THE DEVELOPMENT AND VALIDATION OF A CLASSROOM TEST OF FORMAL REASONING

ANTON E. LAWSON*

*AESOP,*** *Lawrence Hall of Science, University of California, Berkeley, California 94720*

Introduction

The increase in awareness of a need to adapt curricula to the developmental levels of learners has motivated a search for reliable and valid measures of those levels. Individually administered Piagetian tasks, although found to be valid indicators of developmental level (e.g., Bart, 1971; DeVries, 1974; Goldschmid, 1967; Goldschmid & Bentler, 1968; Lawson & Blake, 1976; Lawson, Blake, & Nordland, 1974; Lawson, Nordland, & Kahle, 1975; Lawson & Renner, 1975; Lovell & Shields, 1967), require an experienced interviewer, special materials and equipment, and are too time-consuming for practical classroom use. Therefore, this investigation sought to develop a reliable and valid classroom test of developmental levels–specifically, formal-level reasoning.

Selecting the Testing Format

Initial attempts have been made to develop reliable and valid group-administered measures of formal reasoning (Burney, 1974; Longeot, 1965; Raven, 1973; Rowell & Hoffman, 1975; Shayer & Wharry, 1975; Tisher & Dale, 1975; Tomlinson-Keasey, 1975). These initial attempts have employed a variety of testing formats and have met with varying degrees of success. The selection of a testing format that will retain as many of the positive aspects of the clinical interview as possible yet allow one to test entire groups of students is an extremely important issue. The Burney (1974), Longeot (1965), Raven (1973), Tisher-Dale (1975), and Tomlinson-Keasey (1975) tests are strictly paper-and-pencil measures. Although this format has the advantage of eliminating the need for special equipment, it has important drawbacks. Perhaps the most serious of these is the loss of motivating aspects and the sense of meaningfulness that arise from physical materials and equipment. Further, the pencil-and-paper measures increase the demand on reading and writing skills which, although related to Piagetian operations, are not one and the same. These factors increase the likelihood of poor performances on the pencil-and-paper tests by certain subjects who may be quite capable of success in the more motivating and relaxed clinical setting.

*Present address: Science Education Center, Department of Physics, Arizona State University, Tempe, Arizona 85281.

**Advancing Education through Science-Oriented Programs.

Published by John Wiley & Sons, Inc. 0022-4308/78/0015-0011$01.00

In an attempt to retain as many aspects of the clinical interview as possible, Rowell and Hoffman (1975) developed a testing format in which each student is provided a set of laboratory equipment and a test booklet of instructions and questions. This testing format requires large quantities of laboratory equipment, is quite time-consuming and, therefore, restricts the assessment to a fairly small sample of student reasoning.

Shayer and Wharry (1975) have tried to retain the positive aspects of the clinical interview by using laboratory apparatus in a class demonstration format. Students see the physical materials and they hear the teacher's questions. Their responses are recorded in individual test booklets. This format would seem to be an ideal compromise solution between the clinical interview and the need to test large numbers of students in a short period of time. Shayer and Wharry, however, have not developed a sufficient variety of formal-level problems to make their instrument entirely satisfactory.

The format choosen for the present test involved one set of materials that were used by the investigator to demonstrate situations and pose questions to the class as a whole. Each student responded to the questions individually in writing on his own test booklet. This format retained important aspects of the clinical interview while, at the same time, allowed an entire class of students to be tested with a substantial number and variety of problems.

Problem Statement

The problem of this study was to develop and validate an instrument which would: (1) measure concrete- and formal-operational reasoning; (2) be capable of administration to classes of secondary school and college age students in a relatively short period of time; (3) be easily scored; (4) use a format involving physical materials and requiring as little reading and writing as possible; and (5) include a large enough number and variety of problems to assure a high degree of reliability. Before teachers can be expected to adapt their instructional methods and materials to the developmental levels of their students, such an instrument is needed. Such an instrument could be used as a research tool as well.

Selection of Formal-Operational Items

Formal operations include those reasoning processes that guide the search for and evaluation of evidence to support or reject hypothetical causal propositions. These operations are used in the isolation and control of variables, the combinatorial analysis of possible causal factors (combinatorial reasoning), the weighing of confirming and disconfirming cases (correlational reasoning), the recognition of the probabilistic nature of phenomena (probabilistic reasoning), and the eventual establishment of functional relationships between variables (proportional reasoning).

Items selected for inclusion on the test, therefore, required the isolation and control of variables, combinatorial reasoning, probabilistic reasoning, and proportional reasoning. No items were selected that could be considered to directly require correlational reasoning since, at the time of the test construction, an easily administered demonstration item of this type was not available. In addition, one item involving conservation of weight (Piaget & Inhelder, 1962) and one item involving displaced volume (Karplus & Lavatelli, 1969) were included. Both of these items have been used extensively in clinical interviews and have been found to be good indicators of late concrete and early formal operational reasoning respectively (Lawson & Blake, 1976; Lawson, Blake, & Nordland, 1974; Nordland, Lawson, & Kahle, 1974). They were included for that reason.

Test Construction and Administration

In all, 15 items were constructed for the classroom test. Each item involved a demonstration using some physical materials and/or apparatus. For each item, the demonstration was used to pose a question or call for a prediction. The students responded in writing in individual test booklets. The booklets contained only the questions followed by a number of possible answers. Students were instructed to respond by checking the box next to the best answer and then explain why they chose their answer.

Items were scored correct only if the correct box was checked and an adequate explanation was given for their selection. A brief description of the 15 items follows:

Item 1: The Conservation of Weight (Piaget & Inhelder, 1962)

This item involves two balls of clay of identical size, shape, and weight. The students are shown that the clay balls weigh the same by placing them on opposite ends of a balance beam. One of the balls is flattened into a "pancake" shape and the students are asked about the relative weights of the pieces.

Item 2: Displaced Volume (Karplus & Lavatelli, 1969)

Using two solid metal cylinders of equal size but of different density, the students are shown the level of water displaced by the lighter cylinder and asked to predict the level of water displaced by the heavier cylinder.

Item 3: Proportional Reasoning-1 (Suarez & Rhonheimer, 1974)

Using two plastic cylindrical containers of equal height but with different diameters, the students are shown that a given quantity of water rises 4 units in the wide container and rises a corresponding 6 units when poured into the narrow container. They are then asked to predict how high a given quantity of water that rises 6 units in the wide container would rise if poured into the narrow container.

Item 4: Proportional Reasoning-2 (Suarez & Rhonheimer, 1974)

Using the same plastic containers, 11 units of water are poured into the narrow container and the students are asked to predict how high the water would rise if poured into the wide container. Although the operations employed in responding to this item are presumably the same as those required in Item 3, the numbers are more complex, hence Item 4 should be more difficult.

Item 5: Proportional Reasoning-3 (Inhelder & Piaget, 1958, chap. 11)

Given a balance beam and hanging weights, the students are asked to predict where a 5-unit weight should be hung to balance a 10-unit weight which is hung 7 units of length from the fulcrum.

Item 6: Proportional Reasoning-4 (Inhelder & Piaget, 1958, chap. 11)

Using the same balance beam, the students are asked to predict where a 10-unit weight should be hung to balance a 15-unit weight which is hung 4 units of length from the fulcrum.

Item 7: Controlling Variables-1 (Inhelder & Piaget, 1958, chap. 4)

Using three pendulums (two of equal length but with bobs of 50 g and 100 g, the third longer with a 50-g bob), the students are asked to select which of the pendulums should be used in an experiment to find out if the variable of length effects the period of the pendulum.

Item 8: Controlling Variables-2 (Inhelder & Piaget, 1958, chap. 4)

Using the same three pendulums, the students are asked to select which pendulums should be used in an experiment to find out if the weight of the bob effects the period of the pendulum.

Item 9: Controlling Variables-3 (Wollman, 1975)

Using a ramp and three metal spheres, the students are shown a light sphere rolling down the ramp from a low position, striking and then displacing a target sphere which had been placed at the bottom of the ramp. The students are then asked to select the correct sphere (light or heavy) to release from a high position to find out if the variable of release position effects how far the target sphere will travel after it has been struck.

Item 10: Controlling Variables-4 (Wollman, 1975)

Using a ramp and three metal spheres, the students are shown an experiment in which two metal spheres (A and B) roll down the ramp from the same starting position and strike two target spheres of different densities. They are then asked to decide whether or not the experiment constitutes proof that metal A can displace a target further than metal B.

Item 11: Combinatorial Reasoning-1 (DeLuca, 1977; Sills & Herron, 1976)

Given a metal box with four color-coded switches and a light, the students are shown that the light can be turned on by flipping a certain combination of the switches. They are then asked to list all the possible combinations of the four switches that they would have to try to discover which combination or combinations will turn the light on. This task is logically analogous to the Inhelder and Piaget (1958, chap. 7) chemical combinations task.

Item 12: Combinatorial Reasoning-2 "Permutations" (Longeot, 1965; Piaget & Inhelder, 1975, chap. 8)

Using four objects which represent four stores (a barber shop, a discount store, a grocery store, and a coffee shop), the students are told that the stores are going to be arranged side by side on the ground floor of a new shopping center. The students are asked to list all the possible ways in which the stores could be arranged side by side. (Although it is possible for a student to solve items 11 and 12 by using a rotely applied algorithm, the likelihood of this occurring is small in that algorithms yield total numbers of combinations and/or permutations but do not indicate how to enumerate the combinations and/or permutations as is called for here.)

Item 13: Probability-1

Three yellow wooden squares and three red wooden squares are placed into a sack. The students are asked to predict the chances of drawing out a red square on the first draw.

Item 14: Probability-2

Three red squares, four yellow squares, and five blue squares are placed into a sack. Four red diamond-shaped pieces, two yellow "diamonds," and three blue "diamonds" are also placed into the sack. The students are asked to predict the chances of drawing out a red piece on the first draw.

Item 15: Probability-3

Using the same wooden pieces as in Item 14, the students are asked to predict the chances of drawing a red or blue "diamond" on the first draw.

Method

Subjects

The classroom test was administered to a total of 513 students enrolled in eighth-, ninth-, and tenth-grade classes in a junior and two senior high schools located in two middle- to upper-middle-class suburban communities in the San Francisco Bay area. Classes of students were selected from the entire range of student abilities in all three grades.

A total of 145 eighth-grade subjects (Ss) were selected from the classes of one junior high school science teacher. These students had been randomly assigned to their required science class at the beginning of the school year. The Ss (73 males and 72 females) ranged in age from 12.8 to 15.3 years, mean age = 14.1 years.

A total of 192 ninth-grade Ss (96 males and 96 females, 13.3 to 17.0 years in age, mean age = 15.2 years) were selected from the classes of five high school English teachers. Since English is a required course in the ninth grade, and both remedial and regular English classes were tested, these Ss constituted a reasonably representative sample of the high school's ninth-grade population.

A total of 176 tenth-grade Ss (89 males and 87 females, 15.5 to 18.3 years, mean age = 16.5 years) were selected from the classes of four biology teachers. Since biology is also a required course, these Ss also represented a reasonably representative sample of this high school's tenth grade population. A few eleventh- and twelfth-grade students were included in this sample in that not all students elected to take biology as sophomores.

The eighth- and tenth-grade samples were drawn from a junior and a senior high school which were situated next to each other. All of the junior high school students graduated to this high school. Further, this junior high school is the only feeder school for the high school. Therefore, these two samples were considered to have been drawn from the same population of students. The ninth-grade subjects, on the other hand, were selected from a high school located in an area of higher socioeconomic standing and, therefore, cannot be considered to have been drawn from the same population as the eighth- and tenth-grade samples.

Procedure

The 513 Ss were administered the classroom test. Testing of each class took between 75 and 100 minutes during two consecutive periods. In order to determine the validity of the group test (i.e., assure that the group test measures the same psychological parameter(s) as measured by standard Piagetian tasks), a subsample of 72 Ss were randomly selected and individually administered a battery of four Piagetian tasks. The individual interviews were conducted by three trained interviewers. Each interview lasted about 30 minutes.

The tasks administered were the conservation of weight (Piaget & Inhelder, 1962), displaced volume (Karplus & Lavatelli, 1969), bending rods (Inhelder & Piaget, 1958, chap. 3), and the balance beam (Inhelder & Piaget, 1958, chap. 11). Student responses on the bending rods and the balance beam tasks were summed and classified into developmental levels that reflected the use of concrete, transitional, or formal reasoning. The relationship between scores on the classroom test and the summed bending rods and the balance beam scores was then analyzed using a contingency table and summarized by a correlation coefficient. If the classroom test is a valid measure of the same psychological parameters measured by these interview tasks, then a strong positive relationship should exist between scores on the measures. Responses on the conservation task and the displaced volume task were used in a later analysis to factorially validate the classroom test.

Since each interview task has been employed by previous investigators (Lawson, Nordland, & DeVito, 1974) only brief descriptions of the tasks and materials used will be mentioned here.

Conservation of Weight. Two balls of clay of approximately 50 grams each were presented to S. One ball was then transformed into a "pancake" shape and S was asked about the relative weights of the balls.

Displaced Volume. Two metal cylinders of equal volume but different weight (18 g and 55 g) were handed to S. The equal height and thickness of the metal cylinders were pointed out. The examiner then took the cylinders and lowered the lighter one into one of two test tubes (30 ml) which were partially filled with equal amounts of water. The rise in water level was noted and S was asked to predict the final water level when the heavier cylinder was lowered into the other test tube.

Bending Rods. This task tested S's ability to identify and control variables. Given six flexible metal rods of varying length, diameter, shape, and material which were fastened to a stationary block of wood, S was asked to identify variables and demonstrate proof of the affect of each variable on the amount of bending of the rods when weights were hung from the ends of the rods.

The Balance Beam. Using a balance beam and hanging weights, this task tested S's ability to balance various combinations of weights at various locations along the beam; e.g., given a 15-unit weight hung 4 units of length from the fulcrum, S was asked to predict the proper location to hang a 10-unit weight to achieve a balance. Successful completion of this task implied an understanding of inverse proportionality.

Reliability estimates of these tasks have been established by Lawson, Nordland, and DeVito (1974). Test-retest correlation coefficients ranged between 0.48 and 0.78. Lawson, Nordland, and Kahle (1975) found the reliability of a series of Piagetian tasks, including those used in this study, to be high. Cronbach's Alpha coefficient, a modification of the KR-20 formula for scalable items, was 0.86. Validity of the Piagetian tasks has been determined by numerous investigators.

Results and Discussion

Classroom Test Statistics

The scores for the entire sample on the classroom test ranged from 0 to 15. The distribution of scores approximated a normal curve with a mean of 7.41, standard deviation of 4.27, and standard error of measurement of 2.0. Table I shows the corresponding values for each grade as well as the totals. The Kuder-Richardson 20 estimate of reliability was 0.78. This value, although not as large as might be hoped for, represents an adequate degree of reliability.

The test items were found to be of roughly three levels of difficulty. Items 1, 2, 5, 7, and 8 were the least difficult. The percentage of correct responses ranged from 83.1% on Item 1 to 60.5% on Item 7. Items 9, 10, 13, and 14 were somewhat more difficult. The percentages of correct responses ranged from 48.0% on Item 9 to 43.5% on Item 10. The most difficult items were 3, 4, 6, 11, 12, and 15. Of these, Item 15 was answered correctly by 34.1% of the subjects, and Item 4 was the most difficult with only 23.3% correct responses.

Is the Classroom Test a Valid Measure of Formal Reasoning?

To assess validity of the classroom test as a measure of formal reasoning, three types of evidence were sought. The first type of evidence concerned face validity. A panel of six judges, who were considered experts due to their professional involvement with Piagetian research, responded with 100% agreement that the test items appeared to require concrete and/or formal reasoning. It was concluded, therefore, that the classroom test has face validity.

The second type of evidence concerned the relationship between the classroom test total scores and the level of subject response on the bending rods and balance beam tasks. Pearson product-moment correlations between the classroom test total score and level of response on

TABLE I

Summary Statistics for the Classroom Test of Formal Reasoning

Grade[a]	N	Mean Score	Standard Deviation	Range	Standard Error of Measurement
8	145	5.68	3.23	1-12	1.5
9	192	7.46	4.62	0-15	2.2
10	176	8.04	3.94	1-15	1.8
Total	513	7.41	4.27	0-15	2.0

[a]The eighth- and tenth-grade samples were drawn from the same school population while the ninth-grade sample was drawn from a school located in a community of higher socio-economic standing.

these tasks were 0.75 and 0.65, respectively. [Parametric statistics have been used in the present data analysis although equal interval data are not assumed. Although some may question the efficacy of this, most investigators agree that scale type is irrelevant to the choice of a statistical tool. Parametric procedures are generally advocated due to their greater sensitivity. For a detailed account of this issue, see Gardner (1975).] Both correlation coefficients were statistically significant ($p < 0.001$). If one assumes that these two tasks largely measure the same psychological parameter (i.e., formal thought), then it would be reasonable to combine them to obtain a single interview task score. Since the results of a principal components analysis (reported below) confirmed that the tasks were to some extent measuring the same parameter, the task scores were summed. The correlation between the classroom test total score and this summed bending rods and balance beam task score was 0.76 ($p < 0.001$). This high correlation between the measures documents that the classroom test has convergent validity.

As a third type of evidence of the classroom test's validity, the classroom test and all four interview tasks were submitted to a principal components analysis. Notice that this analysis was based on only 72 cases, therefore, the results should be viewed as tentative at best. Since all of the classroom test items and the interview tasks, with the exception of the conservation of weight, were designed to measure aspects of formal-operational reasoning, the analysis should yield only two principal components: one component identifying concrete reasoning, and one identifying formal reasoning. Prior to the analysis, the classroom test items which purported to measure proportional reasoning (#3–6) were summed for each S to obtain a total proportional reasoning score. Those items designed to measure the ability to control variables (#7–10), combinatorial reasoning (#11–12), and probability (#13–15) were also summed.

These summed scores, the score on the conservation of weight item and the displaced volume item and the four interview tasks were then submitted to the principal-component analysis (Nie et al., 1975). This analysis extracts only components with eigenvalues greater than one (Kaiser, 1960) and uses the varimax method of axes rotation. The results of the analysis are shown in Table II.

Inspection of the table shows that, in fact, three principal factors were extracted which accounted for 66.0% of the total variance. As expected, the classroom test items involving proportions, the control of variables, combinatorial reasoning, and probability loaded heavily on the same factor (0.64–0.77). The bending rods and balance beam interview tasks also loaded on this factor (0.68 and 0.53, respectively). The finding that these classroom test items and these interview tasks loaded heavily on the same factor, supports the hypothesis that they are measuring aspects of the same psychological parameter, i.e., formal-operational reasoning.

Both the displaced volume classroom test item and the displaced volume interview task loaded heavily on a second factor (0.82 and 0.86, respectively). Moderate loading of controlling variables classroom test items and the bending rods and balance beam task scores also were obtained on this factor (0.39–0.53). Both versions of the conservation of weight problem loaded heavily on a third factor (0.89 and 0.85).

These results suggest that three, not two, identifiable psychological parameters are being measured. Factor 1 can be interpreted as "formal reasoning" and factor 3 as "concrete reasoning" as measured by the conservation of weight. Factor 2 represents what can probably be identified as "early formal reasoning" as measured by the displaced volume task and partially measured by the controlling variables classroom test items, and the two Piagetian interview tasks.

Subject responses could be scored as early formal-operational as well as fully

TABLE II

Principal-Components Analysis of Classroom Test Items and Interview Tasks with Varimax Rotation ($n = 72$)

Variable[a]	Variance Account For By:		
	Factor 1 (42.3%)	Factor 2 (13.8%)	Factor 3 (9.9%)
Classroom Test			
Wgt-C			.89
Vol-C		.82	
Proportions	.75[b]		
Controlling Variables	.65	.39	
Combinatorials	.64		
Probability	.77		
Interview Tasks			
Wgt-I			.85
Vol-I		.86	
Bending Rods	.68	.51	
Balance Beam	.53	.53	

[a]Wgt-C = Conservation of weight-classroom version, Vol-C = Displaced Volume-classroom version, Proportions = Proportional Reasoning Items, Controlling Variables = Controlling Variables Items, Combinatorials = Combinatorial Reasoning Items, Probability = Probability Items, Wgt-I = Conservation of Weight-Interview Task, Vol-I = Displaced Volume-Interview Task, Bending Rods = Bending Rods Task, Balance Beam = Balance Beam Task.

[b]Only loadings greater than 0.30 are included.

formal-operational on the interview tasks, so the finding that these tasks loaded moderately on a factor, identified as "early formal reasoning," is not surprising. The result that the controlling variables items loaded in this component may be due to the fact that two of these items (#7–8) involved familiar and intuitively understood variables (e.g., length and weight); hence, these items did not require fully formal-operational reasoning for successful response.

In short, these results support the hypothesis that the classroom test measures aspects of formal reasoning as well as some aspects of concrete reasoning and reasoning that could be considered intermediate. The classroom test can, therefore, be said to have factorial validity as well as face validity and convergent validity.

Classifying Scores on the Classroom Test into Developmental Levels

The classroom test was designed to allow teachers and/or researchers to classify student performance into developmental levels.

Basically, four types of information were used to decide upon a reasonable classification scheme. A contingency table comparing scores on the classroom test with concrete, transitional, and formal-level responses on the combined bending rods and balance beam interview tasks was a primary source of information. Secondary knowledge of what the classroom test items were likely to measure was gained through previous investigations (e.g., Inhelder & Piaget, 1958; Lawson & Renner, 1974; Lawson & Wollman, 1975; Longeot, 1965; Wollman, 1975). The third source of information came from an item analysis of the test items and the fourth from the principal-components analysis.

Although space does not permit a detailed discussion of use of these four sources of information, a brief accounting of the use of the contingency table seems in order. Initially, a contingency table comparing summed score on the bending rods and balance beam interview tasks and scores on the classroom test was constructed. A significant relationship existed between the classroom test scores and the summed tasks scores (contingency coefficient = 0.84. Since the scores on the interview tasks were known to reflect concrete, transitional, or formal-level reasoning, it was possible to inspect the table and note that the majority of the Ss who were classified at the concrete level on the tasks scored from 0 to 5 on the classroom test, the majority of Ss who were classified at the transitional level on the tasks scored from 6 to 11 on the classroom test, while the majority of Ss who were classified as formal on the interview tasks scored from 12 to 15 on the classroom test. The contingency table is shown in Table III. Tentatively, therefore, scores of 0–5 were classified as representing concrete reasoning, scores of 6–11 were classified as transitional, and scores 12–15 were classified as representing formal reasoning. The principal-components analysis, the item analysis, and previous knowledge of what the items measure all suggested that this scoring scheme was reasonable.

In using this scheme it was found that, of the 513 Ss who took the classroom test, 181 (35.3%) responded at the concrete level, 254 (49.5%) responded as transitional to formal-operational, while 78 (15.2%) responded as formal-operational. It should be kept in mind, however, that the classroom test measures different aspects of formal reasoning, that is, different formal-operational schemata (e.g., combinations, proportions, probability). Although a good deal of consistency of student performance was found across these aspects of formal reasoning, they are different nevertheless. Therefore, the total classroom test score should not be viewed as indicating a strictly unidimentional scale.

One further point should be kept in mind when one classifies students on the basis of performance on the classroom test. In reference once again to Table III, let us assume that scores of 0–5 do, in fact, represent concrete reasoning and scores of 6–11 represent the transition to formal operations and so on. Further, let us assume that performance of each student on the interview tasks is near an optimum level. Then, theoretically, all subject responses should be found in the diagonal cells. In Table III, 45 of the 72 responses fit this

TABLE III

Contingency Table Comparing Classroom Test Total Score with Piagetian Level of Response on the Combined Bending Rods and Balance Beam Tasks (n = 72)

		Piagetian Level		
		Concrete	Transitional	Formal
Classroom Test Total Score	0–5	18	10	1
	6–11	5	20	7
	12–15	0	4	7

$\chi^2 = 31.3$, df = 4, $p < .001$
Contingency coefficient = .84

theoretical expectation. Of the responses classified at different levels, 17 were classified at a higher level on the interview tasks while only nine were classified at a higher level on the classroom test. This suggests that, for the majority of cases, subject responses will be classified into the same levels on the classroom test as on the interview tasks, but that the classroom test may slightly underestimate the capabilities of the class as a whole.

Conclusions

The major conclusion of this study is that the same psychological parameters measured by classical Piagetian interview tasks were measured by a series of classroom demonstration test items with a fairly high degree of reliability. The demonstration test items were easily administered to entire classes of students in a short period of time and quickly scored. Further, the classroom test was found to have face validity, convergent validity, and factorial validity.

Implications for Teaching

The classroom test described in this article is intended primarily for use by teachers interested in determining the developmental levels of their students. Use of the test as a source of knowledge about individual student reasoning processes can be extremely worthwhile, but only if (1) the teacher understands the nature of concrete and formal thought and (2) can adapt course content, goals, teaching methods and evaluation procedures to fit these levels of reasoning. [For an in-depth discussion of the implications of the transition from concrete to formal thought for science teaching see Collea et al. (1975); Lawson et al. (1976); and Karplus et al. (1976).] Though this is not the place to explore these issues in depth, a quotation from Ausubel discussing the implications of the transition from concrete to formal thought outlines the central issue. According to Ausubel:

> This developmental shift has far reaching implications for teaching methods and curricular practices in the secondary school.
>
> Once the developing individual reaches the abstract stage of cognitive functioning, he becomes in a large measure an abstract verbal learner. He now acquires most new concepts and learns most new propositions by *directly* (without mediating and constraining influence of concrete-empirical props) apprehending verbally or symbolically stated relationships between previously learned abstractions. To do so meaningfully, he need no longer refer to first-hand, concrete or nonrepresentational experience, nor actually perform any of the abstracting or generalizing operations on the underlying empirical data (pp. 219-220).

Ausubel stresses what the formal-operational learner has become able to do. By the same token, the concrete-operational learner has *not* become able to do these things. As the recent research reported by Lawson and Renner (1975) has documented, secondary school students who are still reasoning at the concrete level are able to learn very little, if any, of what is taught in an abstract verbal way.

Although, to a large extent, the individual who has developed into the formal stage is an abstract verbal learner. This by no means implies that he can never profit from concrete experiences. Indeed, anytime any person moves into an unfamiliar conceptual domain, to acquire meaningful understanding, he must begin with the concrete underlying empirical data of that new discipline. Again, as Ausubel states it:

> It would be very misleading, however, to assert that secondary-school, and even

> older students (Ausubel, of course, is referring only to those secondary-school and older students who have in fact attained formal operations) can *never* profit either from the use of concrete-empirical props to generate intuitive meanings, or from the use of inductive discovery and deductive problem-solving techniques to enhance such meanings. As previously suggested, generally mature students tend to function at a relatively concrete or intuitive level when confronted with a particularly *new* subject-matter area in which they are totally unsophisticated (p. 220).

The identification of students who often fail to use formal reasoning, coupled with appropriate modifications in teaching practice, hopefully could do a great deal for them in eliminating the rote-meaningless memorization of cognitively unassimilable material, and the frustration and intellectual stagnation that result from such a psychologically unpalatable diet. As a number of recent studies suggest, such modifications could significantly aid these students in the development of formal reasoning (e.g., Arons, 1976; Case, 1975; Case & Fry, 1973; Fischbein, Pampu & Minzat, 1970; Lawson & Wollman, 1976; Linn & Thier, 1975; Renner & Lawson, 1973, 1975; Tomlinson-Keasey, 1976).

Use of the classroom test could also benefit those students who are quite accomplished formal reasoners but have been mistakenly placed into classes for remedial learning. This past year a series of Piagetian tasks was administered to 28 students enrolled in a special sophomore biology course for "less able" students. The majority of these students responded to the Piagetian tasks at the concrete or postconcrete level. Three of the students, however, displayed formal reasoning. For this reason they were transferred to a regular biology class and, much to their and the teacher's satisfaction, they performed extremely well.

Of course, the classroom test could also serve as a means of selecting students for special advanced or remedial classes. Such a use, however, requires extreme caution regarding test reliability and keen awareness of the nature of concrete and formal reasoning processes.

References

Arons, A. B. Cultivating the capacity for formal reasoning: Objectives and procedures in an introductory physical science course. *American Journal of Physics*, 1976, **44**(9), 834-838.

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart & Winston, 1968.

Bart, W. M. The factor structure of formal operations. *British Journal of Educational Psychology,* 1971, **41**, 70-79.

Burney, G. M. The construction and validation of an objective formal reasoning instrument. Unpublished doctoral dissertation, University of Northern Colorado, 1974.

Case, R. Gearing the demands of instruction to the developmental capacities of the learner. *Review of Educational Research,* 1975, **45**(1), 59-87.

Case, R., & Fry, C. Evaluation of an attempt to teach scientific inquiry and criticism in a working class high school. *Journal of Research in Science Teaching,* 1973, **10**(2), 135-142.

Collea, F. P., Fuller, R. G., Karplus, R., Paldy, L. G., & Renner, J. W. *Workshop on physics teaching and the development of reasoning.* Stony Brook, N.Y.: American Association of Physics Teachers, 1975.

DeLuca, F. P. Measurement of logical thinking: An electronic equivalent of Piaget's first chemical experiment. Paper presented at the 50th Annual Convention of the National Association of Research in Science Teaching, Cincinnati, March 1977.

DeVries, R. Relationships among Piagetian, IQ and Achievement Assessments. *Child Development,* 1974, **45**, 746-756.

Fischbein, E., Pampu, I., & Minzat, I. Effects of age and instruction on combinatory abilities in children. *The British Journal of Educational Psychology,* 1970, 261-270.

Gardner, P. L. Scales and statistics. *Review of Educational Research,* 1975, **45**(1), 43-57.

Ghiselli, E. E. *Theory of psychological measurement.* New York: McGraw-Hill, 1964.

Goldschmid, M. L. Different types of conservation and non-conservation and their relation to age, sex, IQ, MA, and vocabulary. *Child Development,* 1967, **38**, 1229-1246.

Goldschmid, M. L., & Bentler, P. M. *Manual: Concept assessment kit-conservation.* San Diego: Educational and Industrial Testing Service, 1968.

Inhelder, B., & Piaget, J. *The growth of logical thinking from childhood to adolescence.* New York: Basic Books, 1958.

Iowa Test of Educational Development. *Handbook for teachers and examiners.* Des Moines: University of Iowa, 1972.

Kaiser, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement,* 1960, **20**, 141-151.

Karplus, E. F., & Karplus, R. Intellectual development beyond elementary school: I. Deductive logic. *School Science and Mathematics,* 1970, **70**(5), 398-406.

Karplus, R., & Lavatelli, C. *The developmental theory of Piaget: Conservation.* San Francisco: Davidson Film Producers, 1969.

Karplus, R., Lawson, A. E., Wollman, W., Appel, M., Bernoff, R., Howe, A., Rusch, J. J., & Sullivan, F. *Science teaching and the development of reasoning: A workshop.* Berkeley: Regents of the University of California, 1976.

Lawson, A. E., & Blake, A. J. D. The factor structure of some Piagetian tasks. *Journal of Research in Science Teaching,* 1976, **13**(5), 461-466.

Lawson, A. E., Blake, A. J. D., & Nordland, F. H. Piagetian tasks clarified: The use of metal cylinders. *The American Biology Teacher,* 1974, **36**(4), 209-211.

Lawson, A. E., Carlson, E., Sullivan, F., Wilcox, R. S., & Wollman, W. *Biology teaching and the development of reasoning: A workshop.* Berkeley: Regents of the University of California, 1976.

Lawson, A. E., Nordland, F. H., & DeVito, A. Piagetian formal operational tasks: A crossover study of learning effect and reliability. *Science Education,* 1974, **58**(2), 267-276.

Lawson, A. E., Nordland, F. H., & Kahle, J. B. Levels of intellectual development and reading ability in disadvantaged students and the teaching of science. *Science Education,* 1975, **59**(1), 113-126.

Lawson, A. E., & Renner, J. W. Relationship of science subject matter and developmental levels of learners. *Journal of Research in Science Teaching,* 1975, **12**(4), 347-358.

Lawson, A. E., & Wollman, W. T. Encouraging the transition from concrete to formal cognitive functioning: An experiment. *Journal of Research in Science Teaching,* 1976, **13**(5), 413-430.

Lawson, A. E., & Wollman, W. T. Teaching proportions and intellectual development: An experiment. Unpublished manuscript, University of California, Berkeley, 1975.

Linn, M. C., & Thier, H. D. The effect of experimental science on development of logical thinking children. *Journal of Research in Science Teaching,* 1975, **12**(1), 49-62.

Longeot, F. Analyse statistique de trois tests génétiques collectifs. *Bulletin de l'Institut National D'Etude,* 1965, **20**(4), 219-237.

Lovell, K., & Shields, J. B. Some aspects of a study of the gifted child. *British Journal of Educational Psychology,* 1967, **37**, 201-208.

Lovell, K. Some problems associated with formal thought and its assessment. In D. R. Green, M. P. Ford, and G. B. Flamer (Eds.), *Measurement and Piaget.* New York: McGraw-Hill, 1971.

McKinnon, J. W., & Renner, J. W. Are colleges concerned with intellectual development? *American Journal of Physics,* 1971, **39**, 1047-1052.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill, 1975.

Nordland, F. H., Lawson, A. E., & Kahle, J. B. A study of concrete and formal reasoning ability in disadvantaged junior and senior high school science students. *Science Education,* 1974, **58**(4), 569-575.

Piaget, J., & Inhelder, B. *Le développement des quantités physiques chez l'enfant* (2nd ed.). Neuchatel: Delachaux and Neistel, 1962.

Piaget, J., & Inhelder, B. *The origin of the idea of chance in children.* New York: Norton, 1975.

Raven, R. J. The development of a test of Piaget's logical operations. *Science Education,* 1973, **57**(3), 33-40.

Renner, J. W., & Lawson, A. E. Promoting intellectual development through science teaching. *The Physics Teacher,* 1973, **11**(5), 273-276.

Renner, J. W., & Lawson, A. E. Intellectual development in preservice elementary school teaching: An evaluation. *Journal of College Science Teaching,* 1975, **5**(2), 89-92.

Rowell, J. A., & Hoffman, P. J. Group tests for distinguishing formal from concrete thinkers. *Journal of Research in Science Teaching,* 1975, **12**(2), 157-164.

Russell, H. The measurement of reasoning ability in adolescents. *Ontario Journal of Educational Research,* 1962, **5**(1), 33-40.

Shayer, M., & Wharry, D. Piaget in the classroom: I. Testing a whole class at the same time. Unpublished manuscript, Chelsa College, University of London, 1975.

Sills, T. W., & Herron, J. D. Study of an electronic analog to the combinations of chemical bodies Piagetian task. *Journal of Genetic Psychology,* 1976, **129**, 267-272.

Suarez, A., & Rhonheimer. *Lineare Funktion.* Zurich: Limmat Stiftung, 1974.

Tisher, R. P., & Dale, L. G. *Understanding in science test.* Victoria: Australian Council for Educational Research, 1975.

Tomlinson-Keasey, C. Introduction to formal operations task. Unpublished manuscript, University of Nebraska, 1975.

Tomlinson-Keasey, C. Can we develop abstract thought in college freshmen? Paper presented at the 84th Annual American Psychological Association Convention, Washington, D.C., September, 1976.

Wollman, W. Intellectual development beyond elementary school: VI. Controlling variables: A survey. Unpublished manuscript, University of California, Berkeley, 1975.

Manuscript accepted June 1, 1977