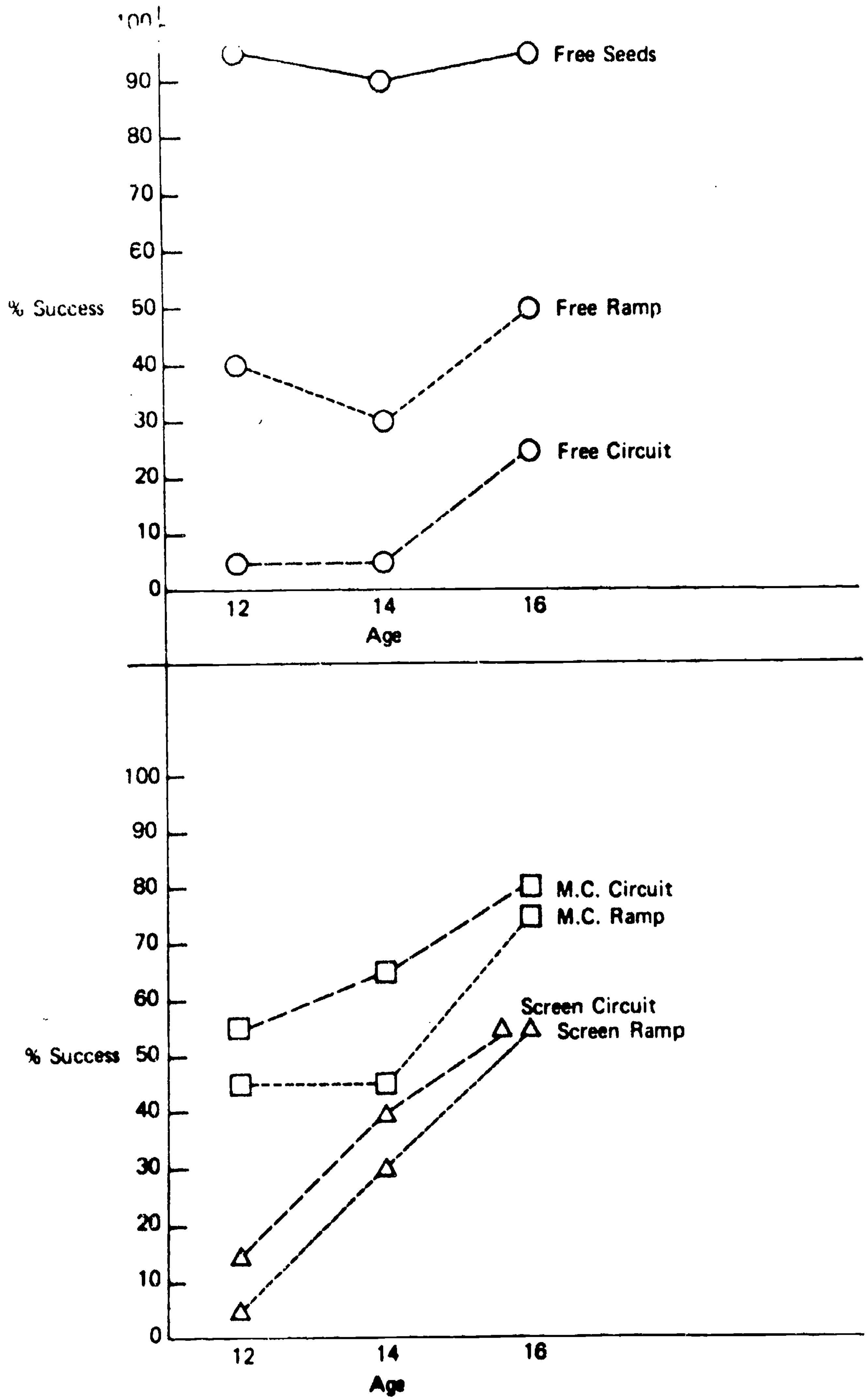
Results for Each Task

Success - Ramp. Figure 3 shows the percent of subjects at each age who were successful on each question for Ramp. The MC question was the easiest, Free and Screen were equally difficult for 12 and 14 year olds but Screen was easier than Free for 16 year olds.

Insert Figure 3

about here

Figure 3 Percentage of subjects succeeding on each question for each problem



Success - Seeds. As can be seen in Figure 3, all but four subjects correctly solved the Seeds Free question.

Age Differences

Free. As shown in Figure 3, there were no age differences for Seeds (nearly everyone succeeded) but older children did better than younger on Free for both Circuit and Ramp: Ramp $z = 2.79$ ($p < .01$); Circuit $z = 2.50$ ($p < .05$).

MC and Screen. There were no significant age differences for MC although there was a trend toward increased success with age as shown in Figure 4: Circuit $z = 1.95$; Ramp $z = 1.78$. There were striking increases with age for Screen: Circuit $z = 3.29$ ($p < .001$); Ramp $z = 3.18$ ($p < .01$).

Sex Differences and Knowledge of Variables

There were no consistent sex differences across problems for any question. On the Circuit, boys did better than girls for Free and MC: Free $z = 2.22$ ($p < .05$); MC $z = 2.21$ ($p < .05$). These differences are due to five 16-year-old boys who correctly answered both questions. For the Ramp, the only sex difference was for Variables, where boys did better than girls: $z = 2.21$ ($p < .05$). Within sexes, 67% answered all variables questions correctly, boys of all ages answered the Variables questions equally well, while older girls did better than younger girls. More complete knowledge of the variables demonstrated by boys was not associated with greater success for boys on any Ramp question either by age, or overall.

The sex differences for Circuit interact with age, since only 16-year-old boys are better than girls. One could hypothesize that this is due to some type of knowledge the older boys have. This information, however, does not also lead boys to be superior to girls on the Screen. Since for Ramp, additional information about variables did not help boys. It appears that either the information helping older boys in Circuit is not similar to the variables information in Ramp or that some other explanation must be found for the success of 16-year-old boys on Circuit.

Comparisons between Tasks

In this section we compare results for the Free question across three different tasks, the MC question across two tasks, and the Screen question across two tasks.

Free question. Although the variables were different, the solution to the Free question was the same for each task. There are significant differences between success rates for each task: Circuit was solved by 7 of 60 (10.7%), Ramp by 24 of 60 (40.0%), and Seeds by 56 of 60 subjects (93.3%). As shown in Figure 3 these differences are consistent across ages. Comparing success rates: Circuit vs. Ramp, $z = 3.32$ ($p < .001$), Ramp vs. Seeds, $z = 5.98$ ($p < .0001$).

The tasks were chosen because they differed in familiarity. Amost every 12-year-old child in a British school has grown plants and is familiar with the variables in the Seeds Problem. Many, perhaps most, children are familiar with inclined planes and marbles, 67% answered questions about the variables correctly. It is very unlikely that children have previous experiences with the Circuit Problem, although they might try to apply information that they have about electricity.

Although familiarity is a logical explanation for the differences in performance in Free, there are other possible explanations. First, perhaps the method used for presenting the task was more informative or gave more information about the variables for one task than for another. For instance, the Ramp Problem, since it was posed as a contest could have invoked a fairness cue or a motivation to win not found in Seeds or Circuit. The Circuit task might have been interpreted by the students as a problem in color matching or the way in which the wires are used in regular electric circuits. Some children did, in fact, combine wires which they believed were found in electrical circuits. When this happened, the experimenters said, "These are not wires from a plug, they were used by chance." This kind of reasoning could have interfered with the subjects' ability to discover the correct solution.

Second, one might suppose that the apparatus as a concrete referent could possibly facilitate performance, but in this case the easiest problem, Seeds, had no apparatus.

Third, the Seeds task is similar to questions asked at school and might suggest a previously learned response.

Fourth, one might suggest that the criterion for success in Circuit was more stringent than for Ramp, especially since so many subjects received a 4 in the Circuit problem. Subjects who scored 4 suggested a six wire solution but recognized that a four wire solution was better, so they clearly did not control the variables by themselves. These subjects do not even appear to be as sophisticated as Ramp subjects who received a 4 for saying that both marbles should be released from high because high is farthest from the target but when constrained to use low for one marble would use high for the other marble.

Fifth, one could argue that Seeds was a concurrent comparison rather than a sequential comparison since two flower pots were used. This would make the Seeds Free question more like the Circuit or Ramp MC question. Seeds is still easier than MC for Circuit or Ramp.

The simplest explanation of the varied difficulty of the three tasks, however, is the striking difference between the familiarity of the situations rather than the differences in method of presentation, apparatus, etc.

Results for Free, then, support the hypothesis that subjects try to solve new problems by drawing upon apparently relevant past experiences. If they actually can utilize previous experiences, they may solve the problem correctly; if they cannot, they fail.

MC and Screen questions. On MC and Screen, there were no differences between Circuit and Ramp as can be seen in Figure 3 (MC $z = 1.12$, Screen z = 0.57). Thus while the Free question was much harder for Circuit than for Ramp, MC was as easy for Ramp as it was for Circuit; Screen was equally difficult for the two problems. These results do not support the familiarity hypothesis advanced to explain results for Free; this relationship will be discussed in detail below. It is interesting to note that while Free was significantly more difficult for Circuit across ages, MC and Screen were slightly easier or equally difficult for Circuit than for Ramp at all ages.

Sequencing. Since the questions were always presented in the same sequence, it could be that one question affected performance on the next. As can be seen in Figure 3 there is no overall increase in success from the first to the third question, so increased familiarity with the apparatus does not make the questions easier. Indeed, the Screen question is about as difficult, or more difficult than the Free question for 12- and 14-year-olds. Additionally, while the nature of the problem affects performance on the Free (first) question, the age of the subjects is the best predictor of performance on the Screen (third) question. These variables appear to be more relevant to performance than the sequence of the questions but the role of sequence clearly deserves further study.

Relationship between Questions

In this section we will first compare the relative success rates for each question. For these comparisons, the role of the task must be considered. Other factors which appear to affect relative success rate will be discussed. Secondly, we will report and discuss the correlation between scores on Free, MC, and Screen for each task.

Free and MC. Comparing success on MC and Free (Figure 3), we have seen that MC is equally difficult for Circuit and Ramp while Free is harder for Circuit than for Ramp. In both Circuit and Ramp, Free is more difficult than MC: Circuit $z = 5.54$ ($p < .0001$); Ramp $z = 1.96$ ($p < .05$). Several explanations for this result are possible. It could be that by the time Circuit subjects are asked the MC, they have been given enough information to understand the task and can now answer correctly. This does not explain why there are age differences for Free but not for MC. If the problem were one of comprehension, one would expect age differences for both questions. An alternative explanation is that the MC format organized the information available for solving the Circuit or Ramp so that the task in both cases had the same requirements: the subject only needed to reason about information that had been organized for him already rather than first organizing the information and then reasoning about it.

Free and Screen. Comparing success on Screen to success on Free (Figure 3), Free is more difficult for Circuit than for Ramp, but Screen is equally difficult. As we noted when comparing MC and Free, the way the information in the question is organized appears to affect the success rate. Thus both MC and Screen present the information in a more organized way than Free. Instead of organizing the information and then reasoning about it, it appears that the subject simply reasons about the information presented. The MC emphasized the information necessary to set up a controlled experiment so it was easier than Free. The Screen emphasized results; this information is not relevant to setting up a controlled experiment. Thus Screen was as difficult as Free for 12 and 14-year-olds but easier for 16-year-olds.

Clearly some change in ability to organize or process information on this question takes place between 14 and 16. It appears that the younger child, forced to pay attention to results, apparently does not attempt to process information about the procedure used to produce the results whereas the older child does consider the procedure. Even when the experiment done behind the screen is repeated without the screen 25% of the children still choose with certainty the marble that wins or the wire that the experimenter mentions when the box buzzes. We do not, of course, know how children would respond if they first saw the experiment without the screen.

In summary, when the subject was presented with the Free question which required him to develop a plan of attack, it appeared that the familiarity of the task was a variable. But for Ramp and Circuit when the subject was presented with MC or Screen which organize the information either in a relevant or irrelevant manner, then familiarity with the task did not affect success rate. For MC and Screen, the organization of the information affected success rate.

Correlations between Questions

Further evidence for the idea that MC is a simplification of Free, and that 12 and 14-year-olds handle Screen differently from 16-year-olds is gained by looking at the correlations between the three questions shown in Table 2.

Free and MC. Although Free is harder than MC, scores are significantly correlated for both tasks: Ramp (Free, MC) = .35 ($p < .01$); Circuit (Free, MC) = .35 ($p < .01$). It appears that some of the same abilities are necessary for performance on both Free and MC.

Free and MC versus Screen. Whereas scores on screen show a dramatic increase with age, scores on Free and MC increase slightly, or not at all, with age. When we look at the relationship between scores on Screen with scores on Free or MC, we see that they are independent for 12- and 14-year-olds, but are related for 16-year-olds as shown by the correlations in Table 3.

Insert Table 3

about here

Furthermore, for 12-year-olds, the number of subjects who fail the Screen but pass the MC far exceeds the number who pass Screen but fail MC. Fourteen and 16-year-olds are just as likely to pass MC and fail Screen as to fail MC and pass Screen. Using McNemar's Test for the significance of changes: for 12-year-olds, Circuit $\chi^2 = 4.08$

Table 3

Relationship between Free, Multiple Choice, and Screen by Age Using Kendall's τ

Circuit

	Multiple Choice				Screen			
	Age				Age			
	12	14	16	All	12	14	16	All
Free	.25	.41**	.26	.35**	.00	.12	.52	.32**
Multiple Choice					.02	.10	.31	.22*

Ramp

	Multiple Choice				Screen			
	Age				Age			
	12	14	16	All	12	14	16	All
Free	.36*	.25	.23	.35**	-10	.19	.38*	.23**
Multiple Choice					-.16	.11	.29	.20*

* p<.05
**p<.01

($p < .05$), Ramp $\chi^2 = 6.13$ ($p < .02$); for 14- and 16-year-olds, there are no significant differences.

In summary, familiarity with the variables was associated with success on Free, but three major results are inconsistent with the hypothesis that familiarity is a major determinant of success. First, the familiarity hypothesis would predict that the same problem would be difficult for MC and Screen as for Free. Instead, the Circuit was most difficult for Free, while Circuit was slightly easier for MC and Screen. Second, the familiarity hypothesis would predict equal relationships across ages for each question. Instead, success on Screen was strongly related to age, while success on Free and MC were not. Third, the familiarity hypothesis would predict equal relationships across questions. Instead, Screen was independent of Free and MC for 12- and 14-year-olds, but was related for 16-year-olds. These results suggest that if familiarity is important, it is not the only important factor in this sort of reasoning. Possible explanations range from independent developmental curves for each of the three questions with a limited role for familiarity, to a catholic information processing model.

Discussion

In this section we discuss the relationship of our findings to those of Inhelder and Piaget, evidence for concrete operations, and the relevance of an information processing approach to explain these results.

Our findings strongly suggest that the nature of the task, including the mode of presenting it, influences the level of performance. What is the significance of that fact for our understanding of formal and concrete operations?

The results of this study suggest that concrete operations is not a unitary trait. Performance on Screen provides the strong relationship to age that one would expect for a change from concrete to formal thought; performance on MC and Free is not so strongly related to age. The correlations in Table 3 offer further evidence that Screen and Free or MC are assessing different types of performance in 12- and 14-year-olds. For the 16-year-olds only, results do suggest a single factor to explain performance on the two tasks.

If we assume that the ability to use the schema "all other things equal" and the progression from the real to the possible occur simultaneously, then it is difficult to reconcile the differences between performance on the three questions for the 12- and 14-year-olds. Suppose we take the real to possible explanation alone. The concrete child cannot go beyond the "real" information but the formal child can. One would expect Screen to be much harder than MC or Free for the concrete child, but not for the formal child. Results for the 12- and 14-year-olds versus the 16-year-olds support this. If ability to use the schema "all others equal" is also progressing, then the purest test of increased ability in this area would be MC. In MC, all the information is organized and the child need only apply the schema. Only the slightest age differences were found for this question. Additionally, the low correlation between Screen and MC, the high success rate of all subjects on Seeds and the fact that significantly more children pass MC and fail Screen than fail MC and pass Screen suggests that even 12-year-olds can apply the "all other things equal" schema when the information is organized for them. It appears that once information is organized, the ability to use the schema does not increase much during adolescence. This finding parallels those reported by Bryant (1974) for transitive relations in young children.

Thus Piaget's ideas about development of the "all other things equal" schema are not consistent with these results while his description of ability to move from the real to the possible is consistent with these findings. In general, the results of these experiments suggest a change in method of processing information between 12 and 16.

Several other theorists have offered explanations of changes in reasoning ability which are congruent with these findings. Lunzer (1963), for example, has discussed changes in ability to accept lack of closure, that is, ability to entertain several possible solutions to a problem or to realize that results can have more than one explanation. Information processing theorists have at various times suggested ideas about inhibiting salience (Hunt, 1961), or increases in computing space (Mc Laughlin, 1963; Pascual-Leone, 1970). One could explain this change by saying that older children had more computing space or inhibited the seemingly salient information in order to process the necessary information and therefore could organize the information in Screen and would be more efficient at Free than younger children.

It is still necessary to explain why there are large differences on Free for Seeds, Ramp, and Circuit. This was, of course, the one question which appeared congruent with a familiarity hypothesis and similar reasoning could be used to reconcile the findings with an information processing theory. One could hypothesize that information is organized better each time it is processed so that questions about information that is familiar are easier to answer. Obviously, when the information is already organized, as in MC, familiarity would not substantially help the processing. However, subjects who scored better on the variables questions in the Ramp Problem did not perform better on Free or Screen. This would imply that instruction should emphasize organizing information in a problem rather than learning facts about variables. Support for this idea comes from the work of Wollman and Karplus (1974) who found that instruction in proportionality does not aid performance on a proportional reasoning task. Further research is needed to substantiate these hypotheses.

The role of a concrete presentation for a task clearly depends on the nature of the presentation. Concrete presentations that emphasize results, like Screen, could make it less likely that a subject would use a logical skill. Questions presented completely verbally, like Seeds, are not necessarily more difficult than those presented with lots of concrete objects. It appears that a careful task analysis of any question or series of questions chosen to assess logical thinking would be advantageous in future research.

The ability to use an "all other things equal" schema, particularly when the information is organized for the child, is independent of age in the range we studied. There appears to be an interaction between the format of the problem and the use of the schema.

The findings of this study suggest that programs which aim to teach logical thought will be most successful if they emphasize the recognition and organization of relevant information rather than if they simply emphasize the "all other things equal" schema. Recent studies (Linn & Thier, 1975; Linn, Chen, & Thier, in press) offer support for these findings.

Reference Note

[1]Detailed descriptions of the categories are available from the authors.

References

Bruant, P. E. Perception and understanding in young children: an experimental approach. London: Methuen and Co., Ltd., 1974.

Hunt, J. McV. Intelligence and experience. New York: Ronald Press Company, 1961.

Inhelder, B., & Piaget, J. The growth of logical thinking: from childhood to adolescence. Parsons, A., & Milgram, S. (trans.). New York: Basic Books, Inc., 1958.

Jackson, S. The growth of logical thinking in normal and subnormal children. British Journal of Educational Psychology, 1965, 35, 255-258.

Karplus, R., & Karplus, E. F. Intellectual development beyond elementary school III - Ratio: a longitudinal study. School Science and Mathematics, 1972, 72, 735-742.

Linn, M. C., & Thier, H. D. The effect of experiential science on development of logical thinking in children. Journal of Research in Science Teaching, 1975, 12, 49-62.

Linn, M. C., Chen, B., & Thier, H. D. Personalization in science: investigations at the middle school level. Instructional Science, in press.

Levine, D. I., & Linn, M. C. Ability to control variables in adolescents: theoretical viewpoints and educational implications. Advancing Education through Science Oriented Programs, PSc Report 10, 1976.

Lovell, K. A follow-up study of Inhelder and Piaget's "The growth of logical thinking." British Journal of Psychology, 1961, 51, 143-153.

Lunzer, E. A. Problems of formal reasoning in test situations. Mussen, P. H. (ed.). European research in cognitive development. SRCD Monographs, 1965, 30, 19-46.

Lunzer, E. A. Formal reasoning: a reappraisal. Proceedings of the Jean Piaget Society, Temple University, 1973, in press.

McLaughlin, G. H. Psycho-logic: a possible alternative to Piaget's formulation. British Journal of Educational Psychology, 1963, 33, 61-67.

Pascual-Leone, J. A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologia, 1970, 63, 301-345.

Science Curriculum Improvement Study. Energy Sources. Chicago: Rand McNally, 1971

White, S. A. Evidence for a hierarchical arrangement of learning processes. Lipsitt, L. P. & Spiker, C. C. (eds.) Advances in child development and behavior, Vol. 2. New York: Academic Press, 1965.

Footnote

The authors would like to thank Dr. A. R. Jonckheere for providing encouragement and intellectual stimulation for all aspects of this study and A. R. Samaroff and Murry Levine for comments on an earlier draft of the manuscript. Special thanks are due to Miss Muriel Carter, the students, and the staff of Southgate School for their cooperation. This research was carried out while one of the authors was a visitor to the Department of Psychology, University College London, and the other author was a University of Michigan Exchange Fellow at University College London. This research was supported, in part, by a National Science Foundation grant to Advancing Education through Science Oriented Programs (AESOP). Requests for reprints should be sent to Marcia C. Linn, Lawrence Hall of Science, University of California, Berkeley, California 94720.