AN ABSTRACT OF THE THESIS OF

Robert A. Logan for the Master of Science

in Psychology presented on April, 1991

Title: Relationships between the Arlin Test of Formal Reasoning and Classroom Multiple-Choice Tests

Abstract approved: [signature]

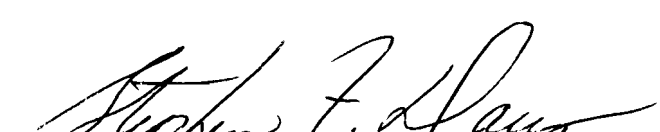
This study examined the relationships between the Arlin Test of Formal Reasoning (ATFR) and a classroom multiple-choice (CMC) test to determine the potential of CMC tests being utilized by classroom instructors to identify concrete and formal reasoning students. This study also examined the relationships between seven demographic factors and ATFR performance scores.

The data were obtained from 258 college students (108 men, 150 women) enrolled in nine sections of introductory psychology courses at a midwestern university. Student mean age was 19.45 years with a range of 17 to 39 years. Within a one week time period both the ATFR and the 70-item CMC test, consisting of 35 conceptual and 35 rote-memory items, were administered to each section of students.

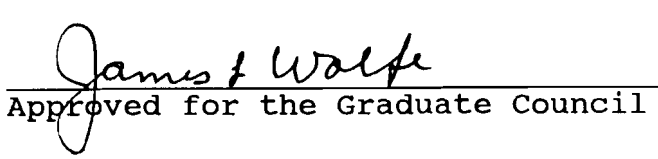
Results obtained from the ATFR indicated that 40 percent of the students scored at the formal reasoning level, 34.5 percent at the concrete level, and 25.5 percent at the transitional level. Significant correlations between ATFR scores and each of the CMC scores were obtained, with student performance on the conceptual items being lower than

on the rote items. One-way analyses of variance (ANOVA) procedures revealed significant main effects of reasoning ability on each of the CMC subtests. One-way ANOVA procedures computed on each of the seven demographic variables revealed no significant main effects for any of the variables on ATFR scores.

It was concluded that a need exists to accommodate concrete thinkers at the college level and that CMC tests have strong potential to be used as ongoing assessment tools of student reasoning ability. In addition, it is suggested that CMC tests utilize conceptual type items to assess student learning. In spite of the lack of effect that demographic factors had on reasoning ability, it is suggested that further research be conducted on the factors of age and gender.



Approved for the Major Division



Approved for the Graduate Council

Relationships between the Arlin Test of Formal Reasoning and Classroom Multiple-Choice Tests

A Thesis

Presented to

the Division of Psychology and Special Education

Emporia State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Robert Alan Logan

April, 1991

ACKNOWLEDGEMENTS

I would like to express my gratitude to my committee, Dr. David Dungan, Dr. Deanna Hawes, and Mr. Howard Carvajal for their guidance and expertise in writing this thesis.

I would also like to thank my wife, Kelly, for her many hours of typing and the unending support she has given me. In addition, I would like to thank my three stepchildren, Nicole, Michael, and Randi, as well as my mother, for their patience, support, and encouragement.

Lastly, I would like to thank Dr. James Parks, and Dr. Loren Tomkins for their experience and assistance as researchers whose help made the completion of this thesis possible.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Introduction

Within the last several decades, cognitive theory has generated an enormous amount of attention and interest from many professional fields. This theoretical framework concerns itself with the study of an individual's ability to think or to engage in various types of mental activities (Neimark, 1975). More specifically, this theoretical approach has made an attempt to explain both how and when various types of mental structures of thought are developed. Thanks to the many insights provided by cognitive theory, researchers, parents, and educators now have a greater appreciation of the capacities and limitations of the types of thinking that are possible at various ages (Flavell, 1977).

One of the most influential of the cognitive theorists was Jean Piaget. Previous study and research done in the assessment of childhood intelligence led Piaget to formulate the first comprehensive theory of cognitive development (Inhelder & Piaget, 1958; Piaget & Inhelder 1969; Piaget, 1972). Piaget maintained that there are four major stages of cognitive development with each stage being age-related and consisting of structural features that permit certain types of thinking.

It is Piaget's last two stages of cognitive development that served as the focus for this research study. In short,

the essential features of these two stages consist of concrete operational reasoning versus formal operational or abstract reasoning. Piaget held that concrete reasoning ends approximately at the beginning of adolescence, thus ushering in the ability to think hypothetically and abstractly.

Statement of Problem

Over the last two decades there has been a growing body of evidence that suggests many late adolescents and adults are still functioning in the concrete stage of thinking (Bart, 1971; Kuhn, 1979; Logan & Dungan, 1990). These findings are not in accordance with Piaget's theory of cognitive development, thus creating some confusion and uncertainty for those who have readily accepted Piaget's assertions. This is especially true for college educators whose primary task is to teach and work with those students who fall within those particular age ranges.

Also, the contradictory results found in the literature pose several other problems that warrant further investigation. For example, if a significant number of college students are functioning below the formal level of reasoning, then a strong need exists to construct and design adequate group tests that will assess the various levels of reasoning. The original interview method utilized by Inhelder and Piaget (1958) is extremely time consuming and expensive, thus rendering it impractical for assessment of

large numbers of students. Although there are currently several group tests of formal reasoning available, most are seriously flawed (Nagy & Griffiths, 1982). Thus, a need still exists for further research and construction of tests that are adequate for individual and group assessment of college student reasoning ability.

One other problem that needs to be addressed is the identification of demographic factors that could possibly influence reasoning abilities. Although previous research has been conducted utilizing some demographic factors (Blackburn, 1980; Buss, 1977; Douglass & Wong, 1977), the results are contradictory and warrant further investigation.

If it can be determined that a significant number of college students are still functioning at a concrete level, then the implications for educators would be enormous. For example, it is currently assumed that college students have the ability to think abstractly, which allows the student to process effectively the abstract material presented in many college courses. This study provides some evidence of whether or not this assumption should be maintained. If it should be determined that this assumption is false, then educators will be faced with the issue of accommodating concrete thinkers at the college level.

Other possible implications could include (a) evaluation of college admission criteria, (b) developing different methods of instruction at both the high school and

college levels, (c) and assessing the effectiveness of classroom evaluation instruments. In regard to the latter point, it is possible that current measures of student learning allow students to perform well on tests without acquiring a thorough understanding of the course material. For example, simple memorization of facts, terms, and lists of information often allow students to perform well on standardized tests without having a real understanding of the conceptual nature of the course material, and thus limit their ability to apply that same material to practical problems relevant to their field of study. This study provides some data on this issue by assessing the students' ability to perform on a test utilizing both factual and conceptual type questions.

Given the overall importance of education in our society, it would seem that a study of this nature would be of great value in addressing the previously mentioned issues. New data are necessary if sound decisions and practical policies are to be made regarding the educational processes that are currently employed. A review of previous research shows both the importance and the need for further studies to be done.

Literature Review

Assessment of cognitive development was originally done on an individual basis by using the clinical interview method (Inhelder & Piaget, 1958). This method of assessment

requires the expertise of a trained evaluator and involves the use of bulky laboratory materials, which is consequently a very time consuming process. It was precisely because of these limitations that theorists and researchers began to focus their attention on developing group paper-and-pencil tests.

Based upon the current research, Gray (1978) produced impressive evidence that Piaget's theory of cognitive development could be used as a foundation for developmentally-based standardized tests. It was further suggested by Gray that individual test items be carefully constructed so that, "emphasis is placed on the differences between the competencies represented by the items and thus on the behavioral changes represented by those differences" (p.2). There are now numerous group paper-and-pencil tests available, which combine the speed and precision of conventional psychometric tests with the measurement of abilities described by Piaget's original interview method (Shayer, Adey & Wylam, 1981).

Tests of Cognitive Development

Lawson (1978) constructed a 15-item test to measure formal reasoning ability in junior and senior high school students. It was Lawson's belief that a test of formal reasoning should consist of items that resemble as closely as possible those tasks that were originally utilized by Piaget in his clinical interview method. The items used in

Lawson's Classroom Test of Formal Reasoning (CTFR) were designed to measure the mental operations used in the isolation and control of variables, proportional reasoning, probabilistic reasoning, and combinatorial reasoning. Subjects used in the study were 513 students, both male and female, selected from grades 8-12 in two suburban communities located in the San Francisco area.

Initial results reported by Lawson suggested that a valid test of formal reasoning had been developed. But, was it really valid? A more recent study (Pratt & Hacker, 1984) using 136 college students was conducted to assess the validity of the original CTFR and a modified version. It was concluded from the results that the test failed to provide a valid measure of formal reasoning. It was further suggested that the CTFR neglected essential aspects emphasized by the original work of Inhelder and Piaget (1958).

Additional evidence indicating a lack of validity for the CTFR was provided by a recent study using 201 junior high school students (Hacker, 1989). Results indicated that the CTFR did not provide a valid measure of the unitary formal reasoning construct. Although Lawson's test did initially appear to have the necessary validity for group assessment of formal reasoning, the more recent research strongly suggests otherwise.

Another group test of logical thinking, combining

written work with classroom demonstration, was developed by Tobin and Capie (1981). The Test of Logical Thinking (TOLT) was designed to measure five modes of formal reasoning: controlling variables, proportional, combinatorial, probabilistic, and correlational reasoning. Each of the ten items requires the participant to select a correct response and a justification from a number of alternatives. One difficulty with this test was the inability of some students to formulate clear justifications for their answers. In an effort to overcome this initial problem, the TOLT was modified to provide multiple-choice justification statements (Tobin & Capie, 1984). Data collected from 682 students indicated high test reliability. Evidence of criterion-related validity was also obtained. However, because the items and format of the TOLT were taken from the work of Lawson, this test is subject to the same criticisms as the CTFR. Another recent study done by Ahlawat and Billeh (1987) compared the psychometric properties of the TOLT, CTFR, and a third test of formal reasoning: Longeot's Test of Logical Thinking. Results indicated a lack of concurrent validity for all three tests. These findings contradict the earlier results of Ward, Nurrenbern, Lucas, and Herron (1981) and Farmer, Farrell, Clark, and McDonald (1982). In these two studies it was concluded that Longeot's test was both a reliable and valid instrument for assessment of formal reasoning ability.

Given the questionable validity of the three previously mentioned tests, it would almost appear as if the group paper-and-pencil tests have not adequately fulfilled the function of Piaget's original interview method. In fact, in an assessment of the relation between traditional formal operations tasks and paper-and-pencil tests, Tschopp and Kurdek (1981) found low correlations between the two and cautioned against viewing formal operations as a unitary construct. Based upon these results, it would appear that the interview is still the better method for assessing formal reasoning ability. However, strong recent support for the potential of standardized assessment of intellectual development was provided by Magolda (1987). A comparison of a standardized production format measure and the interview method resulted in a high positive correlation. So, in spite of the questionable validity of the group tests mentioned thus far, it should not be concluded that all standardized group tests assessing formal reasoning are invalid.

It was precisely because of the lack of suitable paper-and-pencil tests that several additional measures of formal operational reasoning were developed. Research reported by Roberge and Flexer (1982) and Arlin (1982) strongly suggests that there are at least two reliable and valid tests that are currently available for group assessment. Using 300 seventh and eighth graders, Roberge and Flexer (1979)

produced evidence supporting both the reliability and validity of a group assessment test, the Formal Operational Reasoning Test (FORT). For purposes of the present study, it should be noted that the subjects used in the collection of validity data were well below the age of college level students and that the test items only measured three formal operations skills (combinations, propositional logic, and proportionality). On the other hand, the Arlin Test of Formal Reasoning (ATFR) was normed and validated on an adolescent and adult population and measures eight formal operations skills to yield an overall index of formal reasoning ability (Arlin, 1982, 1984). It would appear that the ATFR would be a very practical and valid test in assessing the reasoning abilities of college students because of the norms which are provided for students beyond the high school level.

It should be pointed out at this point that the group tests mentioned thus far do not constitute an exhaustive list. There are several other tests that need to be considered. A reliability and validity study done by Patterson and Milakofsky (1980) suggested that the Inventory of Piaget's Developmental Tasks (IPDT) yields results similar to those obtained by individual interviews. Additional research done more recently by Bitner (1988, 1989) and Piburn (1989) suggests that both the Group Assessment of Logical Thinking (GALT) and the Propositional

Logic Test (PLT) have the potential to be reliable and valid measures of formal reasoning ability.

However, from a practical perspective, each of these three tests has certain limitations. The IPDT consists of 72 items, thus requiring a significantly larger amount of time to complete. In addition, the test authors themselves caution against widespread use because of inadequate norms for each age level. Because of their recent development, both the GALT and the PLT are still lacking an adequate amount of supporting empirical evidence. Additional research examining the reliability and validity of these two tests still needs to be done. Additionally, both of these tests lack norms for students beyond the high school level.

While several attempts have been made to develop group tests of formal reasoning, most have been seriously flawed by assessing less than the eight formal schemes making up the formal reasoning construct (Nagy & Griffiths, 1982). The ATFR is the only paper-and-pencil test that measures all eight of the formal schemes: multiplicative compensations, correlations, probability, combinatorial reasoning, proportional reasoning, forms of conservation, mechanical equilibrium, and frames of reference (Arlin, 1984). The ATFR consists of 32 items organized into eight subtests, with each of the subtests representing one of the eight formal schemes.

A multitrait-multimethod validity study (Arlin, 1982)

using 394 military recruits, indicates that the ATFR is both a reliable and valid measure of formal reasoning. Since its introduction, several favorable reviews of the ATFR have been published (Arter & Salmon, 1987; Fakouri, 1985). In a highly critical review of the ATFR, Santmire (1985) concluded, "the concept of formal operational reasoning is probably robust enough that the total score assessment provided by the ATFR is reasonably well correlated with level of formal operational functioning" (p.83).

Additionally, the ATFR consists of several other advantages that need to be pointed out. It can be administered in a fairly short amount of time (30-45 minutes) to large groups of students. Also, it is objective and easy to score, and it does not require special training for its administration. All items are presented in a four-response multiple-choice format, and the answers are recorded by the subjects on a standard answer sheet. A portion of each test page consists of a line drawing, which represents the problem in graphic form. This is then followed by the multiple-choice items which relate to that drawing. These features, in addition to the favorable reviews previously mentioned, qualify the ATFR as the most valid and practical instrument to be used in the present study of college student reasoning abilities. Therefore, through the process of elimination, it would seem in order that the instrument to be utilized in the present study be

that of the ATFR.

Testing Learning and Cognitive Development

The primary purpose of the present study is to determine the performance of concrete and formal thinkers on a multiple-choice test consisting of both factual and conceptual questions. In a review of the educational literature concerning measurement of cognitive skills and test construction, Coffman (1988) suggested that the focus should be on assessing abilities beyond simple recall of information. This is basically the same thing that Bloom, Englehart, Furst, Hill, and Krathwohl (1956) suggested with their proposed taxonomy for use in assessing educational goals and methods. The six hierarchical levels in Bloom's taxonomy consist of knowledge at the bottom, followed successively by comprehension, application, analysis, synthesis, and evaluation. In relation to Piaget's stages of cognitive development, it would appear that formal operational reasoning would be necessary to complete those tasks at the latter five levels of Bloom's taxonomy.

Cornbleth (1986) recommended that test items should include social studies subject matter when designing tests for measurement of critical thinking skills. This type of material often consists of abstract principles and concepts. She also points out that test items should assess application skills rather than simple rote memory. Again, formal reasoning skills would appear to be necessary to

complete those tasks of an applied nature. Simple memorization of factual types of information would be insufficient for completion of these tasks.

Again, relating this to Bloom's taxonomy, it would logically appear that those items located at the knowledge level would be of the rote memory type. Those items located in the five levels above the knowledge level would therefore be of a conceptual nature. Additional suggestions and guidelines for construction of critical thinking tests are given by Norris (1986, 1988). Each of these studies strongly suggests that if formal reasoning is to be assessed, then test items need to measure more than one level of thinking skills. Also, it is aptly pointed out in each of these studies that a strong need still exists for further research in this area.

Given the need for further research, and the possible relationship between conceptual thinking and formal reasoning, it would certainly seem in order that correlations between reasoning ability and conceptual questions be assessed. The present study makes an attempt to provide new information on this issue.

Factors Affecting Reasoning

Assessing the formal reasoning ability of college students was not the only purpose of the present research endeavor. Previous research on intellectual development has strongly implicated the role of several demographic factors

(Blackburn, 1980; Buss, 1977; Niaz, 1985, 1987, 1989). Therefore, the third purpose of the current study was to assess the relationships between the formal reasoning ability of college students and possible influencing demographic factors. Because of the conflicting results found in the available literature, it would appear that several of these factors warrant further investigation.

One research study conducted by Douglass and Wong (1977) found significant effects on reasoning scores for the factors of culture, age, and gender. Chinese and American students were given three Piagetian tasks of formal operations. Significant effects were demonstrated with Americans, older subjects, and males performing at more advanced levels. Also, studies done by Sinnott and Guttman (1978) and Sinnott (1975) using older adults, revealed significant effects of age, income, education, perceived health, and gender on performance level. Again, the males scored significantly higher. During this same time frame, two additional studies done by Lawson (1975) and Graybill (1975) provided similar results indicating that males score significantly higher than females on tests of formal reasoning. Several other recent studies done by Bitner (1987, 1989), Logan and Dungan (1990), and Williams (1989) all revealed that males scored higher than females in formal reasoning ability.

Although these findings appear to strongly favor males,

there is additional research suggesting otherwise. When examined from a different perspective, Piburn and Baker (1989) found that females turn away from science for reasons that have little or nothing to do with their reasoning ability. In fact, several other studies done by Tomlinson-Keasey (1972), Linn (1981), and Peskin (1980) revealed that the sex differences found in other studies were due to differences in interests rather than ability.

Other demographic factors that appear to be important are those of age (Hooper, Hooper, Colbert, & McMahan, 1986; Hooper, Hooper, & Colbert, 1985; Webb, 1974), culture (Mwamwenda & Mwamwenda, 1989), socioeconomic level (Kuhn, 1976) and education. This last factor has been the focus of a study assessing quality of education (Penn, Jacob & Brown, 1988) and its relationship to reasoning ability. Results indicated that both the type and quality of previous education were significant factors in the development of formal reasoning abilities. Other studies have focused on different aspects of this same factor. Schwebel (1975) found almost no relationship between logical thinking and college selection criteria (high school rank and SAT scores) for males and a low relationship for females. Niaz (1985) examined relationships between grades and reasoning ability. Results indicated that the probability of passing a course increased considerably relative to reasoning ability. Given these results, it would seem in order that additional

research is needed to assess the formal reasoning ability of college students.

Purpose of this study

The purpose of this study, then, was threefold. First, the Arlin Test of Formal Reasoning (ATFR) was administered to college students to determine the proportion of students that were functioning at each of the concrete and formal operational stages. It was hypothesized that a significant number of students would score below the formal reasoning level. Previous research done by Schwebel (1975), Niaz (1985, 1987, 1989) and Roberge and Flexer (1979) using paper-and-pencil tests of formal reasoning indicates that a large percentage of college students score below the formal level. With proper identification, these students could be taught in more concrete ways (Renner & Paske, 1977; Schneider & Renner, 1980) or taught to apply formal reasoning strategies (Danner & Day, 1977; Keeley, Browne, & Kreutzer, 1982).

The second purpose of this study was to assess the relationships and differences between ATFR scores and scores obtained on a multiple-choice test consisting of equal amounts of rote memory and conceptual questions. Based upon research regarding thinking skills, Ahlawat and Billeh (1987) and Hirschorn (1988) suggested that classroom tests need to assess critical thinking ability. Additionally, Cornbleth (1986) suggested that tests measuring student

learning and thinking skills must consist of questions beyond simple rote memory. In the present study, this was accomplished by applying Bloom's taxonomy level to the construction of multiple-choice questions.

It was hypothesized that students assessed by the ATFR at the concrete level would score significantly lower than those assessed at the formal level on conceptual type questions. Results supporting this hypothesis would indicate that classroom tests need to utilize conceptual questions when measuring student learning. Also, if significant relationships could be established between ATFR scores and performance on classroom multiple-choice tests, then CMC tests could serve as tools for ongoing assessment of student reasoning skills.

The third purpose of this study was to examine individual differences by assessing the relationships between scores on the ATFR and various demographic factors. It was hypothesized that formal reasoning ability would be significantly related to each of these factors. Previous research has yielded conflicting results for sex (Lawson, 1979; Peskin, 1980; Williams, 1989), age (Patterson & Milakofsky, 1978; Sinnott, 1975), socioeconomic status (Penn, Jacob, & Brown, 1988; Sinnott & Guttman, 1978) and occupational interest (Linn, 1981; White & Ferstenberg 1978). It was because of these conflicting results that further examination of these and other demographic factors

was needed to assess their possible effect on reasoning ability.

CHAPTER 2

METHOD

Subjects

The sample population for this study consisted of 108 men and 150 women enrolled in introductory psychology courses at a midwestern university. Student mean age was 19.45 years with a range of 17 to 39 years. Extra course credit was given to those students who participated in the study. Signed student consent forms, explaining the purpose of the study, were obtained from each participating student. In addition, the consent forms also pointed out that subjects were free to discontinue their participation at any time during the study. To ensure confidentiality, a number was assigned to each subject, and names were not used at any point during the study. In accordance with specific university policies, approval for this study was obtained from the Review Board for Treatment of Human Subjects before beginning the study.

Instruments

The Arlin Test of Formal Reasoning (ATFR) was used as an objective and valid measure of formal reasoning ability. The ATFR consists of 32 items presented in a 4-response multiple-choice format. A portion of each test booklet page consists of a line drawing, which represents the reasoning problem in graphic form. This is then followed by the multiple-choice items which relate to the various drawings.

A multi-trait multi-method validity study (Arlin, 1982) indicates that this objective test is a valid and reliable measure of formal operations. The two methods employed were the paper-and-pencil version of the ATFR and cross-validation of the instrument by individual clinical interviews with a random sample of the same subjects. Test-retest reliabilities ranged from $\underline{r}$ = .76 to $\underline{r}$ = .89 depending on the version of the test and the time period between testings. It was pointed out by the test author that because of the developmental nature of the concepts being tested, these reliabilities were as high as might be expected.

The second means of instrumentation employed in this study was a psychology classroom multiple-choice (CMC) test constructed of 35 rote-memory and 35 conceptual questions with a four-choice format. To increase the validity of this measure, selection of the two types of questions was done in accordance with the guidelines provided by the authors of the test bank for the required textbook used in the introductory psychology courses. Each multiple-choice item was defined by the test bank authors as either factual (rote) or applied (concept).

Procedures

The ATFR was administered during regular classroom time according to the specific directions provided in the test manual. Administration of this test in each of the

introductory psychology courses was completed by only one examiner over the span of one week. Before beginning the test, subjects were requested to provide, on the back of their answer sheets, information regarding age, gender, overall family size, birth order, parental education, and previous education of the student. Subjects were then given the ATFR test booklets, and administration was conducted according to the standardized procedures provided in the test manual. One week later, subjects were given the psychology CMC test as part of their regular course requirements.

Variables

Twelve variables producing scorable data were used in the present study. Student reasoning performance (Variable 1) was determined by classifying individual ATFR scores at one of the five levels of reasoning as determined by the test manual author. The five levels were: Concrete (0-7 points), High Concrete (8-14 points), Transitional (15-17 points), Low Formal (18-24 points), and High Formal (25-32 points).

Three additional variables used in this study were the subject's overall score on the CMC test (Variable 2), and the scores on each of the two types of multiple-choice questions (Variables 3,4) included in the CMC test. Variable 5 was an overall reasoning ability score consisting of two levels of concrete and formal. Concrete and formal

reasoning ability were determined by using the previous five levels of Variable 1. The Concrete group was determined by combining the bottom two levels of Variable 1 together (0-14), and the Formal group was determined by combining the top two levels together (18-32). To protect against any possible overlap between the two groups on Variable 5, the Transitional group, or middle level of Variable 1 (15-17), was discarded.

Possible effects of demographic factors on reasoning performance were assessed by employing seven demographic variables with two or more levels for each. The procedures used for dividing each of the demographic variables into various levels were done in the following manner. Age data were divided into two levels of high (21 years and above) and low (under 21 years). Family size data were divided into three levels of small (4 or less), medium (5 or 6) and large (7 or more). Birth order data were divided into three levels of youngest, middle, and oldest or only child. Parental education data for each parent were divided into three levels of low (high school or less), medium (some college or vo-tech), and high (college graduate). Previous education data were divided into two levels of new (first year in college) and experienced (previous college experience).

Statistical Design

The statistical technique used to determine the number

of students functioning at each of the five reasoning levels was to simply convert the number of individual ATFR scores at each level into percentages.

Relationships between the two assessment instruments were determined by computing three Pearson product-moment correlation coefficients. First, a coefficient was computed to determine the overall relationship between individual performance on both the administered tests. Second, a coefficient was computed to determine the relationship between individual ATFR scores and CMC rote-memory scores. Third, a coefficient was computed to determine the relationship between ATFR scores and CMC concept scores. These three correlation coefficients were computed to find relationships so that it could be determined whether or not a classroom test could be used as a general assessment tool of individual student reasoning ability.

To examine the possible differences between concrete and formal reasoning individuals and their performance on each of the two types of multiple-choice questions, three one-way analyses of variance (ANOVA) procedures were used. For each of these analyses, Variable 5 was the independent variable and Variables 2, 3, and 4 were the dependent variables. These analyses were computed to provide specific information about whether or not formal reasoning subjects score better than concrete reasoning subjects on conceptual type questions.

Analyses of the different levels of each demographic factor were done using seven one-way ANOVA procedures. For each of the seven analyses, the various demographic factors (Variables 6 to 12) served as the independent variables and the ATFR scores (Variable 1) were the dependent variable.

Chapter 3

RESULTS

The number of students scoring at each of the five ATFR levels of reasoning was obtained and converted into percentages. These results are shown in Table 1.

Table 1

Percentage of Students Scoring at Each of the Five ATFR Levels of Reasoning (N = 258).

Level of Reasoning	N	Percentage
Concrete	6	2.5
High Concrete	83	32.0
Transitional	66	25.5
Low Formal	89	34.5
High Formal	14	5.5

It can be seen that only 40 percent of the sample population scored in the formal reasoning range, with only 5.5 percent of the subjects scoring at the highest possible level. It should be noted that the 25.5 percent of subjects who scored at the Transitional level were not considered as formal reasoning students.

In addition to the score obtained for each subject on the ATFR, three other scores were obtained from administration of the CMC Test. Test ranges, means,

standard deviations, and correlations between the ATFR and the CMC Test are presented in Table 2.

Table 2

Mean Test Scores, Standard Deviations, and Correlations between ATFR and CMC Scores.

Test	M	SD	Correlation
Arlin	16.39	4.80	
CMC Total	36.02	6.75	.41*
CMC Rote	19.44	3.66	.36*
CMC Concept	16.58	3.99	.36*

*$p<.001$

It can be seen that a significant correlation was obtained between the ATFR scores and the CMC total scores, $r = .41$, $p<.001$. In addition, significant and identical correlations were obtained between the ATFR scores and those obtained from each of the two CMC subtests, $r = .36$, $p<.001$. Although these two correlations were identical, it should be noted that the mean performance score for the 35 concept items was 2.86 points lower than the scores obtained for the rote items.

To determine the effects of reasoning ability on CMC performance scores, three one-way analyses of variance (ANOVA) procedures were conducted. Mean scores and ANOVA

results for each analysis are shown in Table 3.

Table 3

Mean Scores and ANOVA Results of Reasoning Ability on CMC Performance Scores (N=192).

Test	Reasoning Ability				
	Concrete		Formal		
	N	M	N	M	F
CMC Total	88	32.80	104	38.80	43.33*
CMC Rote	88	17.89	104	20.90	35.76*
CMC Concept	88	14.91	104	17.89	30.84*

*$p<.001$

It can be seen that a significant main effect of reasoning ability on CMC Total scores was obtained; $F(1, 190 = 43.33$, $p<.001$, with the formal reasoning group scoring significantly higher than the concrete group. Significant main effects of reasoning ability on CMC Concept scores; $F(1, 190) = 30.84$, $p<.001$, and on CMC Rote scores; $F(1, 190) = 35.76$, $p<.001$, were also obtained. For each analysis it was found that the formal reasoning group scored significantly higher than the concrete group. In spite of the fact that three analyses were conducted, adjustment of the alpha level was deemed unnecessary because the obtained

significance level for each analysis was less than .001. It should also be noted that mean performance scores for both reasoning groups were consistently lower on concept items.

To determine the effects of each of the demographic variables on ATFR performance scores, seven one-way ANOVA's were computed. Mean scores and ANOVA results are presented in Table 4.

Table 4

Mean Scores and ANOVA Results of Demographic Variables on ATFR Scores (N=258).

Variable	Group	N	M	F
Age	Low	216	16.31	.04
	High	42	16.79	
Gender	Male	108	17.02	3.23
	Female	150	15.93	
Family Size	Small	108	15.99	.65
	Medium	108	16.72	
	Large	42	16.55	
Birth Order	Youngest	94	15.70	2.59
	Middle	57	17.52	
	Oldest	107	16.38	
Father's Education	Low	93	16.80	.56
	Medium	90	16.07	
	High	75	16.27	
Mother's Eduction	Low	97	16.70	.34
	Medium	96	16.15	
	High	65	16.28	
Previous Education	New	206	16.42	.04
	Experienced	52	16.27	

The ANOVA results revealed no main effects for any of the seven demographic variables on ATFR scores. Although multiple analyses were conducted, the alpha level for each analysis was not adjusted because of the fact that no significant differences were obtained at the .05 level. In spite of the lack of significant differences obtained for each of the demographic variables, it should be noted that the means of the gender and birth order factors showed noticeable differences between the levels.

CHAPTER 4

DISCUSSION

Piaget's theory of cognitive development proposes that concrete operational reasoning ends approximately at the beginning of adolescence, thus ushering in the ability to think hypothetically and abstractly (Piaget & Inhelder, 1969; Piaget, 1972). However, research studies done by Bart (1971), Kuhn (1979), Logan and Dungan (1990), and others, has indicated that many late adolescents and adults are still functioning in the concrete stage of thinking.

The results of the present study indicate that only forty percent of the students tested on the ATFR were functioning at the formal operational reasoning level. These results are not in agreement with Piaget's theory of cognitive development, thus providing additional evidence that there is presently a strong need to accommodate concrete thinkers at the college level. However, the key factor is first to identify these students, so that they may be taught in more concrete ways (Renner & Paske, 1971; Schneider & Renner, 1980) or taught to apply formal reasoning strategies (Danner & Day, 1977; Keeley, Browne & Kreutzer, 1982).

Inhelder and Piaget (1958) originally utilized the clinical interview method to identify specific stages of cognitive development. However, this method of individual assessment is a very time consuming process which involves

the use of a trained evaluator. These limitations have rendered the clinical interview method an impractical tool for assessment of large groups of students, thus creating a need for the development of group paper-and-pencil tests. Although there are currently several group tests of formal reasoning available, most are seriously flawed in one way or another (Nagy & Griffiths, 1982). However, based upon several critical reviews by Fakouri (1985) and Santmire (1985), it appears that the ATFR (Arlin, 1982, 1984) is both a valid and reliable instrument of formal reasoning ability.

The present study attempted to establish relationships between the ATFR and CMC tests, so that the latter tests could be utilized as ongoing assessment tools of student reasoning abilities. The significant relationships obtained between the two types of tests, provided some evidence that CMC tests could be used in the above mentioned manner. The significant correlations obtained between the ATFR and each of the CMC scales indicates practical significance for CMC tests. In the construction of any test that measures complex human factors, practical significance is indicated when correlations account for even small amounts of variance. Although higher correlations would have been more desirable, it can still be concluded that CMC tests, utilizing both conceptual and rote-memory items, have the potential to measure general reasoning ability.

In spite of the identical correlations obtained between

the ATFR and the two CMC subtests, student mean scores were almost three points higher on the rote-memory items. These results indicate that conceptual items were considerably more difficult to complete successfully than were the rote-memory items. These results are in accordance with the findings of Bloom (1956), Cornbleth (1986), and Norris (1986, 1988), each of whom suggested that classroom tests need to assess more than one type of thinking.

The present study also provided strong support for the hypothesis that formal reasoning students would perform better than concrete thinkers on conceptual items. The significant main effects of reasoning ability on each of the CMC scores showed quite clearly that formal or abstract reasoning ability increases accuracy of performance on both types of CMC items. These findings are in agreement with those of Niaz (1985) who found that reasoning ability was significantly related to overall grades. These results indicate that current measures of student learning need to utilize conceptual type items if higher order cognitive processes are to be measured. The present results also provide additional evidence that CMC tests can be utilized by classroom instructors to differentiate between concrete and formal reasoning students.

The last issue addressed by this study was the possible effect of demographic variables on reasoning ability. The absence of main effects for each of the seven demographic

variables indicates that classroom instructors need not be concerned with these demographic factors when attempting to identify reasoning abilities of individual students. These results contradict the findings of Lawson (1975), Peskin (1980), and Williams (1989) who each reported that males scored significantly higher than females on formal reasoning tests. The results of the present study also contradict the findings of Patterson and Milakofsky (1978) and Sinnott (1975) who reported significant effects of age on reasoning ability. The remaining five demographic factors were examined as exploratory research, and as previously mentioned there were no main effects obtained.

In summary, the results of this study support the assertion that a large number of college students are currently functioning below the formal operational level of reasoning, and that the lack of ability to think in abstraction impairs performance on conceptual type questions. This would seem to indicate that a strong need exists to accommodate concrete thinkers at the college level. The present results also provide evidence that CMC tests can be utilized as ongoing assessment tools to identify individual student reasoning skills.

The lack of effect of each of the seven demographic factors on reasoning ability suggests that these specific individual differences do not have any significant impact on overall cognitive development. Because the present study

was conducted on a select population, it is suggested that further research be conducted on both the factors of age and gender to shed further light on the conflicting results that presently exist. In addition, the results of this study also provide some structure and framework for further research to be conducted on CMC tests in relation to assessment of reasoning ability for college students.

References

Ahlawat, K.S., & Billeh, V.Y. (1987). The comparative investigation of the psychometric properties of three tests of logical thinking. Journal of Research in Science Teaching, 24(2), 93-105.

Arlin, P.K. (1982). A multitrait-multimethod validity study of a test of formal reasoning. Educational and Psychological Measurement, 42, 1077-1088.

Arlin, P.K. (1984). Arlin test of formal reasoning: test manual. East Aurora, New York: Slossan Educational Publications.

Arter, J.A., & Salmon, J. (1987). Assessing higher order thinking skills. Portland, OR.: Northwest Regional Education Laboratory.

Bart, W.M. (1971). The factor structure of formal operations. British Journal of Educational Psychology, 41, 70-77.

Bitner, C. (1987). The GALT: A measure of logical thinking ability of seventh through twelfth grade students. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

Bitner, C. (1988). Is the GALT a reliable instrument for measuring the logical thinking abilities of students in grade six through twelve? Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

Bitner, B.L. (1989). Developmental patterns in logical reasoning of students in grades six through ten: Increments and plateaus. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

Blackburn, J.A. (1980). The influence of personality, curriculum, and memory correlates on formal reasoning in young adults and elderly persons. Journal of Gerontology, 39(2), 207-209.

Bloom, B.S., Englehart, M.D., Furst, G.H., Hill, W.H., & Krathwohl, D.R. (1956). Taxonomy of educational objectives: Handbook I, the cognitive domain. New York: David McKay Company, Inc.

Buss, A.R. (1977). Piaget, Mack, and Buck-Morss on cognitive development: A critique and reinterpretation. Human Development, 20(2), 118-128.

Coffman, W.E. (1988). Measurement of thinking skills: An historical perspective. Iowa Testing Programs Occasional Papers, No. 34.

Cornbleth, C. (1986). Assessing skills and thinking in Social Studies. Paper commissioned by the Study Group on the National Assessment of Student Achievement.

Danner, F.W., & Day, C.M. (1977). Eliciting formal operations. Child Development, 48(4), 1600-1606.

Douglass, J.D., & Wong, A.C. (1977). Formal operations: Age and sex differences in Chinese and American children.

Child Development, 48(2), 689-692.

Fakouri M.E. (1985). Arlin Test of Formal Reasoning. In D.J. Keyser & R.C. Sweetland (Eds.), Test Critiques, Vol. 2, (pp. 40-44). Kansas City, MO: Test Corporation of America.

Farmer, W.A., Farrell, M.A., Clark, R.M., & McDonald, J. (1982). A validity study of two paper-pencil tests of concrete and formal operations. Journal of Research in Science Teaching, 19(6), 475-485.

Flavell, J.H. (1977). Cognitive development. Englewood Cliffs, N.J.: Prentice Hall.

Gray, W.M. (1978). Standardized tests based on developmental theory. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31).

Graybill, L. (1975). Sex differences in problem-solving ability. Journal of Research in Science Teaching, 12 (4), 341-346.

Hacker, R.G. (1989). The construct validities of some group tests of intellectual development. Educational and Psychological Measurement, 49(1), 269-275.

Hirschorn, M.W. (1988). Liberal education urged for all in high school, not just the students who are college bound. Chronicle of Higher Education, 34(37), 33-34.

Hooper, F.H.; Hooper, J.O., & Colbert, K.K. (1985). Personality and memory correlates of intellectual

functioning in adulthood: Piagetian and psychometric assessments. Human Development, 28(2), 101-107.

Hooper, J.O.; Hooper, F.H.; Colbert, K., & McMahan, R. (1986). Cognition, memory, and personality in elderly students. Educational Gerontology., 12(3), 219-229.

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

Keeley, S.M., Browne, M.N., & Kreutzer J.S. (1982). A comparison of freshman and seniors on general and specific essay tests of critical thinking. Research in Higher Education, 17(2), 139-154.

Kuhn, D. (1976). Relation of two Piagetian stage transitions to IQ. Developmental Psychology., 72(2), 157-161.

Kuhn, Deanna (Ed.) (1979). Intellectual development beyond childhood, new directions in child development: Vol.5. San Francisco: Jossey Bass.

Lawson, A.E. (1975). Sex differences in concrete and formal reasoning ability as measured by manipulative tasks and written tasks. Science Education, 59(3), 397-405.

Lawson, A.E. (1978). The development and validation of a classroom test of formal reasoning. Journal of Research in Science Teaching, 15(1), 11-24.

Lawson, A.E. (1979). Relationships among performance on group administered items of formal reasoning. Perceptual

and Motor Skills, 48, 71-78.

Linn, M.C. (1981). Male-female differences in formal thought. Paper presented at the Biennial Meeting of the Society for Research in Child Development.

Logan, R.A., & Dungan,D. (1990). An examination of college student reasoning abilities. Paper presented at the Annual Convention of the Association for Psychological and Educational Research in Kansas.

Magolda, M.B. (1987). Comparing open-ended interviews and standardized measures of intellectual development. Journal of College Student Personnel, 28(5), 443-448.

Mwamwenda, T.S., & Mwamwemda. B B. (1989). Formal operational thought among African and Canadian college students. Psychological Reports, 64(1), 43-46.

Nagy, P., & Griffiths, A.K. (1982). Limitations of recent research relating Piaget's theory to adolescent thought. Review of Educational Research, 513-556.

Neimark, Edith D. (1975). Intellectual development during adolescence. In Francis Degen Horowitz (Ed.) Review of research in child development, Vol. 4, (pp. 541-594). Chicago, IL: University of Chicago Press.

Niaz, M. (1985). Evaluation of formal operational reasoning by Venezuelan freshman students. Research in Science and Technological Education, 3(1), 43-50.

Niaz, M. (1987). The role of cognitive factors in the teaching of science. Research in Science and

Technological Education, 5(1), 6-16.

Niaz, M. (1989). The role of cognitive style and its influence on proportional reasoning. Journal of Research in Science Teaching, 26(3), 221-235.

Norris, S.P. (1986). Evaluating critical thinking ability. History and Social Science Teacher, 21(3), 135-146.

Norris, S.P. (1988). Research needed on critical thinking. Canadian Journal of Education, 13(1), 125-137.

Patterson, H.O., & Milakofsky, L. (1978). A paper-pencil inventory for the assessment of Piaget's tasks. Paper presented at the Annual Meeting of the American Psychological Association.

Patterson, H.O., & Milakofsky, L. (1980). A paper-and-pencil inventory for the assessment of Piaget's tasks. Applied Psychological Measurement, 4(3), 341-353.

Penn, N.E.; Jacob, T.C., & Brown, M. (1988). Comparison between Gorham's Proverbs Test and the revised Shipley Institute of Living Scale for a Black population. Perceptual and Motor Skills, 66(3), 839-845.

Peskin, J. (1980). Female performance and Inhelder and Piaget's tests of formal operations. Genetic Psychology Monographs, 101(2), 245-256.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. Human Development, 15, 1-12.

Piaget, J., & Inhelder, B. (1969). The psychology of the child. Helen Weaver (Trans.). New York: Basic Books.

Piburn, M.D. (1989). Reliability and validity of the Propositional Logic Test. Educational and Psychological Measurement, 49(3), 667-672.

Piburn, M.D., & Baker, D.R. (1989). Sex Differences in formal reasoning ability: Task and interviewer effects. Science Education, 73(1), 101-113.

Pratt, C., & Hacker, R.G. (1984). Is Lawson's Classroom Test of Formal Reasoning valid? Educational and Psychological Measurement, 44(2), 441-448.

Renner, J.W., & Paske, W.C. (1977). Comparing two forms of instruction in college physics. American Journal of Physics, 45(9), 851-859.

Roberge, J.J., & Flexer, B.K. (1979). Further examination of formal operational reasoning abilities. Child Development, 50(2), 478-484.

Roberge, J.J., & Flexer, B.K. (1982). The Formal Operational Reasoning Test. Journal of General Psychology, 106(1), 61-67.

Santmire, T.E. (1985). Review of the Arlin Test of Formal Reasoning. In J.V. Mitchell, Jr. (Ed.), The Ninth Mental Measurements Yearbook, Vol 1, (pp. 81-83). Lincoln, NE: University of Nebraska Press.

Schneider, L.S., & Renner, J.W. (1980). Concrete and formal teaching. Journal of Research in Science Teaching, 17(6), 503-517.

Schwebel, M. (1975). Formal operations in first-year

college students. The Journal of Psychology, 91, 133-141.

Shayer, M., Adey, P., & Wylam, H. (1981). Group tests of cognitive development: Ideals and realization. Journal of Research in Science Teaching, 18(2), 157-168.

Sinnott, J.D. (1975). Everyday thinking and Piagetian operativity in adults. Human Development, 18(6), 430-443.

Sinnott, J.D., & Guttman, D. (1978). Piagetian logical abilities and older adults abilities to solve everyday problems. Human Development, 21(5-6), 327-333.

Tobin, K.G., & Capie, W. (1981). The development and validation of a group test of logical thinking. Educational and Psychological Measurement, 41(2), 413-423.

Tobin, K.G., & Capie, W. (1984). The Test of Logical Thinking. Journal of Science and Mathematics Education in Southeast Asia, 7(1), 5-9.

Tomlinson-Keasey, C. (1972). Formal operations in females from eleven to fifty-four years of age. Developmental Psychology, 6(2), 364.

Tschopp, J.K., & Kurdek, L.A. (1981). An assessment of the relation between traditional and paper-pencil formal operations tasks. Journal of Research in Science Teaching, 18(1), 87-91.

Ward, C.R., Nurrenberg, S.C., Lucas, C., & Herron, J.D.

(1981). Evaluation of the Longeot Test of Cognitive Development. Journal of Research in Science Teaching, 18(2), 123-130.

Webb, R.A. (1974). Concrete and formal operations in very bright six to eleven year olds. Human Development, 18(4), 292-300.

White, K.M., & Ferstenberg, A. (1978). Professional specialization and formal operations: The balance task. Journal of Genetic Psychology, 133(1), 97-104.

Williams, W.L. (1989). A comparative study of logical thinking skills: West German data. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

TO: All Graduate Students Who Submit a Thesis or Research Problem/Project as Partial Fulfillment of the Requirements for an Advanced Degree

FROM: Emporia State University Graduate School

I, ROBERT A. LOGAN, hereby submit this thesis/report to Emporia State University as partial fulfillment of the requirements for an advanced degree. I agree that the Library of the University may make it available for use in accordance with its regulations governing materials of this type. I further agree that quoting, photocopying, or other reproduction of this document is allowed for private study, scholarship (including teaching) and research purposes off a nonprofit nature. No copying which involves potential financial gain will be allowed without written permission of the author.

Robert A. Logan
Signature of Author

4-25-91
Date

RELATIONSHIPS between the ARLIN TEST OF FORMAL REASONING AND CLASSROOM MULTIPLE-CHOICE TESTS.
Title of Thesis/Research Project

Jackie Tolbert
Signature of Graduate Office Staff Member

April 25, 1991
Date Received

Distribution: Director, William Allen White Library
Graduate School Office
Author