R E P O R T   R E S U M E S

ED 014 765   24   CG 000 915

AN EXPERIMENT TO IMPROVE THE REASONING ABILITY OF SEVENTH-GRADE STUDENTS. FINAL REPORT.

BY- HOWELL, EDGAR N.
TEMPLE UNIV., PHILADELPHIA, PA.

REPORT NUMBER BR-6-8110   PUB DATE 15 JUN 67
CONTRACT OEC-6-10-361
EDRS PRICE MF-$0.50 HC-$2.68 65P.

DESCRIPTORS- *GRADE 7, *RESEARCH, ANALYSIS OF VARIANCE, INTELLIGENCE QUOTIENT, TESTS OF SIGNIFICANCE, SEX (CHARACTERISTICS), THOUGHT PROCESSES, *LOGIC, SYMBOLIC LANGUAGE, *TEACHING, *PRODUCTIVE THINKING, LORGE THORNDIKE INTELL. TESTS-LEVEL 4, TEST OF INFERENCE PATTERNS, INFERENTIAL REASONING ANALYSIS

A UNIT ON EFFECTIVE THINKING, WITH THE EXPERIMENTAL FOCUS ON INFLUENTIAL THINKING, WAS PREPARED AND TESTED. A TEXT, "EFFECTIVE THINKING," WAS PREPARED. AN INEXPERIENCED TEACHER USED THE TEXT IN A NINE WEEK COURSE WITH SEVENTH GRADE STUDENTS. ONLY A POST-TEST CONTROL GROUP DESIGN WAS USED. THE TWO FACTORS STUDIED WERE TREATMENT AND SEX. THE TEST OF INFERENCE PATTERNS (TIP), USED BY HOWELL IN AN EARLIER STUDY, WAS THE CRITERION MEASURE AT THE END OF THE FIRST NINE WEEK PERIOD. STUDENT IQ'S WERE ALSO AVAILABLE. THE INFERENTIAL REASONING ANLYSIS (IRA) BY HOWELL WAS THE CRITERION MEASURE USED FOR THE SECOND NINE WEEK PERIOD. ANALYSIS OF VARIANCE OF TOTAL TEST SCORES ON TIP YIELDED SIGNIFICANT DIFFERENCES BETWEEN THE EXPERIMENTAL AND CONTROL GROUPS. THERE WAS NO SIGNIFICANT DIFFERENCE BETWEEN THE SCORES OF MALE AND FEMALE STUDENTS OR BETWEEN TREATMENT AND SEX. USING A COVARIANCE ADJUSTMENT, IT WAS FOUND THAT IQ HAD AN EFFECT ON THE CRITERION. ANALYSIS OF VARIANCE OF PART I, II, IV, AND TOTAL SCORES OF THE IRA PRODUCED SIGNIFICANTLY HIGHER SCORES FOR THE EXPERIMENTAL GROUP. SIGNIFICANT DIFFERENCES WERE FOUND BETWEEN THE SEXES FOR TOTAL IRA SCORES. NO SIGNIFICANT INTERACTION WAS FOUND BETWEEN TREATMENT AND SEX. IN GENERAL, RESULTS INDICATED THAT SEVENTH GRADERS CAN PROFIT FROM INSTRUCTION IN INFERENTIAL REASONING, AND THAT HIGHER IQ STUDENTS WILL PROFIT MORE FROM THIS INSTRUCTION. (SK)

FINAL REPORT

PROJECT NO. 6-8110
CONTRACT NO. OE 6-10-361

AN EXPERIMENT TO IMPROVE THE REASONING ABILITY OF SEVENTH-GRADE STUDENTS

JUNE 1967

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

OFFICE OF EDUCATION
BUREAU OF RESEARCH

AN EXPERIMENT TO IMPROVE THE REASONING ABILITY OF SEVENTH-GRADE STUDENTS

Proj. No. 6-8110
Contract No. OE 6-10-361

Edgar N. Howell

June 15, 1967

Temple University
Philadelphia, Pa.

CONTENTS

INTRODUCTION

Effective thinking is generally desired as an outcome of the educative process. The development of the ability to think has been rated as an important purpose of American secondary education by the Educational Policies Commission of the National Education Association. The explosive production of new subject matter, especially in science and mathematics, during the last few decades has effected an almost overwhelming burden of content for students to master. Yet, learning to think and mastering subject content ought not be diametrically opposed. Once the fundamental structure of a subject, such as the undefined terms and postulates in mathematics, is learned, there is a clear need for effective thinking to discover and prove propositions inferred by the assumptions. Much classroom activity is devoted to the presentation and mastery of content. Effective thinking is usually a by-product of instruction; it is generally learned implicitly and not until the tenth grade. The problem was to determine whether certain aspects of effective thinking could be successfully learned in a direct manner as early as the seventh grade.

The seventh-grade curriculum at Hudson Junior High School in Wisconsin has included a daily period for student exploration into topics not introduced into the standard courses. This exploratory program was divided into four parts: speech, art, logic and either industrial education or home economics. All students took each of the four courses with boys taking industrial education and girls taking home economics. Each course was of nine weeks duration. The logic course has been the center of discussion in this report, for it represented an attempt to teach students how to arrive at well-reasoned conclusions to the problems they face. Such conclusions could be attained if the students knew how to think clearly, to separate fact from opinion, to recognize and evaluate propaganda, and to reason deductively and inductively. These desirable student behaviors represented correlates to the objectives of the logic course.

Logic was taught for the first time during the 1964-5 school year. The content for the course was selected from standard college textbooks on logic, for these were the only sources known to the teacher at that time. The vocabulary was laden with traditional terms from classical logic. In the Spring and Fall of 1965, Howell [7] was invited to study whether the logic course, as presented, prepared the students to correctly identify valid conclusions of selected inference patterns (cf. Appendix A). Analysis of the experimental data justified the assertion that the logic course did not provide a suitable medium of transfer to satisfactory achievement in inferential reasoning. The teacher of this course realized that the materials chosen for use were not being understood by the students. These evaluations pointed to the need for reorganizing the course content and preparing suitable written materials for the students.

A search of the literature revealed only a few studies on the logical reasoning ability of students in grades 6, 7, and 8.

Hyram [9] experimented with students of grades 7 and 8 who had an IQ of at least 110. He instructed the experimental group in the nature of thinking, the tools of thinking, the nature of definition, the nature of eductive inference, the nature of deductive inference, the nature of experimentation, and common errors in reasoning. Significant differences were found between the matched experimental and control groups when measured by a self-constructed test of reasoning ability. Hyram's major conclusions were that correct thinking is a function of one's knowledge of the principles of logic, and that seventh- and eighth-grade students can be taught to think logically by means of instructional units in logic.

Maw [13] investigated whether intermediate-grade children from middle and upper-middle suburban neighborhoods could improve their critical thinking abilities under instruction. She found highly significant differences in favor of the experimental group on her Test of Critical Thinking. She concluded that the lessons were effective in improving those thinking skills tested in her instrument.

Corley [2] studied the ability of students of grades 6, 7, 8, and 10 to achieve well, both in an introduction to methods of reaching general conclusions and in an introduction to demonstrative geometry. He found that sixth-grade students were able to apply the methods of intuition, inductive reasoning, and deductive reasoning moderately well in both geometric and non-geometric situations. He also noted that this ability advanced rapidly during the sixth grade and then increased slowly from the seventh to tenth grades.

Howell [6] questioned whether junior high school students in an accelerated mathematics program could recognize correct conclusions to selected inference patterns even though they had received no formal training in logic. Analysis of his Test of Selected Inference Patterns revealed support for the hypotheses that growth in inferential reasoning ability without formal instruction in logic improves slightly with increasing grade level and that this ability favors neither sex. More detailed inspection of the test revealed that the seventh-grade students understood three of the ten inference patterns, viz., and-elimination, hypothetical syllogism, and rule of detachment; that the eighth-grade students understood the preceding three and also the pattern called and-introduction; and that the ninth-grade students understood the former four as well as the rule of contraposition.

The preceding study indicated that junior high school students of above-average ability in mathematics learn inference patterns at a progressive but slow rate without formal instruction. The other studies indicated that students of grades 6, 7, and 8 can learn elementary aspects of effective thinking with formal instruction. The question arose as to what extent all seventh-grade students could profit from explicit instruction in inference patterns. Therefore, another purpose of this investigation was to determine experimentally whether seventh-grade students could profit from instruction in patterns of inferential reasoning.

The objectives of this study were to: (1) prepare a textbook on effective thinking which contained a unit on inference patterns, (2) select or prepare a suitable instrument to measure ability to reason inferentially, and (3) establish an experimental situation to test student ability to reason inferentially.

METHOD

The method section of this report has been divided into three parts to match the foregoing objectives. They are: (1) the text, (2) the tests, and (3) the experimental procedure.

A. The Text.

The initial phase of the text preparation involved the establishment of a definition of effective thinking. Several were examined for suitability and similarity. The writing team selected that proposed by Ennis [3] in which he equated critical (or effective) thinking with the correct assessing of statements. This implied the need for a thorough examination of each statement and of the entire set of statements leading to the conclusion. Hence the text, Effective Thinking [8], was subdivided into two major parts: statements, and reasoning. The chapters of the first part dealt with observation, authority, and statement errors; those of the second part concerned inductive, deductive, and faulty reasoning.

The text was prepared during June, July, and August of 1967 by a writing team consisting of Edgar N. Howell, Elmer A. Mellum, and Noel Schumacher. Dr. Howell was an associate professor of mathematics and mathematics-education at Wisconsin State University at River Falls. Mr. Mellum was the seventh-grade science teacher at Hudson Junior High School. Mr. Schumacher was the principal of Hudson Junior High School and had taught social studies prior to assuming his administrative post. Delphine Johnson, language-arts teacher at Hudson Junior High School, helped with the final revision of the text.

One of the important guidelines in the preparation of this material was that of restraining the selection of terms and the style of writing to the level of the student. To adjust for the almost practical impossibility of achieving this, it was assumed that the teacher of the logic course would inherit the responsibility of guiding the students through the more difficult parts.

A copy of Effective Thinking has been submitted with this final report.

B. The Tests.

A search of the usual commercial sources revealed a lack of standardized tests of reasoning ability appropriate for use with seventh-grade students. Therefore, two locally devised instruments were used as measures in the experiment. The first was a scale called Test of Inference Patterns, which has been frequently abbreviated TIP throughout this report. The second was named Inferential Reasoning Analysis and abbreviated IRA. A copy of Test of Inference Patterns has been placed in Appendix B, while a copy of Inferential Reasoning Analysis has been placed in Appendix C.

TIP was a revision of Test of Selected Inference Patterns, a copyrighted instrument prepared by Howell for use in his doctoral dissertation [6]. TIP was designed for and employed first in a previous experiment [7] at Hudson Junior High School. Eight valid and two invalid inference patterns were chosen for use in TIP; they have been listed in Appendix A. The test was divided into two parts: in the first, the patterns were presented in statement form; in the second, the statements were replaced by capital letters. Within each part there were three occurrences of each pattern. Numbering of the exercises was done randomly. The assignment of either a correct or an incorrect response to each valid exercise was made at random. Although the invalid patterns could not require a unique response, the choice of an apparently correct or incorrect response was made randomly.

Since valid and invalid inference patterns were intermingled in TIP, it was conjectured that student learning of the valid patterns might be more accurately displayed on a test which contained valid patterns only at the beginning. Therefore, in December 1966, the project director prepared a new instrument, Inferential Reasoning Analysis, which consisted of four parts. Parts I and II were similar to those of TIP, except that they lacked examples of three patterns used in TIP, viz., not-introduction, converse pattern, and inverse pattern. Not-introduction, although valid, was deleted from IRA because the students were not taught this pattern and because it was not explained in Effective Thinking. However, four examples of each of the seven valid inference patterns occurred in each of the first two parts of IRA. Part III contained ten exercises, each requiring the use of two valid inference patterns to achieve a complete correct response. Five possible responses were assigned to each example. Only two of these were correct and one used more information from the premisses than the other. The 'more correct' response was allotted two points, the other one point. Part IV contained five examples of the invalid patterns mixed randomly with five examples of valid ones.

C. The Experimental Procedure.

All 157 students of the 1966-7 seventh-grade class at Hudson Junior High School were required to take the logic course. Four exploratory program sections were to be formed, and one section would take logic per quarter. Therefore, only one experimental treatment could be tried during that quarter on one section and another section would be available as a control. To minimize possible reactive arrangements and to avoid the loss of an instructional period, it was decided not to give a pretest. Under these circumstances the posttest-only control group design [1] was judged the most appropriate for the experiment. Since a different section would be taking logic during the second quarter, it was decided to conduct the experiment twice, the second one being a modification rather than a replication of the first.

The administration and faculty desired to have heterogeneous grouping within all sections of the seventh grade. To achieve this they had devised a ranking procedure whereby each student was assigned a composite score based upon the following criteria: (1) his score on the Otis Quick-Scoring Mental Ability Tests, Form A, given in the fourth grade, (2) his score on the Iowa Tests of Basic Skills given during the last month of the sixth grade, and (3) his score on a five-point scale of sixth-grade classroom performance evaluated by his teacher. Students were then ranked on the basis of these composite scores, starting with the highest. Groups of four were then formed. Within each of these groups, students were assigned randomly to the exploratory sections. Students lacking a complete set of scores were assigned to the sections in a completely random manner.

One of the four exploratory sections was chosen at random to take the logic course during the first quarter; this section was called the experimental group for that quarter. Another section was selected randomly to be the control group for that quarter; most of the students in this section took the exploratory course in speech during the first nine-week period. Since it was decided to run the experiment twice, one of the two remaining sections was chosen to take the logic course during the second quarter. The other section became the control group for that quarter; most of its students took art at that time.

Three weeks of the quarter were allotted for instruction on deduction. Thus 15 class periods, each of 42-minute duration, were available to the teacher for presentation and testing on syllogisms, inference patterns, and proof.

At the end of the first quarter TIP was administered to both treatment groups. IRA was administered to the remaining sections at the close of the second quarter in mid-January of 1967.

To increase the power of the significance test [1], analysis of covariance was chosen for the statistical examination of the experimental data. The Lorge-Thorndike Intelligence Tests [10,11], both verbal and nonverbal batteries, were selected as covariate measures. These tests were considered appropriate for this purpose both because the authors had designed these instruments to measure reasoning ability and because they have been found to be good indices of scholastic aptitude [12]. The nonverbal battery [11] was used separately from the verbal battery [10] as a covariate to determine whether there would be different effects on the criteria since at least one-tenth of the seventh-grade students were receiving remedial reading instruction.

Student scores on the verbal and nonverbal batteries of the Lorge-Thorndike Intelligence Tests and on all parts and totals of both TIP and IRA were first examined by analysis of variance so that changes due to the covariance adjustment could be observed. A 2×2 factorial design was chosen with factors of treatment and sex. The two levels of treatment were experimental and control; those of sex were male and female. Since the cell frequencies were not expected to be equal under the sampling technique, the analyses of variance and covariance were performed by the method of unweighted means [14]. The 5% level of significance was chosen for all statistical tests in this experiment.

To determine the degree of association between the independent variables (treatment and sex) and the dependent variable (IQ or TIP or IRA), omega-squared [5] was calculated for each obtained F-ratio in the analyses of variance and covariance.

Several post hoc tests were performed. The reliability coefficients of TIP and IRA for both the experimental and control groups were determined by the split-half method with correction for attenuation [4]. Correlation coefficients [4] were calculated for all pairs of scales and subscales used in the experiment. These were tested for significance from zero and for significance between treatment groups. To determine which inference patterns were understood by treatment groups and subgroups, the proportion of students correctly responding to five or six occurrences of each pattern on TIP and to six or seven or eight on IRA were computed. Selected pairs of these proportions were examined for significant differences [4] between contrasting groups.

RESULTS

The results from all analyses of the data collected during the first experiment have been reported in their entirety. Following this, the results from the second experiment have been reported in their entirety.

The F-ratios and estimated omegas-squared from the analyses of variance and covariance of the tests used in the first experiment have been tabulated in Table 1. The summary statistics, test for homogeneity of variance, and summary tables for the analyses of variance and covariance which form the basis for Table 1 have been placed in Appendix D, Tables 13 through 16. Bartlett's test for the homogeneity of variance [14] failed to reveal any significant differences among cell variances for the five scales or subscales involved in the first experiment.

The results of the analyses of variance and covariance in the first experiment led to the rejection of the null hypotheses of no significant differences between:

1) treatment means on Part II of TIP,
2) treatment means on Part II of TIP when adjusted by the verbal battery scores on the Lorge-Thorndike test,
3) treatment means on Part II of TIP when adjusted by the nonverbal battery scores on the Lorge-Thorndike test,
4) treatment means of TIP,
5) treatment means of TIP when adjusted by the verbal battery scores on the Lorge-Thorndike test, and
6) treatment means of TIP when adjusted by the nonverbal battery scores on the Lorge-Thorndike test.

The null hypothesis for no interaction between treatment and sex on the verbal battery of the Lorge-Thorndike Intelligence Tests was also rejected.

The degree of association between independent variables and dependent variables were insignificant except for the treatment effect measured by Part II of TIP and the total score of TIP.

Both F-ratios and omegas-squared for the Part II scores on TIP and the total scores on TIP increased under the covariance adjustments for intelligence quotients whether verbal or nonverbal. However, the increase was larger when the verbal battery scores were used as the covariate.

TABLE 1. F-RATIOS AND ESTIMATED OMEGAS-SQUARED FROM THE ANALYSES OF VARIANCE AND OF COVARIANCE IN THE FIRST EXPERIMENT.

Variable		Source of Variation					
Covariate	Criterion	Treatment		Sex		Interaction	
		F	O^2	F	O^2	F	O^2
	LT-V	.56	.00	.07	.00	4.29	.00
	LT-NV	.00	.00	.09	.00	.97	.00
	TIP-I	1.90	.01	.00	.00	.32	.00
LT-V	TIP-I	3.64	.03	.00	.00	.15	.00
LT-NV	TIP-I	2.17	.02	.00	.00	.06	.00
	TIP-II	11.07	.12	.16	.00	.22	.00
LT-V	TIP-II	16.30	.17	.34	.00	.22	.00
LT-NV	TIP-II	13.86	.14	.36	.00	.00	.00
	TIP-T	8.86	.09	.05	.00	.36	.00
LT-V	TIP-T	15.46	.16	.20	.00	.30	.00
LT-NV	TIP-T	11.52	.12	.19	.00	.02	.00

$F(1,75) = 3.97$ at the 5.0% level of significance
$F(1,75) = 6.98$ at the 1.0% level of significance
$F(1,75) = 11.79$ at the 0.1% level of significance

LT-V = Lorge-Thorndike Intelligence Tests, Verbal Battery
LT-NV = Lorge-Thorndike Intelligence Tests, Nonverbal Battery

TIP-I = Test of Inference Patterns, Part I
TIP-II = Test of Inference Patterns, Part II
TIP-T = Test of Inference Patterns, Total Test

The reliability of TIP for the 40 students in the experimental group and that for the same number in the control group were calculated by the split-half method and corrected by use of the Spearman-Brown Prophecy Formula. The obtained reliability coefficients have been tabulated in Table 2.

TABLE 2. SPLIT-HALF RELIABILITY COEFFICIENTS[a] FOR THE TEST OF INFERENCE PATTERNS.

Test of Inference Patterns	Experimental	Control
Part I	.45	.29
Part II	.75	.13
Total	.78	.11

[a] Coefficients have been attenuated by use of the Spearman-Brown Prophecy Formula.

Since reliability coefficients obtained by the split-half method were product-moment correlation coefficients, they could be tested for significance of difference from zero. Those for the control group were found to be not significantly different from zero.

Correlations have been calculated separately for the treatment groups on all pairs of the following scales: Lorge-Thorndike Intelligence Tests, Verbal Battery; Lorge-Thorndike Intelligence Tests, Nonverbal Battery; TIP - Part I; TIP - Part II; and TIP - Total Test. The correlation coefficients for the experimental group have been placed in Table 3, while those for the control group have been placed in Table 4.

All correlation coefficients for the experimental group were found to be significantly different from zero. However, for the control group the following correlation coefficients were found to be not significantly different from zero: those between the Lorge-Thorndike nonverbal battery and each of the three sets of scores on TIP, that between Parts I and II of TIP, and that between the Lorge-Thorndike verbal battery and Part II of TIP. Parts I and II of TIP were observed to be measuring different characteristics because at most 31% of the variance of one was accounted for by the other.

TABLE 3. MATRIX OF CORRELATION COEFFICIENTS[b] ON THE INDICATED VARIABLES FOR THE EXPERIMENTAL GROUP IN THE FIRST EXPERIMENT.

	LT-V	LT-NV	TIP-I	TIP-II	TIP-T
Lorge-Thorndike Verbal	1.00	.65	.54	.58	.64
Lorge-Thorndike Nonverbal		1.00	.54	.53	.60
Test of Inf. Pat. Part I			1.00	.55	.82
Test of Inf. Pat. Part II				1.00	.92
Test of Inf. Pat. Total					1.00

[b] For df = 38, $r \geq .312$ is required for significance from zero at the 5% level.

TABLE 4. MATRIX OF CORRELATION COEFFICIENTS[b] ON THE INDICATED VARIABLES FOR THE CONTROL GROUP IN THE FIRST EXPERIMENT.

	LT-V	LT-NV	TIP-I	TIP-II	TIP-T
Lorge-Thorndike Verbal	1.00	.59	.35	.26	.39
Lorge-Thorndike Nonverbal		1.00	.02	.29	.23
Test of Inf. Pat. Part I			1.00	.17	.70
Test of Inf. Pat. Part II				1.00	.82
Test of Inf. Pat. Total					1.00

[b] For df = 38, $r \geq .312$ is required for significance from zero at the 5% level.

The corresponding correlation coefficients for the treatment groups were tested for significance of difference for independent samples. The resulting z-scores have been placed in Table 5. At the 5% level of significance only two z-scores were found to be significant: that for the Lorge-Thorndike nonverbal battery and Part I of TIP, and that for the Lorge-Thorndike nonverbal battery and the total score on TIP.

TABLE 5. z-SCORES[c] FROM TESTING FOR A SIGNIFICANT DIFFERENCE BETWEEN CORRESPONDING CORRELATION COEFFICIENTS IN TABLE 3 AND TABLE 4.

	LT-V	LT-NV	TIP-I	TIP-II	TIP-T
Lorge-Thorndike Verbal	.00	.42	1.03	1.71	1.49
Lorge-Thorndike Nonverbal		.00	2.51	1.25	1.97
Test of Inf. Pat. Part I			.00	1.92	1.25
Test of Inf. Pat. Part II				.00	1.86
Test of Inf. Pat. Total					.00

[c] $z \geq 1.96$ is required for significance at the 5% level.

Since significant treatment differences were found in the analyses of variance and covariance, the Test of Inference Patterns was further examined to determine which patterns were understood by students of various groups and to determine which patterns contributed to making the results significant. A student was said to have understood a pattern if he responded correctly to five or six occurrences of that pattern. The proportion of such students in several categories have been compiled in Table 6. Students were placed in the high IQ category if the sum of their IQ's on the verbal and nonverbal batteries of the Lorge-Thorndike Intelligence Tests was greater than 220. The median IQ for the students in this first experiment was 110 on each battery.

TABLE 6. PROPORTION OF STUDENTS IN THE INDICATED CATEGORIES WHO RESPONDED CORRECTLY TO FIVE OR SIX EXAMPLES OF EACH INFERENCE PATTERN ON A TEST OF INFERENCE PATTERNS.

Categories	Patterns[d]									
	A	B	C	D	E	F	G	H	I	J
Experimental	.65	.60	.55	.10	.60	.68	.23	.50	.20	.00
Male	.65	.61	.44	.14	.65	.70	.30	.44	.22	.00
Female	.65	.59	.71	.06	.53	.65	.12	.59	.18	.00
High IQ	.80	.75	.70	.15	.75	.80	.30	.60	.25	.00
Low IQ	.50	.45	.40	.05	.45	.55	.15	.40	.15	.00
Control	.48	.78	.13	.05	.35	.70	.33	.35	.03	.00
Male	.40	.80	.05	.10	.25	.50	.40	.20	.00	.00
Female	.55	.75	.20	.00	.45	.90	.25	.50	.05	.00
High IQ	.53	.79	.11	.11	.53	.79	.26	.47	.00	.00
Low IQ	.43	.76	.14	.00	.19	.62	.38	.24	.05	.00

[d] Each of capital letters A through J refers to the inference pattern so designated in Appendix A.

Proportions between the following categories were found to be significant for the indicated patterns:

1) experimental over control on patterns C and E,
2) experimental male over control male on patterns C and E,
3) experimental female over control female on pattern C,
4) control female over control male on patterns F and H,
5) experimental high IQ over control high IQ on pattern C,
6) control low IQ over experimental low IQ on pattern B,
7) experimental high IQ over experimental low IQ on pattern A,
8) control high IQ over control low IQ on pattern E.

Pattern C (or-introduction) and pattern E (hypothetical syllogism) contributed the most to the significant difference between the experimental and control groups in the first experiment.

The F-ratios and estimated omegas-squared from the analyses of variance and covariance of the tests used in the second experiment have been tabulated in Table 7. The summary statistics, tests for homogeneity of variance, and summary tables for the analyses of variance and covariance which form the basis for Table 7 have been placed in Appendix E, Tables 17 through 21. Bartlett's test for homogeneity of variance [14] revealed a significant difference only among the four cell variances for the nonverbal battery of the Lorge-Thorndike Intelligence Tests. However, the calculated chi-square of 8.05 represented a slight excess over the tabled value of 7.8 at the 5% level of significance. This moderate departure from the assumption of equal variances has been claimed not to seriously affect the sampling distribution of the resulting F statistic [14].

Analyses of variance on the verbal and nonverbal batteries of the Lorge-Thorndike Intelligence Tests yielded no significant differences for the main effects of treatment or sex and for the interaction between treatment and sex.

The results of the analyses of variance and covariance on Inferential Reasoning Analysis and its subscales led to the rejection of the null hypotheses of no significant difference between treatment means in all cases except for the analysis of variance on Part I of IRA. Thus, the instructional unit on deductive reasoning was effective in producing highly significant differences on the inferential reasoning scale.

Significant F-ratios were obtained for the main effect of the sex factor in the following analyses of variance (ANOVA) and analyses of covariance (ANOCOVA):

1) ANOVA on the Part I scores of IRA,
2) ANOCOVA on the Part I scores of IRA with nonverbal IQ as covariate,
3) ANOVA on the Part I and II scores of IRA,
4) ANOCOVA on the Part I and II scores of IRA with nonverbal IQ as covariate,
5) ANOVA on the Part III scores of IRA,
6) ANOCOVA on the Part III scores of IRA with verbal IQ as covariate,
7) ANOCOVA on the Part III scores of IRA with nonverbal IQ as covariate,
8) ANOVA on IRA,
9) ANOCOVA on IRA with verbal IQ as covariate, and
10) ANOCOVA on IRA with nonverbal IQ as covariate.

There was no significant interaction between treatment and sex in the analyses of variance and covariance on Inferential Reasoning Analysis and its subscales.

TABLE 7. F-RATIOS AND ESTIMATED OMEGAS-SQUARED FROM THE ANALYSES OF VARIANCE AND OF COVARIANCE IN THE SECOND EXPERIMENT.

Variable		Source of Variation					
Covariate	Criterion	Treatment		Sex		Interaction	
		F	O^2	F	O^2	F	O^2
	LT-V	.00	.00	1.48	.01	.08	.00
	LT-NV	.31	.00	.04	.00	1.19	.00
	IRA-I	3.74	.03	4.69	.04	.11	.00
LT-V	IRA-I	6.32	.07	3.15	.02	.03	.00
LT-NV	IRA-I	8.12	.08	8.27	.08	.17	.00
	IRA-II	39.46	.33	2.26	.01	.19	.00
LT-V	IRA-II	49.50	.39	1.08	.00	.11	.00
LT-NV	IRA-II	58.15	.42	3.51	.02	.02	.00
	IRA-I+II	20.28	.19	4.74	.04	.26	.00
LT-V	IRA-I+II	36.66	.32	3.25	.02	.18	.00
LT-NV	IRA-I+II	42.81	.33	9.68	.07	.08	.00
	IRA-III	11.93	.11	8.37	.08	1.18	.00
LT-V	IRA-III	14.73	.14	6.66	.06	1.12	.00
LT-NV	IRA-III	22.71	.20	14.42	.12	.27	.00
	IRA-IV	12.25	.13	1.18	.00	.00	.00
LT-V	IRA-IV	12.74	.14	.67	.00	.00	.00
LT-NV	IRA-IV	13.01	.14	1.27	.00	.00	.00
	IRA-T	23.64	.21	8.10	.07	.69	.00
LT-V	IRA-T	38.86	.32	6.94	.04	.68	.00
LT-NV	IRA-T	55.05	.37	17.90	.12	.00	.00

$F(1,72) = 3.97$ at the 5.0% level of significance
$F(1,72) = 7.00$ at the 1.0% level of significance
$F(1,72) = 11.83$ at the 0.1% level of significance

LT-V = Lorge-Thorndike Intelligence Tests, Verbal Battery
LT-NV = Lorge-Thorndike Intelligence Tests, Nonverbal Battery

IRA-I = Inferential Reasoning Analysis, Part I
IRA-II = Inferential Reasoning Analysis, Part II
IRA-I+II = Inferential Reasoning Analysis, Parts I and II
IRA-III = Inferential Reasoning Analysis, Part III
IRA-IV = Inferential Reasoning Analysis, Part IV
IRA-T = Inferential Reasoning Analysis, Total Test

The degree of association between treatment and Inferential Reasoning Analysis was substantial. An omega-squared of .39 was obtained for the treatment with Part II of IRA under covariance adjustment for verbal IQ. Thus, the independent variable (treatment) has been estimated to account for 39% of the variance of the dependent variable (Part II of IRA) adjusted by the covariate (verbal IQ). As witnessed by the relative sizes of the omegas-squared in the treatment column, the stronger association between treatment and Part II accounted for much of the estimated indices on Parts I and II and on the total test.

The degree of association between sex and inferential reasoning was small. The indices of association were observed to be greater on Parts I or III than on Parts II or IV. There was no measurable degree of association for the interaction effect with inferential reasoning.

For the main effect of treatment all F-ratios and estimated omegas-squared increased when the various scales of IRA were adjusted for either verbal or nonverbal IQ. The increases were slightly larger when the nonverbal IQ scores were used as the covariate. For the main effect of sex the F-ratios and estimated omegas-squared from the analyses of variance were reduced under covariance adjustment for verbal IQ but increased for nonverbal IQ.

The reliability of Parts I and II of IRA was determined separately for the 40 students in the experimental group and the 37 students in the control group. The reliability coefficients calculated by the split-half method and attenuated by the Spearman-Brown Prophecy Formula have been placed in Table 8 on the next page. Since these reliability coefficients were obtained by product-moment correlation, they were tested for significance from zero. The reliability coefficient of .305 for the control group on Part II was slightly less than the table value of .325 required for significance from zero.

TABLE 8. SPLIT-HALF RELIABILITY COEFFICIENTS[a] FOR PARTS I AND II OF INFERENTIAL REASONING ANALYSIS.

Inferential Reasoning Analysis	Experimental	Control
Part I	.75	.60
Part II	.55	.305
Parts I + II	.76	.66

[a] Coefficients have been attenuated by use of the Spearman-Brown Prophecy Formula.

Pearson product-moment coefficients of correlation have been calculated separately for the treatment groups on all pairs of the following scales: Lorge-Thorndike Intelligence Tests, Verbal Battery; Lorge-Thorndike Intelligence Tests, Nonverbal Battery; IRA-Part I; IRA-Part II; IRA-Parts I and II; IRA-Part III; IRA-Part IV; and IRA-Total Test. The correlation coefficients for the experimental group have been placed in Table 9, while those for the control group have been placed in Table 10.

All correlation coefficients for the experimental group were found to be significantly different from zero at the preassigned level of 5%. With respect to the control group, however, when Part IV of Inferential Reasoning Analysis was one of the scales, the obtained coefficients of correlation were not significantly different from zero.

Inspection of the correlation coefficients for the pairwise comparisons of the four parts of IRA yielded a maximum coefficient of .60 between Part I and Part III for the experimental group. Each of the subscales was measuring a different characteristic of the group because at most 36% of the variance of any part was accounted for by the variance of another.

TABLE 9. MATRIX OF CORRELATION COEFFICIENTS[b] ON THE INDICATED VARIABLES FOR THE EXPERIMENTAL GROUP IN THE SECOND EXPERIMENT.

	LT-NV	IRA-I	IRA-II	IRA-I+II	IRA-III	IRA-IV	IRA-T
LT-V	.62	.67	.40	.65	.54	.45	.66
LT-NV		.50	.56	.61	.64	.50	.71
IRA-I			.47	.87	.60	.44	.81
IRA-II				.83	.41	.48	.69
IRA-I+II					.59	.53	.88
IRA-III						.56	.89
IRA-IV							.70

[b] For df = 38, $r \geq .312$ is required for significance from zero at the 5% level.

TABLE 10. MATRIX OF CORRELATION COEFFICIENTS[e] ON THE INDICATED VARIABLES FOR THE CONTROL GROUP IN THE SECOND EXPERIMENT.

	LT-NV	IRA-I	IRA-II	IRA-I+II	IRA-III	IRA-IV	IRA-T
LT-V	.71	.63	.56	.71	.37	-.05	.59
LT-NV		.65	.45	.67	.51	-.20	.63
IRA-I			.41	.89	.45	-.01	.74
IRA-II				.77	.38	.11	.66
IRA-I+II					.50	.04	.84
IRA-III						.18	.87
IRA-IV							.28

[e] For df = 35, $r \geq .325$ is required for significance from zero at the 5% level.

The corresponding correlation coefficients for the treatment groups were tested for significance of difference for independent samples. The resulting z-scores have been placed in Table 11. At the 5% level the null hypothesis of no significant difference was rejected for Part IV of Inferential Reasoning Analysis paired with the following: Lorge-Thorndike verbal battery, Lorge-Thorndike nonverbal battery, Part I of IRA, Parts I and II of IRA, and total test of IRA.

TABLE 11. z-SCORES[c] FROM TESTING FOR A SIGNIFICANT DIFFERENCE BETWEEN CORRESPONDING CORRELATION COEFFICIENTS IN TABLES 9 AND 10.

	LT-NV	IRA-I	IRA-II	IRA-I+II	IRA-III	IRA-IV	IRA-T
LT-V	-.68	.29	-.88	-.47	.91	2.25	.49
LT-NV		-.95	.63	-.43	.82	3.17	.62
IRA-I			.31	-.37	.88	2.03	.68
IRA-II				.71	.15	1.74	.23
IRA-I+II					.54	2.32	.65
IRA-III						1.90	.37
IRA-IV							2.44

[c] $z \geq 1.96$ is required for significance at the 5% level.

Since significant treatment differences were found in the analyses of variance and covariance, Inferential Reasoning Analysis was further examined to determine which patterns were understood by students of various groups and to determine which patterns contributed to making the results significant. A student was said to have understood a pattern if he responded correctly to six or seven or eight occurrences of that pattern. The proportions of such students in several categories have been compiled in Table 12. A student was placed in the high IQ category if the sum of his IQ's on the verbal and nonverbal batteries of the Lorge-Thorndike Intelligence Tests was greater that 220. The median IQ for the students in this second experiment was 110 on each battery.

TABLE 12. PROPORTION OF STUDENTS IN THE INDICATED CATEGORIES WHO RESPONDED CORRECTLY TO SIX, SEVEN OR EIGHT EXAMPLES OF EACH VALID INFERENCE PATTERN ON PARTS I AND II OF THE INFERENTIAL REASONING ANALYSIS.

Categories		Patterns[d]						
		A	B	C	E	F	G	H
Experimental	Total	.73	.83	.68	.75	.93	.25	.50
	Male	.60	.80	.60	.75	.95	.20	.40
	Female	.85	.85	.75	.75	.90	.30	.60
	High IQ	.90	.90	.90	.90	1.00	.42	.79
	Low IQ	.57	.76	.48	.62	.86	.10	.24
Control	Total	.76	.76	.19	.35	.70	.49	.30
	Male	.65	.65	.13	.23	.65	.50	.32
	Female	.93	.93	.27	.53	.80	.47	.27
	High IQ	.94	.83	.28	.50	.94	.72	.50
	Low IQ	.58	.68	.11	.21	.47	.26	.11

[d] Each of the capital letters A through H refers to the inference pattern so designated in Appendix A.

Proportions between the following categories were found to be significant for the indicated patterns used in Parts I and II of IRA:

1) experimental over control on patterns C,E, and F, but control over experimental on pattern G,
2) experimental males over control males on patterns C, E, and F, but control males over experimental males on pattern G,
3) experimental females over control females on patterns C and H,
4) experimental males over control females on pattern C, but control females over experimental males on pattern A.
5) experimental females over control males on patterns C and E,
6) control females over control males on patterns A, B, and E,
7) experimental high IQ over experimental low IQ on patterns A, C, G, and H,
8) experimental high IQ over control high IQ on patterns C and E,
9) experimental high IQ over control low IQ on patterns A, C, E, F, and H,
10) control high IQ over experimental low IQ on patterns A and G,
11) experimental low IQ over control low IQ on patterns C,E and F,
12) control high IQ over control low IQ on patterns A, F, G, and H.

DISCUSSION

The sampling procedure employed in this investigation was done without replacement and from a finite population of 157. This procedure did not affect the sample mean; but for a sample size of 40, the sample variance could have been changed by a factor of 1.06 and the standard error of the mean by .86. No adjustments were made in any statistical tests to reflect this possible source of error. The analysis of variance procedure, however, is sufficiently robust that moderate departures from its assumptions do not seriously affect the validity of inferences derived from the experimental data [4,5].

Another limitation on the experimental aspect of this project was the availability and use of only one person as the teacher of the logic course. In addition, the appointed teacher was a first-year teacher and had no special preparation in logic. He was given the textbook, Effective Thinking, and a suggested timetable of topics, and then expected to conduct the course in an effective manner. This lack of guidance was partly on purpose and partly beyond the control of the principal investigator. Some unmeasurable amount of credit for significant differences must be attributed to the professional attitude and competence of this teacher in leading his students to understand the content and to appreciate the importance of effective thinking.

An original intent of this project was to incorporate the elements of effective thinking learned in the logic course into the regular classroom setting. This would have been done had the plan of four class sections for all courses been feasible. Due to a larger enrollment than expected the administration decided to establish five sections for the regular courses and four for the exploratory program. The assignment of students to the regular course sections followed the same technique as that for the exploratory program except that groups of five were formed. Thus the logic material was learned in a separate setting with very few references made to situations in the regular content areas by the teacher of the logic course.

The experimental design did not permit direct comparison of the results of the first experiment with those of the second. However, such a comparison has been made in the remaining discussion subject to possible error from these threats to invalidity: history and maturation [1]. The main history threat was that students in the first experiment might have informed those in the second of the nature of the measuring instruments. This could be one explanation for the significant difference between sexes in the second experiment. Consultation with the teacher of the logic course revealed no unusual occurrences that might contribute to a history effect. Regarding maturation, Howell [6] determined that eighth-grade accelerated mathematics students who had received no formal instruction in inference patterns understood one more inference pattern (and-introduction) than did their seventh-grad counterparts. A period of eleven weeks lapsed from the administration of TIP to that of IRA.

Comparison of the control categories in Tables 6 and 12 showed a significant difference on one pattern only, viz., pattern A (and-introduction). This difference could have been caused by randomness, maturation, history, or the fact that IRA did not contain invalid patterns. Since the invalid patterns used in TIP involved conditions and not conjunctions, the last possibility was considered unplausible. The IRA experimental group gained a significant increase over thE TIP experimental group on two patterns: and-elimination (B) and rule of detachment (F). In addition to the previously suggested causes of difference, previous experience by the teacher might have aided the observed gain on both patterns. The absence of invalid patterns in IRA might have effected better performance on rule of detachment (F). In TIP both the rule of detachment (F) and its invalid analogue, converse rule (I), were presented to the students. This inclusion of pattern I in TIP without previous exposure to it in the logic class might have caused a reduction of confidence by the experimental group in responding correctly to pattern F.

Analysis of covariance, whether with verbal or nonverbal IQ as covariate, produced highly significant differences between the treatment groups for the Part II scores on both TIP and IRA and for the total scores on TIP and the Part I+II scores on IRA. However the degree of association between treatment and measuring instruments was at least twice as great for IRA as for TIP. Again, confrontation with previously unencountered and unlearned invalid inference patterns might have caused this difference along with the previously discussed threats. Significant treatment differences were obtained on Part I of IRA but not on Part I of TIP.

Examination of the correlations between Part I and Part II of TIP and between the corresponding parts of IRA revealed a maximum coefficient of .55. Therefore, these two parts measured some common characteristic which accounted for at most 31% of the variance of each, but they also measured different characteristics. No names have been attached to them at this time. A factor analytic study has been planned in the near future to determine the quantity and nature of the characteristics or factors operating in the tests.

The low reliability coefficients for TIP and the low proportion of student understanding on all but two inference patterns in TIP indicated that the students in the control group were quite inconsistent in their replies to the examples of the remaining eight patterns. The very low correlation between the Part I and the Part II scores of TIP revealed that less than 3% of the variance of Part I was attributable to Part II.

Investigation of the various mean scores reported in Table 17 evidenced higher means for the female students when compared with the male students within the same treatment category. These differences significantly favored the female students on Part III of IRA with either covariate and on Part I of IRA when nonverbal IQ was used as the covariate. The significant differences on these parts influenced the significance between sexes on the Parts I+II scores and the total test scores. On Part III of IRA the control male students scored a mean of only 12.59 compared with means of over 17 for the other three groups. Lack of attention to directions or disinterest might have caused such a low mean. The experimental male students at least had the benefit of some instruction in and practice on proof and could respond to the directions with more enthusiasm

Post-experimental interviews confirmed the conjecture that low comprehension of pattern G, elimination by cases, was due to a misunderstanding of the nature of the operator 'or'. This pattern had been taught with the exclusive meaning of 'or', i.e., P or Q means P, or Q, and not both, rather than with the inclusive meaning, i.e., P or Q means P, or Q, or both.

CONCLUSIONS, IMPLICATIONS, and RECOMMENDATIONS

The following conclusions were derived from the experimental results of this investigation:

1. Seventh-grade students were able to successfully learn and correctly recognize several valid inference patterns when provided with formal instruction on this topic. This result confirmed the conclusion reached by Corley, Hyram, and Maw that students of grades 6, 7, and 8 can learn elementary aspects of effective thinking.

2. Seventh-grade students who did not receive instruction in inference patterns were not sure of the meaning of the logical operator 'or' nor did they recognize a correct conclusion to a chain of related conditional statements.

3. IQ was definitely related to performance on inferential reasoning. The higher IQ students scored consistently better on both tests than those with lower IQ's.

4. Instruction in only valid inference patterns was not sufficient for the students to correctly identify invalid inferences.

5. No consistent conclusion seemed warranted concerning possible differences between the abilities of students of the opposite sexes to reason inferentially. However, the female students tended to perform better on one criterion's parts which contained only valid patterns in their statement form.

6. Some presently unidentified and distinct characteristics were operating on the statement form and the letter form of the inference patterns.

These findings suggest that educators charged with the responsibility for constructing curricula and improving instruction seriously reconsider the inclusion of explicit instruction in inferential reasoning at the junior high school level. Students without such instruction have exhibited not only serious deficiencies in valid reasoning but also ready acceptance of conclusions attained through invalid procedures. A unit on effective thinking can be taught separately as it was in this experiment or it can be presented within the framework of a selected content area or it can be integrated into all the experience areas of a junior high school curriculum.

This researcher recommends th a curriculum and instruction project be undertaken to prepare suitable materials and experimentally test an effort to integrate aspects of effective thinking into the entire junior high school program.

SUMMARY

This study was undertaken at the seventh-grade level to prepare and experimentally test a unit on effective thinking. The experimental focus was on inferential reasoning, a sub-unit within effective thinking. A previous study by Howell had indicated that accelerated mathematics students of the junior high school grades learned inference patterns at a rate of one per year without formal instruction. Studies by Corley, Hyram, and Maw reported that students of grades 6 through 10 were able to learn elements of effective thinking with instruction.

The goals of this project were threefold: (1) to prepare a textbook on effective thinking which contained a section devoted to inferential reasoning; (2) to establish an experimental situation capable of determining the ability of seventh-grade students to correctly recognize valid inference patterns; and (3) to select or prepare an instrument to measure inferential reasoning ability.

A text, called Effective Thinking, was prepared during the summer months of 1966 and contained chapters on observation, authority, statement errors, inductive reasoning, deductive reasoning, and faulty reasoning. The chapter on deductive reasoning was devoted mainly to valid inference patterns. Two examples of valid inference patterns are:

Or-introduction	Hypothetical Syllogism
P. ――― P or Q.	If P, then Q. If Q, then R. ――― If P, then R.

The letters P and Q represent any statement.

The experiment was conducted with 157 seventh-grade students enrolled at Hudson Junior High School in Hudson, Wisconsin during the first semester of the 1966-7 academic year. All students were required to take a nine-week course, called Logic, as part of the exploratory program. Four sections were formed by stratified random sampling; each section was randomly assigned a nine-week period for the logic course. One member of the faculty, an inexperienced teacher, was assigned to teach the course, using Effective Thinking as the text. The posttest-only control group design was considered the most appropriate experimental design. Two factors were chosen: treatment and sex. There were two levels of treatment, experimental and control; there were two of sex, male and female. In the absence of suitable published tests, Test of Inference Patterns (abbreviated TIP and used by Howell in a prior study) became the criterion measure at the end of the first nine-week period. TIP consisted of two parts, each containing three examples of each of the ten inference patterns. The examples of the second part were similar to those shown above, whereas in the first part each letter was replaced by a statement. The experimental data were examined statistically by means of the analysis of variance and then by the analysis of covariance, first with verbal IQ and then with nonverbal IQ as the covariate. Student IQ's were obtained at the start of the experiment from the Lorge-Thorndike Intelligence Tests, Level 4. Other *post hoc* tests were performed to determine the reliability of TIP, the correlations between all pairs of scales and subscales employed, and the proportion of students who showed understanding of each of the ten inference patterns in TIP.

Since TIP contained two invalid inference patterns along with eight valid ones, and since another section was available to take the logic course during the second nine-week period, a decision was made to repeat the experiment with the modification of a different criterion. The new measure tested knowledge of valid inference patterns first and then that of both valid and invalid patterns. This scale was prepared by Howell and called Inferential Reasoning Analysis (abbreviated IRA). Parts I and II of IRA were similar to those of TIP except that each part contained four examples of each of seven valid patterns used in TIP. Part III of IRA contained ten exercises requiring the correct application of two valid inference patterns to find the conclusion. Part IV consisted of five valid and five invalid inference patterns. Student scores on IRA at the end of the second period were analyzed in the same way as those on TIP.

Analysis of variance of the total test scores on TIP resulted in significant differences between the experimental and control groups. There was no significant difference between the scores of the male and female students nor was there significant interaction between treatment and sex. Under separate covariance adjustment for verbal or nonverbal IQ larger significant F-ratios were obtained for the treatment factor. The estimated omegas-squared, a measure of the association between treatment and criterion, also increased under the covariance adjustments. These increments attested to the fact that IQ had an effect on the criterion, TIP, under the experimental treatment of the logic course. The sex factor and the interaction remained relatively unaffected by covariance analysis.

A low reliability coefficient on TIP and comprehension proportions of less than .50 on eight of the ten patterns suggested that many of the students in the control group were guessing at several responses. Contrast of comprehension proportions between the treatment groups furnished statistically significant differences in favor of the experimental group on two patterns: or-introduction and hypothetical syllogism.

Analysis of variance on the valid-patterns-only parts of IRA (Parts I+II) and on the mixed part (Part IV) and on the total scores produced significantly higher scores for the experimental group over the control group. Under analysis of covariance, the F-ratios and omegas-squared increased appreciably for the Parts I+II and the total test scores but only slightly for the Part IV scores. Significant differences were found between the sexes on the analyses of variance and covariance with both covariates for the total IRA scores. But on Parts I+II similar differences were found on the analyses of variance and of covariance with nonverbal IQ only. No significant differences were found on Part IV.

Examination of the mean cell scores on Part III revealed that most of the sex difference was due to the unusually low performance by the control male students. No significant interaction between treatment and sex was found in the analyses of the IRA scores.

The reliability coefficient for the experimental and control groups on Parts I+II of IRA were .76 and .66, respectively. Comprehension proportions of over .50 were obtained on three patterns by both groups and on two more patterns by the experimental group. Significantly better comprehension performance was attained by the experimental group on three patterns (or-introduction, hypothetical syllogism, and rule of detachment), but the control group did better on one pattern (elimination by cases). This last was later determined to be caused by teacher instruction in an improper meaning of the connective 'or'.

The analyses of variance and of covariance on both TIP and IRA produced the greatest F-ratios and estimated omegas-squared in favor of the experimental group on the Part II scores. Parts I and II of both tests were determined to be sharing at most one-third of their variances and thus measuring much that was different.

The results of the experiment confirmed the findings of other researchers that students of grades 6 through 10 can learn elements of effective thinking. In particular, the seventh-grade students profited from instruction in inferential reasoning and the higher IQ students profited more from this instruction. Although some differences between the sexes were found on the parts of IRA which did not contain invalid patterns, no general conclusion can be derived for the sex factor. Instruction on valid inference patterns was not a sufficient condition for the students to recognize invalid patterns.

These conclusions imply the need for those educators who are responsible for curriculum and instruction to reexamine the role of inferential reasoning at the junior high school level. The control students of this experiment showed little comprehension of several elementary aspects of deductive thinking. But they can correct these misconceptions when properly instructed.

REFERENCES

1. Campbell, Donald T., and Stanley, Julian C. 'Experimental and Quasi-Experimental Designs for Research on Teaching,' Handbook of Research on Teaching (Ed. by N. L. Gage). Skokie, Ill.: Rand McNally and Co. 1963. p. 171-246.

2. Corley, Glyn J. An Experiment in Teaching Logical Thinking and Demonstrative Geometry in Grades Six Through Ten. Unpublished doctoral dissertation, George Peabody College for Teachers. 1959. 146p.

3. Ennis, Robert H. 'A Definition of Critical Thinking,' The Reading Teacher. XVII, May 1964. p. 599-612.

4. Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Co., Inc. 1959. 347p.

5. Hays, William L. Statistics For Psychologists. New York: Holt, Rinehart and Winston. 1963. 719p.

6. Howell, Edgar N. Recognition of Selected Inference Patterns by Secondary School Mathematics Students. Unpublished doctoral dissertation, University of Wisconsin. 1965. 157p.

7. Howell, Edgar N. Effect of a Classical Logic Course to Prepare Seventh-Grade Students in Inferential Reasoning. Unpublished report, Institutional Grant No. 0150-6-64, Wisconsin State University, River Falls. September 1966. 40p.

8. Howell, Edgar N.; Mellum, Elmer A.; and Schumacher, Noel. Effective Thinking. Unpublished interim report, Project No. 6-8110, Office of Education, U. S. Department of Health, Education, and Welfare. August 1966. 123p.

9. Hyram, George H. 'An Experiment in Developing Critical Thinking in Children,' Journal of Experimental Education. XXVI, December 1957. p. 125-132.

10. Lorge, Irving, and Thorndike, Robert L. The Lorge-Thorndike Intelligence Tests, Level 4, Form A, Verbal Battery. Boston: Houghton Mifflin Co. 1954. 8p.

11. Lorge, Irving, and Thorndike, Robert L. The Lorge-Thorndike Intelligence Tests, Level 4, Form A, Nonverbal Battery. Boston: Houghton Mifflin Co. 1954. 8p.

12. Lorge, Irving, and Thorndike, Robert L. Examiner's Manual for the Lorge-Thorndike Intelligence Tests. Boston: Houghton Mifflin Co. 1957. 32p.

13. Maw, Ethel W. An Experiment in Teaching Critical Thinking in the Intermediate Grades. Unpublished doctoral dissertation, University of Pennsylvania. 1959. 210p.

14. Winer, Ben J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Co., Inc. 1962 672p.

APPENDIX A

Valid Inference Patterns

A. And-Introduction

P.
Q.
―――――
P and Q.

B. And-Elimination

P and Q. / P.   P and Q. / Q.

C. Or-Introduction

P.
―――――
P or Q.

D. Not-Introduction

If P, then Q.
If P, then not Q.
―――――
Not P.

E. Hypothetical Syllogism

If P, then Q.
If Q, then R.
―――――
If P, then R.

F. Rule of Detachment
Modus Ponens

If P, then Q.
P.
―――――
Q.

G. Elimination by Cases
Disjunctive Syllogism

P or Q.
Not P.
―――――
Q.

P or Q.
Not Q.
―――――
P.

H. Rule of Contraposition
Modus Tollens

If P, then Q.
Not Q.
―――――
Not P.

Invalid Inference Patterns

I. Converse Pattern

If P, then Q.
Q.
―――――
P.

J. Inverse Pattern

If P, then Q.
Not P.
―――――
Not Q.

APPENDIX B

TEST OF INFERENCE PATTERNS

by

Edgar N. Howell, Ph.D.

Wisconsin State University
River Falls, Wisconsin

January 1966

Do NOT open this booklet until you are told to do so by your examiner.

PART I

Directions: On the next three pages you will find exercises like this:

61. Lemo is a city.
St. Paul is not a state.

Lemo is a city and St. Paul is not a state.

Your task is to determine whether or not the statement below the line, i.e., the statement 'Lemo is a city and St. Paul is not a state', can be concluded correctly given the two statements above the line. If you reason that the conclusion is correct, then circle the 'Yes' on the answer paper next to the number corresponding to the number of the question (In this case 61); if you think that the statement below the line does not correctly follow from the one(s) above the line, then circle the 'No' on the answer paper. Thus your answer for the above example would look like this: 61. Yes No or like this: 61. Yes No . If you wish to change your answer, simply put an X through your first answer and circle the other one like this: 61. Yes No .

Remember: You decide whether or not the conclusion is correct only on the basis of the statements above the line. All statements above the line are to be accepted as true, i.e., the statement 'Lemo is a city.' means that Lemo really is a city.

There are three laws of logic with which you ought to be familiar.

1. A statement cannot be both true and false at the same time.
2. A statement is either true or false and there is no other possibility.
3. The negation of a negative statement is an affirmative statement. For example, the statement 'Jane is not unhappy' is equivalent to the statement 'Jane is happy'.

Keep these rules in mind as you determine whether each of the conclusions is correct or incorrect. Take a moment to reread these directions.

Do NOT turn the next page until the examiner tells you to do so.

1. Wisconsin is a state.
St. Paul is a city.

St. Paul is a city and Wisconsin is not a state.

2. If the paper is not delivered, then I will not know the news.
The paper was delivered.

I know the latest news.

3. A square is a rectangle.

A circle is not a rectangle or a square is a rectangle.

4. If I use bait, then I will catch a fish.
If I use bait, then I will not catch a fish.

I use bait.

5. This pencil is broken and this book is new.

This book is not new.

6. I did not go to the library today.
If today is Tuesday, then I must go to the library.

Today is not Tuesday.

7. That dog's name is Barny.

The dog's name is Barny or tomorrow is Sunday.

8. If the flat tire is fixed, then I will ride my bike.
The flat is not fixed.

I am not riding my bike.

9. The sky is cloudy or the coat is green.
The sky is not cloudy.

The coat is not green.

10. If the temperature is over 85 degrees, then I will swim.
If the temperature is over 85 degrees, then I will not swim.

The temperature is over 85 degrees.

11. We went hunting.
If the car starts, then we will go hunting.

The car started.

12. If it does not rain, then we will go to the game.
We went to the game.

It did not rain.

13. That car is not a Ford.
That is a Ford or it is a Chevrolet.

That car is not a Chevrolet.

Turn to the next page NOW.

14. If John plays tennis, then he wears sneakers.
If class is over, then John plays tennis.

If class is over, then John wears sneakers.

15. If I miss the bus, then I will be late for school.
If I am late for school, then I will be punished.

If I miss the bus, then I will be punished.

16. Apples are fruit.
If legons are not vegetables, then apples are not fruit.

Legons are not vegetables.

17. That hat is green and this coat is red.

That hat is not green.

18. If it is raining, then I will go fishing.
If it is raining, then I will not go fishing.

It is not raining.

19. One bushel of apples is rotten and a second bushel is green.

The second bushel of apples is not green.

20. Helen is a sophomore.
Linda is not a junior.

Linda is not a junior and Helen is a sophomore.

21. There is no school on Saturday.
There is no school on Sunday.

There is school on Saturday and there is no school on Sunday.

22. Angle ABC is a right angle.
If angle ABC is a right angle, then its measure is 90 degrees.

The measure of angle ABE is not 90 degrees.

23. If I get an A in mathematics, then I will make the honor roll.
I got an A in mathematics.

I did not make the honor roll.

24. If I carry my umbrella, then It will not rain.
It is not raining.

I am carrying my umbrella.

25. Joe does not like pretzels.
Jim likes popcorn or Joe likes pretzels.

Jim likes popcorn.

Turn to the next page NOW.

26. The milk was delivered.
If the milk is not delivered, then we will not eat cereal.

We ate cereal.

27. If Fido chases Cleo, then there will be a cat-and-dog fight.
There is no cat-and-dog fight.

Fido is not chasing Cleo.

28. John Glenn is a famous astronaut.

Jim plays hockey or John Glenn is a famous astronaut.

29. The sun is shining.
If the sun is shining, then I can see my shadow.

I cannot see my shadow.

30. It it snows, then the roads must be cleared.
If the roads are cleared, then cars may be driven safely.

If it snows, then cars may not be driven safely.

Do NOT turn to the next page until the examiner tells you to do so.

PART II

Directions: We will now find it convenient to abbreviate simple sentences; otherwise some of the following statements would be very long and complicated. Up to now you considered inference patterns such as

It is raining.
The sun is not shining.
―――――――――――――――
It is raining and the sun is not shining.

Let us agree to replace the statement 'It is raining.' by the letter 'P' and the statement 'The sun is shining.' by the letter 'Q'. Then the above argument would look like this:

P.
Not Q.
―――――
P and not Q.

We place the period at the end of the line to remind you that 'P' stands for a sentence. In the remaining exercises, each of the letters P, Q, R, S, or T stands for any simple sentence.

NOTE: When we write 'not P and Q', does the word 'not' negate (i.e., nullify or deny) just the simple sentence 'P' or does the 'not' negate the compound sentence 'P and Q'? To avoid any confusion, we place parentheses around 'P and Q' when 'not' negates the compound sentence, i.e., we write 'not (P and Q)'. Hence, in 'not P and Q' the 'not' negates only 'P'.

Remember: Each statement above the line is to be accepted as true. You answer the question 'Does the conclusion correctly follow?' Mark the answer the same way as before. Take a moment to reread these directions.

Do NOT turn to the next page until
the examiner tells you to do so.

31. Not P.
If P, then Q.
Not Q.

32. R.
If P and Q, then R.
Not (P and notQ).

33. If not P, then not Q.
P.
Q.

34. If P, then Q or R.
Not (Q or R).
Not P.

35. If P, then Q.
Not Q.
Not P.

36. Not Q.
Not P.
Q and not P.

37. If not P, then not Q.
If not P, then Q.
Not P.

38. Not Q.
P or Q.
P.

39. Not Q.
P.
P and not Q.

40. Not P.
If P, then Q or R.
Q or R.

41. Not P.
If not P, then Q.
Not Q.

42. Not P or not Q.
P.
Q.

43. If not Q, then not R.
If P, then not Q.
If P, then not R.

44. P or not Q.
R.
(P or not Q)and R.

45. If P, then Q.
Q.
P.

46. If P, then Q.
P.
Q.

47. P and Q.
Not P.

48. If R and S, then T.
If P or Q, then R and S.
If P or Q, then not T.

49. If P or Q, then not R.
If P or Q, then R.
P or Q.

50. If P, then Q.
If Q, then R.
If P, then not R.

Turn to the next page NOW.

51. Not P and Q.

P.

52. P or Q.

P or Q or R.

53. If P and not Q, then R.
P and not Q.

Not R.

54. Not Q.
If P, then not Q.

Not P.

55. If P, then not Q.
If P, then Q.

Not P.

56. P or not Q.
Not P.

Not Q.

57. Not P.

P or not Q.

58. Not P and not Q.

Not Q.

59. Q.

P or Q.

60. Not Q.
If not P, then Q.

Not P.

ANSWER SHEET FOR TIP

Name ______________________ Sex ______________

School ______________________ Grade ______________

Homeroom ______________________

Directions: For each inference pattern circle either 'Yes' or 'No'

Part I			Part II			Part III		
1.	Yes	No	21.	Yes	No	41.	Yes	No
2.	Yes	No	22.	Yes	No	42.	Yes	No
3.	Yes	No	23.	Yes	No	43.	Yes	No
4.	Yes	No	24.	Yes	No	44.	Yes	No
5.	Yes	No	25.	Yes	No	45.	Yes	No
6.	Yes	No	26.	Yes	No	46.	Yes	No
7.	Yes	No	27.	Yes	No	47.	Yes	No
8.	Yes	No	28.	Yes	No	48.	Yes	No
9.	Yes	No	29.	Yes	No	49.	Yes	No
10.	Yes	No	30.	Yes	No	50.	Yes	No
11.	Yes	No	31.	Yes	No	51.	Yes	No
12.	Yes	No	32.	Yes	No	52.	Yes	No
13.	Yes	No	33.	Yes	No	53.	Yes	No
14.	Yes	No	34.	Yes	No	54.	Yes	No
15.	Yes	No	35.	Yes	No	55.	Yes	No
16.	Yes	No	36.	Yes	No	56.	Yes	No
17.	Yes	No	37.	Yes	No	57.	Yes	No
18.	Yes	No	38.	Yes	No	58.	Yes	No
19.	Yes	No	39.	Yes	No	59.	Yes	No
20.	Yes	No	40.	Yes	No	60.	Yes	No

APPENDIX C

INFERENTIAL REASONING ANALYSIS

by

Edgar N. Howell, Ph.D.

Temple University
Philadelphia, Pa.

January 1967

Do NOT open this booklet until you are told to do so by your examiner.

Part I

Directions: On the next three pages you will find exercises like this: Example A.

Lemo is a city and Anfer is a nation.
Anfer is a nation and Lemo is a city.

Your task is to determine whether or not the statement below the line (called the conclusion) can be correctly concluded from the statement above the line (called the premiss). If you reason that the conclusion is correct, then circle the 'Yes' on the answer sheet next to the number corresponding to the number of the question (in this case A). If you think that the conclusion does not correctly follow from the premiss, then circle the 'No' on your answer sheet. Circle one of these now. For Example A you should have circled 'Yes' because the conclusion does follow from the premiss. Whenever you wish to change an answer, simply put an X through your first answer and then circle the other one.

Example B.

Lemo is a city and Anfer is a nation.
Anfer is not a nation and Lemo is a city.

Circle an answer next to 'B' on your answer sheet. The correct answer to this example is 'No' because the conclusion does not follow from the premiss.

Remember: You decide whether or not the conclusion is correct only on the basis of the statements above the line, viz., the premisses. All premisses are to be accepted as true. The premiss 'Lemo is a city' means that Lemo really is a city.

Here are three laws of logic with which you ought to be familiar:

1. A statement cannot be both true and false at the same time.
2. A statement is either true or false and there is no other possibility.
3. The negation of a negative statement is an affirmative statement. For example, the statement 'Jane is not unhappy' is the same as the statement 'Jane is happy'.

Keep these rules in mind as you determine whether each of the conclusions is correct. Take a minute to reread these directions.

In each of the following exercises you are to answer the question: Does the conclusion follow correctly from the premiss(es)?

Do NOT turn to the next page until the examiner tells you to do so.

1. Wisconsin is a state.
St. Paul is a city.

St. Paul is a city and Wisconsin is not a state.

2. The TV was on and the volume was too loud.

The TV was on.

3. Jean did not go to the library today.
If today is Tuesday, then Jean goes to the library.

Today is not Tuesday.

4. If ice melts, then water will be formed.
If the temperature is 40 degrees, then ice melts.

If the temperature is 40 degrees, then water will not be formed.

5. The newspaper was delivered.
The paperboy did not collect.

The newspaper was delivered and the paperboy did not collect.

6. Lincoln was a U. S. president.

Lee was not a senator or Lincoln was a U. S. president.

7. The pencil was broken and the pen was out of ink.

The pen was not out of ink.

8. The sky is cloudy or the grass is burned.
The sky is not cloudy.

The grass is burned.

9. If I carry my umbrella, then it will not rain.
I am carrying my umbrella.

It is not raining.

10. If John plays tennis, then he must wear sneakers.
If class is over, then John plays tennis.

If class is over, then John must wear sneakers.

Turn to the next page **NOW**.

11. That hat is green and this one is red.

    That hat is not green.

12. If Fido chases Cleo, then there will be a cat-and-dog fight.
    There is no cat-and-dog fight.

    Fido is not chasing Cleo.

13. This book is brand new or it has been used.
    This book has not been used.

    This book is not brand new.

14. There is no school on Saturday.
    There is no school on Sunday.

    There is school on Saturday and there is no school on Sunday.

15. Helen is a senior.
    Linda is not a junior.

    Linda is not a junior and Helen is a senior.

16. If the flat tire is fixed, then Paul will ride his bike.
    Paul is not riding his bike.

    The flat tire was fixed.

17. Marie won the dance contest.

    Marie did not win the dance contest or Bill danced with her.

18. This car is not a Ford.
    This car is a Ford or it is a Buick.

    This car is not a Buick.

19. If I get an A in mathematics, then I will make the honor roll.
    I got an A in mathematics.

    I did not make the honor roll.

20. That dog's name is Barney.

    That dog's name is Barney or Henry's car is a Chevy.

Turn to the next page NOW.

21. The sun is shining.
<u>If the sun is shining, then shadows can be seen.</u>

Shadows cannot be seen.

22. Most apples in the first basket are green and most in the second are rotten.

Most apples in the second basket are rotten.

23. If it snows, then the roads must be cleared.
If the roads are cleared, then cars may be driven safely.

If it snows, then cars may not be driven safely.

24. A square is a rectangle.

A circle is not a square or a square is not a rectangle.

25. Julie did not go swimming.
If the temperature is over 85 degrees, then Julie goes swimming.

The temperature was over 85 degrees.

26. If Mary misses the bus, then she will be late for school.
If Mary is late for school, then she will be punished.

If Mary misses the bus, then she will be punished.

27. Joe does not like pretzels.
Jim likes popcorn or Joe likes pretzels.

Jim likes popcorn.

28. $2 \times 3 = 6$.
If $2 \times 3 = 6$, then $6/2 = 3$.

$6/2 = 3$.

Do NOT turn to the next page until the examiner tells you to do so.

Part II

Directions: We will now find it convenient to abbreviate simple sentences. In Part I you examined inference patterns such as

It is raining.
The sun is not shining.
It is raining and the sun is not shining.

In Part II let us agree to replace the statement 'It is raining' by the letter 'P' and the statement 'The sun is shining' by the letter 'Q'. Then the above inference pattern would look like this:

P.
Not Q.
P and not Q.

We place the period at the end of each line to remind you that 'P' stands for a sentence. In the remaining exercises, each of the letters P, Q, or R stands for any simple sentence.

Remember: Each statement above the line is to be accepted as true. You must answer the question:

Does the conclusion correctly follow?

Mark the answer sheet in the same way as in Part I. Take a minute to reread these directions.

Do NOT turn to the next page until the examiner tells you to do so

29. P or Q.
Not Q.

Not P.

30. P.
If P, then Q.

Not Q.

31. P.
Not P or Q.

Not Q.

32. If Q, then R.
If P, then Q.

If P, then not R.

33. Not P.
Not Q.

Not Q and P.

34. Not P or not Q.
P.

Not Q.

35. Not Q.
If P, then Q.

Not P.

36. If Q, then R.
If not P, then Q.

If not P, then R.

37. If not P, then Q.
Not Q.

P.

38. Not P and Q.

Q.

39. Not P and not Q.

Not P.

40. Not Q.
P.

P and not Q.

41. P.
Not Q.

P and Q.

42. Not P.

Not P or not Q.

43. Not P.

Not Q or P.

44. Not Q.

Q or P.

45. If P, then not Q.
Q.

P.

46. P and not Q.

Q.

47. If P, then not Q.
If not Q, then not R.

If P, then R.

48. Not P.
If not P, then not Q.

Not Q.

49. If not P, then Q.
Not P.

Q.

Turn to the next page NOW.

50. P.

Q or P.

51. Not P.
P or not Q.

Not Q.

52. Q.
If not P, then not Q.

Not P.

53. P and Q.

Not P.

54. P.
Q.

P and Q.

55. If P, then not Q.
P.

Q.

56. If not P, then Q.
If Q, then not R.

If not P, then not R.

Do NOT turn to the next page until the examiner tells you to do so.

Part III

Directions: On the next two pages you will find ten exercises like the following example:

	Pattern	Response Set
Example C.	P or Q.	(1) Not S.
	Not Q and S.	(2) P and S.
		(3) Not P.
		(4) P.
		(5) Not given.

Each exercise contains exactly two correct answers in its response set. One correct answer uses more information from the premisses of the pattern than the other. The former receives 2 points, the latter 1 point; incorrect answers receive 0 points. On the Part III section of the answer sheet, record what you think are the two correct answers for Example C. Do this by placing an x in two of the five sets of parentheses.

You should have marked your answer sheet in this way:

	1	2	3	4	5
Example C.	( )	(x)	( )	(x)	( )

Responses 2 and 4 were correct with response 2 using more information from the premisses.

One of the two correct responses can always be found among the first four items in each response set. Sometimes, however, a correct second response does not occur among the first four. Then you should select the fifth response 'Not given' for your second correct answer.

Exercises having more than two responses marked will receive no credit. If you think you have chosen an incorrect answer, fill in the space between the parentheses with ink or lead and then mark your other choice clearly with an x.

Remember: The statements above the line (the premisses) are to be accepted as true. You are to answer the question:

Which conclusions correctly follow?

Take a minute to reread these directions.

Do NOT turn to the next page until the examiner tells you to do so.

Pattern	Response Set
57. Not Q. P or Q. <u>R.</u>	(1) Not R. (2) P and R. (3) Q and P. (4) P. (5) Not given.
58. If P, then Q. P. <u>If Q, then R.</u>	(1) Not R. (2) Q. (3) R. (4) If Q, then P. (5) Not given.
59. If P, then Q. If Q, then R. <u>If R, then S.</u>	(1) S. (2) If P, then S. (3) If Q, then S. (4) If R, then Q. (5) Not given.
60. Not Q and R. <u>P or Q.</u>	(1) P. (2) P and Q. (3) Q. (4) Not Q. (5) Not given.
61. If R, then not P. If not P, then not Q. <u>Q.</u>	(1) Not Q. (2) R. (3) Not R. (4) P. (5) Not given.
62. If P and Q, then not R. P. <u>Q.</u>	(1) Not Q or R. (2) P and Q. (3) R. (4) Not R. (5) Not given.

Turn to the next page NOW.

Pattern	Response Set
63. P or Q. If Q, then not R. <u>Not P.</u>	(1) Q. (2) Not Q. (3) P. (4) R. (5) Not given.
64. R and not Q. <u>If P, then Q.</u>	(1) R. (2) Not P. (3) P. (4) Q. (5) Not given.
65. If P, then Q. Not Q. <u>If not P, then R.</u>	(1) Q or P. (2) Not R. (3) R. (4) Not P. (5) Not given.
66. P and Q. <u>If P, then R.</u>	(1) R. (2) Q. (3) Not R. (4) Not P. (5) Not given.

Do NOT turn to the next page until the examiner tells you to do so.

Part IV

Directions: In the previous three parts of this test the premisses were selected in such a way that their truth implied a true conclusion. Such an inference pattern is called valid. If an inference pattern is not valid, then it is called invalid.

On the next page there are ten exercises involving inference patterns. Some are valid, others are invalid. Invalid patterns cannot have a correct conclusion. Valid patterns do have a correct conclusion, but the one given may not be correct. Therefore, there are three possible answers to each exercise:

1. Valid-Yes
2. Valid-No
3. Invalid.

'Valid-Yes' means that the pattern is valid and that the given conclusion is the correct one. 'Valid-No' means that the pattern is valid but the given conclusion is not correct. You are to determine which of the three choices correctly describes each exercise. Mark your answer with an x in the appropriate set of parentheses.

Take a minute to reread these directions.

Do NOT turn to the next page until the examiner tells you to do so.

67. P.
———
P or Q.

68. If P, then Q.
P.
———
Q.

69. P or Q.
———
P.

70. If P, then Q.
Q.
———
P.

71. If P, then Q.
Not Q.
———
Not P.

72. P and Q.
———
Q.

73. P.
———
P and Q.

74. If P, then Q.
Q.
———
P.

75. If P, then Q.
If R, then S.
———
If P, then S.

76. If P, then Q.
If Q, then R.
———
If R. then P.

Answer Sheet for INFERENTIAL REASONING ANALYSIS

Name ______________________________ Sex __________

Date ____________________ Quarter taking logic 1 2 3 4

A.	Yes	No	B.	Yes	No
		Part I			Part II
1.	Yes		29.	Yes	No
2.	Yes	No	30.	Yes	No
3.	Yes	No	31.	Yes	No
4.	Yes	No	32.	Yes	No
5.	Yes	No	33.	Yes	No
6.	Yes	No	34.	Yes	No
7.	Yes	No	35.	Yes	No
8.	Yes	No	36.	Yes	No
9.	Yes	No	37.	Yes	No
10.	Yes	No	38.	Yes	No
11.	Yes	No	39.	Yes	No
12.	Yes	No	40.	Yes	No
13.	Yes	No	41.	Yes	No
14.	Yes	No	42.	Yes	No
15.	Yes	No	43.	Yes	No
16.	Yes	No	44.	Yes	No
17.	Yes	No	45.	Yes	No
18.	Yes	No	46.	Yes	No
19.	Yes	No	47.	Yes	No
20.	Yes	No	48.	Yes	No
21.	Yes	No	49.	Yes	No
22.	Yes	No	50.	Yes	No
23.	Yes	No	51.	Yes	No
24.	Yes	No	52.	Yes	No
25.	Yes	No	53.	Yes	No
26.	Yes	No	54.	Yes	No
27.	Yes	No	55.	Yes	No
28.	Yes	No	56.	Yes	No

Name ____________________

	1	2	3	4	5
Example C.	( )	( )	( )	( )	( )

Part III

	1	2	3	4	5
57.	( )	( )	( )	( )	( )
58.	( )	( )	( )	( )	( )
59.	( )	( )	( )	( )	( )
60.	( )	( )	( )	( )	( )
61.	( )	( )	( )	( )	( )
62.	( )	( )	( )	( )	( )
63.	( )	( )	( )	( )	( )
64.	( )	( )	( )	( )	( )
65.	( )	( )	( )	( )	( )
66.	( )	( )	( )	( )	( )

Part IV

	Valid-Yes	Valid-No	Invalid
67.	( )	( )	( )
68.	( )	( )	( )
69.	( )	( )	( )
70.	( )	( )	( )
71.	( )	( )	( )
72.	( )	( )	( )
73.	( )	( )	( )
74.	( )	( )	( )
75.	( )	( )	( )
76.	( )	( )	( )

APPENDIX D

TABLE 13. SUMMARY STATISTICS FOR THE TESTS USED IN THE FIRST EXPERIMENT.

Test	Statistic	Experimental Male	Experimental Female	Control Male	Control Female
Lorge-Thorndike Verbal	Number	23	17	20	20
	Mean	111.21	104.58	107.50	112.55
	Variance	147.81	138.38	192.47	148.15
Lorge-Thorndike Nonverbal	Number	23	17	20	20
	Mean	111.08	107.23	108.25	110.30
	Variance	190.62	270.56	142.19	113.16
Test of Inf. Pat. Part I	Number	23	17	20	20
	Mean	18.69	18.29	17.50	17.80
	Variance	7.22	10.72	7.73	4.48
Test of Inf. Pat. Part II	Number	23	17	20	20
	Mean	20.34	20.29	17.10	17.85
	Variance	21.05	17.59	6.51	12.13
Test of Inf. Pat. Total	Number	23	17	20	20
	Mean	39.04	38.58	34.60	35.65
	Variance	43.95	41.00	14.88	21.18

TABLE 14. BARTLETT'S TEST FOR THE HOMOGENIETY OF THE CELL VARIANCES ON THE TESTS USED IN THE FIRST EXPERIMENT.

Test	Chi-Square[f]
Lorge-Thorndike Intelligence Tests	
Verbal Battery	.60
Nonverbal Battery	3.71
Test of Inference Patterns	
Part I	3.23
Part II	6.92
Total	7.30

[f] For df = 3 a chi-square > 7.8 is required for significance at the 5% level.

TABLE 15. SUMMARY TABLES FOR THE ANALYSES OF VARIANCE ON THE TESTS USED IN THE FIRST EXPERIMENT.

Source of Variation	Sum of Squares	df	Mean Square	F
Lorge-Thorndike Verbal				
Treatment	89.03	1	89.03	.56
Sex	12.31	1	12.31	.07
Interaction	674.24	1	674.24	4.29
Within cells	11937.99	76	157.07	
Lorge-Thorndike Nonverbal				
Treatment	.27	1	.27	.00
Sex	16.05	1	16.05	.09
Interaction	172.13	1	172.13	.97
Within cells	13374.84	76	175.98	
Test of Inf. Pat. - Part I				
Treatment	14.11	1	14.11	1.90
Sex	.05	1	2.39	.00
Interaction	2.43	1	2.43	.32
Within cells	562.59	76	7.40	
Test of Inf. Pat. - Part II				
Treatment	160.14	1	160.14	11.07
Sex	2.39	1	2.39	.16
Interaction	3.19	1	3.19	.22
Within cells	1099.09	76	14.46	
Test of Inf. Pat. - Total				
Treatment	269.35	1	269.35	8.86
Sex	1.74	1	1.74	.05
Interaction	11.19	1	11.19	.36
Within cells	2308.42	76	30.37	

F(1,76) = 3.97 at the 5% level of significance
F(1,76) = 6.98 at the 1% level of significance

TABLE 16. SUMMARY TABLES FOR THE ANALYSES OF COVARIANCE OF THE FIRST EXPERIMENT.

Source of Variation	Sum of Squares	df	Mean Square	F
Cov.: LT-V, Crit.: TIP-I				
Treatment	21.76	1	21.76	3.64
Sex	.01	1	.01	.00
Interaction	.92	1	.92	.15
Within cells	447.77	75	5.97	
Cov.: LT-V, Crit.: TIP-II				
Treatment	192.32	1	192.32	16.30
Sex	4.07	1	4.07	.34
Interaction	2.71	1	2.71	.22
Within cells	884.66	75	11.79	
Cov.: LT-V, Crit.: TIP-T				
Treatment	343.47	1	343.47	15.46
Sex	4.56	1	4.56	.20
Interaction	6.79	1	6.79	.30
Within cells	1665.34	75	22.20	
Cov.: LT-NV, Crit.: TIP-I				
Treatment	14.38	1	14.38	2.17
Sex	.00	1	.00	.00
Interaction	.39	1	.39	.06
Within cells	496.23	75	6.61	
Cov.: LT-NV, Crit.: TIP-II				
Treatment	161.80	1	161.80	13.86
Sex	4.26	1	4.26	.36
Interaction	.00	1	.00	.00
Within cells	875.44	75	11.67	
Cov.: LT-NV, Crit.: TIP-T				
Treatment	272.67	1	272.67	11.52
Sex	4.50	1	4.50	.19
Interaction	.51	1	.51	.02
Within cells	1774.75	75	23.66	

$F(1,75) = 3.97$ at the 5% level of significance
$F(1,75) = 6.98$ at the 1% level of significance

APPENDIX E

TABLE 17. SUMMARY STATISTICS FOR THE TESTS USED IN THE SECOND EXPERIMENT.

Test	Statistic	Experimental Male	Experimental Female	Control Male	Control Female
Lorge-Thorndike Verbal	Number	20	20	22	15
	Mean	108.70	111.45	107.81	112.33
	Variance	209.48	175.52	152.06	121.09
Lorge-Thorndike Nonverbal	Number	20	20	22	15
	Mean	112.35	108.25	110.68	113.46
	Variance	204.23	116.40	302.98	84.55
Inf. Reas. Anal. Part I	Number	20	20	22	15
	Mean	21.75	23.20	19.95	21.93
	Variance	11.88	9.64	14.61	10.35
Inf. Reas. Anal. Part II	Number	20	20	22	15
	Mean	20.65	21.30	16.54	17.73
	Variance	8.23	7.16	5.49	7.49
Inf. Reas. Anal. Parts I + II	Number	20	20	22	15
	Mean	42.40	44.35	36.50	39.66
	Variance	29.41	22.66	24.64	28.09
Inf. Reas. Anal. Part III	Number	20	20	22	15
	Mean	18.00	20.15	12.59	17.33
	Variance	39.26	17.71	21.68	29.52
Inf. Reas. Anal. Part IV	Number	20	20	22	15
	Mean	4.45	4.80	3.18	3.60
	Variance	2.99	2.16	2.15	1.97
Inf. Reas. Anal. Total	Number	20	20	22	15
	Mean	64.85	69.30	52.27	60.40
	Variance	135.39	78.32	62.96	95.11

TABLE 18. BARTLETT'S TEST FOR THE HOMOGENIETY OF THE CELL VARIANCES ON THE TESTS USED IN THE SECOND EXPERIMENT.

Test	Chi-Square[f]
Lorge-Thorndike Intelligence Tests	
Verbal Battery	1.26
Nonverbal Battery	8.05
Inferential Reasoning Analysis	
Part I	.97
Part II	.85
Part I + II	.38
Part III	3.43
Part IV	.92
Total	3.12

[f] For df = 3 a chi-square > 7.8 is required for significance at the 5% level.

TABLE 19. SUMMARY TABLES FOR THE ANALYSES OF VARIANCE ON THE TESTS USED IN THE SECOND EXPERIMENT.

Source of Variation	Sum of Squares	df	Mean Square	F
Lorge-Thorndike Verbal				
Treatment	.03	1	.03	.00
Sex	248.87	1	248.87	1.48
Interaction	14.63	1	14.63	.08
Within cells	12203.77	73	167.17	
Lorge-Thorndike Nonverbal				
Treatment	59.36	1	59.36	.31
Sex	8.14	1	8.14	.04
Interaction	223.43	1	223.43	1.19
Within cells	13638.82	73	186.83	
Inf. Reas. Anal. - Part I				
Treatment	44.20	1	44.20	3.74
Sex	55.42	1	55.42	4.69
Interaction	1.31	1	1.31	.11
Within cells	860.83	73	11.79	
Inf. Reas. Anal. - Part II				
Treatment	277.42	1	277.42	39.46
Sex	15.92	1	15.92	2.26
Interaction	1.36	1	1.36	.19
Within cells	513.13	73	7.02	
Inf. Reas. Anal. - Parts I + II				
Treatment	528.03	1	528.03	20.28
Sex	123.42	1	123.42	4.74
Interaction	6.97	1	6.97	.26
Within cells	1900.18	73	26.02	
Inf. Reas. Anal. - Part III				
Treatment	318.98	1	318.98	11.93
Sex	223.95	1	223.95	8.37
Interaction	31.68	1	31.68	1.18
Within cells	1951.20	73	26.72	

TABLE 19 (Continued)

Source of Variation	Sum of Squares	df	Mean Square	F
Inf. Reas. Anal. - Part IV				
Treatment	28.71	1	28.71	12.25
Sex	2.78	1	2.78	1.18
Interaction	.02	1	.02	.00
Within cells	171.02	73	2.34	
Inf. Reas. Anal. - Total				
Treatment	2174.56	1	2174.56	23.64
Sex	745.74	1	745.74	8.10
Interaction	63.77	1	63.77	.69
Within cells	6714.71	73	91.98	

$F(1,73) = 3.97$ at the 5% level of significance
$F(1,73) = 7.00$ at the 1% level of significance

TABLE 20. SUMMARY TABLES FOR THE ANALYSES OF COVARIANCE ON INFERENTIAL REASONING ANALYSIS AS CRITERION AND LORGE-THORNDIKE INTELLIGENCE TEST, VERBAL BATTERY AS COVARIATE.

Source of Variation	Sum of Squares	df	Mean Square	F
Cov.: LT-V, Crit.: IRA-I				
Treatment	44.21	1	44.21	6.32
Sex	22.07	1	22.07	3.15
Interaction	.23	1	.23	.03
Within cells	503.64	72	6.99	
Cov.: LT-V, Crit.: IRA-II				
Treatment	277.43	1	277.43	49.50
Sex	6.10	1	6.10	1.08
Interaction	.64	1	.64	.11
Within cells	403.50	72	5.60	
Cov.: LT-V, Crit.: IRA-I+II				
Treatment	528.07	1	528.07	36.66
Sex	46.85	1	46.85	3.25
Interaction	2.62	1	2.62	.18
Within cells	1036.84	72	14.40	
Cov.: LT-V, Crit.: IRA-III				
Treatment	319.00	1	319.00	14.73
Sex	144.38	1	144.38	6.66
Interaction	24.38	1	24.38	1.12
Within cells	1559.15	72	21.65	
Cov.: LT-V, Crit.: IRA-IV				
Treatment	28.72	1	28.72	12.74
Sex	1.51	1	1.51	.67
Interaction	.00	1	.00	.00
Within cells	162.24	72	2.25	
Cov.: LT-V, Crit.: IRA-T				
Treatment	2174.70	1	2174.70	38.86
Sex	388.38	1	388.38	6.94
Interaction	38.23	1	38.23	.68
Within cells	4028.54	72	55.95	

$F(1,72) = 3.97$ at the 5% level of significance
$F(1,72) = 7.00$ at the 1% level of significance

TABLE 21. SUMMARY TABLES FOR THE ANALYSES OF COVARIANCE ON INFERENTIAL REASONING ANALYSIS AS CRITERION AND LORGE-THORNDIKE INTELLIGENCE TEST, NONVERBAL BATTERY AS COVARIATE.

Source of Variation	Sum of Squares	df	Mean Square	F
Cov.: LT-NV, Crit.: IRA-I				
Treatment	61.03	1	61.03	8.12
Sex	62.09	1	62.09	8.27
Interaction	1.27	1	1.27	.17
Within cells	540.59	72	7.50	
Cov.: LT-NV, Crit.: IRA-II				
Treatment	302.57	1	302.57	58.15
Sex	18.29	1	18.29	3.51
Interaction	.11	1	.11	.02
Within cells	374.59	72	5.20	
Cov.: LT-NV, Crit.: IRA-I+II				
Treatment	618.04	1	618.04	42.81
Sex	139.79	1	139.79	9.68
Interaction	1.21	1	1.21	.08
Within cells	1039.44	72	14.43	
Cov.: LT-NV, Crit.: IRA-III				
Treatment	384.30	1	384.30	22.71
Sex	244.05	1	244.05	14.42
Interaction	4.60	1	4.60	.27
Within cells	1218.10	72	16.91	
Cov.: LT-NV, Crit.: IRA-IV				
Treatment	30.10	1	30.10	13.01
Sex	2.95	1	2.95	1.27
Interaction	.01	1	.01	.00
Within cells	166.55	72	2.31	
Cov.: LT-NV, Crit.: IRA-T				
Treatment	2536.90	1	2536.90	55.05
Sex	825.08	1	825.08	17.90
Interaction	.29	1	.29	.00
Within cells	3317.74	72	46.07	

$F(1,72) = 3.97$ at the 5% level of significance
$F(1,72) = 7.00$ at the 1% level of significance