new hampshire college

GRADUATE SCHOOL OF BUSINESS

TRANSDUCTIVE REASONING AND THE TEACHING OF CONDITIONAL LOGIC TO CHILDREN

DR. ROBERT H. SEIDMAN

RESEARCH REPORT 86-2
JANUARY 1986

Occasional Paper Series

Transductive Reasoning and The Teaching of Conditional Logic To Children

Dr. Robert H. Seidman

Research Report 86-2
January 1986

Robert H. Seidman

TRANSDUCTIVE REASONING AND THE TEACHING OF CONDITIONAL LOGIC TO CHILDREN

TABLE-OF-CONTENTS

TABLE-OF-CONTENTS (continued)

TRANSDUCTIVE REASONING AND THE TEACHING OF CONDITIONAL LOGIC TO CHILDREN

A. Introduction

The ability to make acceptable inferences is central to almost every educational activity. This kind of reasoning is so closely related to problem solving and intelligence that, as Sternberg (1982) points out, it is often very difficult to tell the three apart.

This report focuses upon a body of literature pertaining to children's conditional reasoning ability, a type of deductive reasoning.[1] The literature is surveyed from two main perspectives: 1) natural-cultural influences - to what extent can children do conditional reasoning without direct instruction in it and 2) influences of instruction - what are the affects of direct and indirect instruction in the principles of conditional reasoning upon the child's ability to do conditional reasoning?

In addition, an important thread which runs through many of the studies is exposed. This thread shows that a great number of children utilize the conditional statement as if it were a biconditional statement. That is to say, they interpret the conditional statement in a biconditional manner. Furthermore, there is some evidence to indicate that this interpretation is not a truth-functional one.

Conditional logic[2] is a type of propositional logic that utilizes the logical connectives, "if", "if-then", "only if", and "if and only if" to connect antecedent and consequent propositions (represented by "p" and "q", respectively). I focus my study upon the "if-then" connective since "if p then q" (called the conditional statement) is a fairly commonly used locution in ordinary discourse. In addition, it is often used to reflect deductive necessity: a conclusion follows necessarily from a premise or set of premises by virtue of the formal structure of the premise(s) and conclusion without appeal to empirical evidence.

Logicians have agreed that the binary truth function representing the conditional statement renders the statement true for all truth value combinations of its antecedent and consequence except when p is true and q is false. This interpretation of the conditional statement is called material implication (or material conditional). Ennis and Paulus (1965) identify twelve principles of conditional logic. However, only the four most widely researched principles are considered in this review. They can be viewed as corresponding to four conditional arguments.

Consider the conditional statement as the first premise in a two premise argument which has as its second premise the affirmation or denial of either the antecedent or consequence of the first premise. The conclusion of the argument is either the affirmation or denial of the proposition which is not present in the second premise. Of these four possible arguments two are valid arguments (i.e., the conclusion follows logically from the premises) and two are invalid (i.e., the conclusion does not follow logically from the premises). The arguments and the conditional principles corresponding to them are (given the conditional statement, "if p then q," as the first premise):

1. Forward Conditional. (Modus Ponens) The affirmation of the antecedent (p) implies the affirmation of the consequent (q).

2. Inversion. The denial of the antecedent ($\bar{p}$) does not by itself imply the denial of the consequent ($\bar{q}$).

3. Conversion. The affirmation of the consequent (q) does not by itself imply the affirmation of the antecedent (p).

4. Contraposition. (Modus Tollendo Tollens) The denial of the consequent ($\bar{q}$) implies the denial of the antecedent ($\bar{p}$).

Principles 1 and 4 are called validity principles because a particular conclusion follows necessarily from the premises. Principles 2 and 3 are called fallacy principles because no conclusion follows necessarily from the premises and to assert one is to draw a falacious inference.

The correct utilization of the principles of conditional logic in ordinary discourse is an important part of reasoning logically. Logical reasoning (i.e., the drawing of valid inferences and correctly judging whether or not a statement follows from other statements necessarily) is thought to be an essential part of critical thinking. See Ennis (1962) and Smith (1957).[3] Conditional statements play a significant role in discourse associated with such things as explanation and argumentation. Laws, causal relationships, possibilities as well as evidence relationships are frequently expressed in terms of conditional statements.

The interpretation of the conditional statement is not without controversy. There is some empirical evidence that the conditional statement is interpreted in a manner other than material conditional.

Studies show that many children tend to interpret the conditional statement as if it were a biconditional statement.

The "if p only if q" locution (called a biconditional statement) is considered by logicians to be true whenever the truth values of the antecedent and consequence are identical and false whenever these truth values are different (material equivalence or biconditional interpretation). Thus, if the conditional statement ("if p then q") is interpreted in a biconditional manner then the conclusions drawn fallaciously under the material conditional interpretation become validly drawn conclusions. Thus, Principles 2 and 3 cease to be fallacy principles.

In addition to the interpretation controversy, there exists a controversy as to whether or not the "if-then" connective can truly be considered a truth-functional connective. It is important to note that the literature reviewed here deals with the ordinary language usage and interpretation of conditional statements and does not consider the formal aspects of logical conditionals. See Barker (1969) for a formal analysis.

The studies reviewed here show, for the most part, that children do much better on the validity principles than on the fallacy principles and efforts at teaching the principles of conditional logic have largely been unsuccessful. However, there is some evidence of increasing success on both validity and fallacy principles with age and developmental level.

There is, in these studies, a thread that suggests that many children in the concrete operational state of cognitive development interpret the conditional statement in a biconditional sense. This would account, to some degree, for the differential performance on the validity and fallacy principles. This is not to say, however, that children necessarily interpret the conditional statement truth-functionally. I note that some of the results of empirical research can be explained on non-logical (non-truth-functional) grounds. In fact, there is some evidence that "transductive" reasoning operates in these circumstances. According to Piaget, transductive reasoning is not true deduction because the tendency is to juxtapose things - to make associate "and" connections rather than ascribe logical necessity to the situation.

Sixteen studies spanning a twenty year period are presented in varying degrees of detail and are separated and synthesized by three Analysis sections.

B. The Research-I

1. Pre-1960 Research

Although a few studies show that children between the ages of 5 and 10 years can at least do some propositional logic, none of the studies focus upon any of the individual principles of conditional logic. See Bonser (1910), Burt (1919), Winch (1921), Moore (1929), Woodcock (1941) and Donaldson (1963). Although Donaldson's study is chronologically post-1960, its spirit is pre-1960. In fact, psychologists appear to have sorely neglected the area of logical reasoning in general. For a review of the logical reasoning literature to 1929 see Moore (1929), to 1943 see Robitaille (1943). Wilmoth (1973) reviews the literature since 1920 and includes children's logical reasoning. Morgan and Morgan summarize:

> During the last 50 years no systematic and comprehensive approach has been undertaken by them [psychologists] toward these problems [logical reasoning]. We made a careful search of the literature since 1927 and found 21 references to experimental studies of logical reasoning, and we were rather generous in our interpretation of what constitutes an experimental study. (Morgan and Morgan, 1953, 400)

We begin this survey with the study by Shirley Hill (1961) which was highly touted by Suppes and Binford (1965), revised by O'Brien and Shapiro (1968; 1970) and which began a tradition highlighted by Ennis's Cornell Critical Thinking Project. Out of the Hill-Ennis tradition came a number of research endeavors, each adding something to our knowledge of the young child's conditional reasoning abilities. A slightly different approach began in Europe with Matalon (1962) and Peel (1967). This approach postulated non-standard truth functional explanations for children's conditional reasoning abilities.

Finally, an approach by Wason and Johnson-Laird and their followers, although dealing mainly with adult subjects, established yet another tradition which is useful in viewing the problem. From the studies examinmed here we come to see the important role that the proper interpretation and assessment of the conditional statement plays. Evidence is gathered for a biconditional-like interpretation of the conditional statement rather than the standard material conditional interpretation.

2. <u>Hill (1960)</u>

The purpose of Hill's study was

> ... to examine the abilities of first, second and third grade children to derive valid logical inferences from sets of verbal premises.
> ...[and to answer the question] to what extent have these rules of inference been incorporated into the cognitive equipment of children in the early elementary school years? (Hill, 1960, 1-2)

It should be noted that Hill's subjects were asked merely to "recognize" logically valid conclusions and did not actually derive these conclusions in any constructive sense. Also, the rules of formal logic provided Hill with a conceptual model for "correct" deductive thinking:

> It is assumed that the relations dealt with by formal logicians are not just the esoteric possessions of professional logicians but are commonly accepted modes of correct thinking and that growth in logically correct thinking in children is the result of the acquisition ... of the basic principles of logical theory. (ibid., 6)

We shall see later that this assumption by Hill leads to great difficulties in trying to reconcile the empirical evidence reported in some of the studies surveyed here.

Hill's 100 item test consisted of three subtests: classical syllogisms, logical quantification with two-place predicates, and sentential logic (60 items). Our only concern is with the latter subset. All of these subtest items consisted of two premises, a question and two possible responses ("YES" and "NO"). The test items required the subject to make a discrimination between a <u>necessary</u> conclusions (i.e., valid) and the negation of a necessary conclusion. These two types of conclusions are evenly distributed over the 60 test items.

A typical test item in the sentential logic subtest was:

If this is Room 9, then it is the fourth grade.
This is Room 9.
Is it fourth grade?

a. YES b. NO

The content of the items were familiar but not suggestive and the vocabulary was deemed appropriate for the grades tested. A variant of the above test had the second premise and both the antecedent and consequent of the first predicate negated.

The overall test reliability coefficient was 0.90 with no subtest reliabilities reported. Hill's sample consisted of 270 middle-class and middle-income six, seven and eight year olds who were divided into three groups of 25 males and 25 females each. A fourth group of 20 eight year olds who were advanced readers was utilized to check for reading effect. The Standard Group (I) received no correction to test responses. The Reinforcement Group (II) received the correct answer after each response and the Baseline Group (III) received no reinforcement but the first premise in each item was eliminated to check for content effect. Table 1, below, shows the categories of sentential logic that Hill tested.

TABLE 1

CATEGORIES OF SENTENTIAL LOGIC TESTED BY HILL

1. Modus Ponendo ponens
2. Modus tollendo ponens
3. Modus tollendo tollens
4. Law of hypothetical Syllogism
5. Hypothetical Syllogism and Tollendo Tollens
6. Tollendo tollens and tollendo ponens
7. Ponendo pones and tollendo tollens

Hill found that the mean percentage of correct responses for ages 6, 7 and 8 were 74.30, 80.42 and 85.54, respectively. In addition, Hill found that the percentage of correct responses over all principles and over all grades were above 74 in all but three cases (and in these cases none was lower than 65). This led Hill to conclude:

> The high percentage of successful responses of all standard groups ... suggests that the children in the age group six through eight years are able to recognize valid conclusions derived from hypothetical premises. (ibid., 50)

For both the overall sentential logic scores and for the individual test items, Hill found highly significant F ratios beyond the 0.001 level for age. The same held true for the other two subtests, leading Hill to conclude that logical ability increases with age. Also, Hill found that Groups I and II differed significantly at the 0.01 level on the overall test and concluded that "determinant reinforcement" improved performance in recognizing logically valid conclusions and that Group II "learned" to use logical principles. In addition, an item analysis revealed that the addition of a negated proposition had a significant effect (beyond the 0.001 level) on the difficulty of the item.

Hill concluded that her research showed that children in the concrete operations stage have the "... ability to recognize valid conclusions derived from hypothetical premises." (ibid., 69) Finally, Hill's analysis showed that there were no statistically significant differences between male and female mean scores.

Suppes (1964; 1965) and Suppes and Binford (1965), citing Hill's study claimed that six, seven and eight year olds:

> ... are able to deal very effectively with verbal premises that call for hypothetical reasoning and are by no means limited to 'concrete' operations. (Suppes and Binford, 1965, 34)

3. O'Brien and Shapiro (1968)

O'Brien and Shapiro believed that the particular "behavioral manifestations" of hypothetical-deductive thought examined by Hill were too narrow. Hill's subjects were called upon only to recognize logical conclusions (or negations of logical conclusions) and not to test the conclusion's logical necessity. O'Brien and Shapiro utilized two measuring instruments. Test A was identical to Hill's test and Test B was the same as Test A, except that:

> (1) 33 of the original 100 items were 'opened up' so that no necessary conclusions followed from the premises, and (2) for every item in Test B, a 'NOT ENOUGH CLUES' option was added to the 'YES' and 'NO' response clues provided in Test A. (O'Brien and Shapiro, 1968, 533)

For example, the sample item given in the Hill study now looks like this:

If this is Room 9, then this is fourth grade.
This is not Room 9.
Is it fourth grade?

a. YES b. NO c. NOT ENOUGH CLUES

Hill's test had 50 items keyed "YES" and 50 items keyed "NO". Test B had 34, 33 and 33 items keyed "YES," "NO", and "NOT ENOUGH CLUES," respectively. Twenty-five subjects (12 male and 13 female) from a school in an upper-middle-class suburb were randomly assigned to each of two tests at each age-grade level. Unlike the Hill study, this study did not evaluate the effect of reinforcement on logical abilities and did not consider the effects of visual and oral test presentations.

Test A analysis confirmed almost all of Hill's findings except that a Scheffe's 'a posteriori' test showed signficant differences only between the means of the 6 year olds and those of the seven and eight year olds. This result held for the sentential logic items. For Test B, no statistically significant difference was detected between age level means except for the 33 altered items (significant at the 0.05 level). The difference between the subject's performance on the two tests and on the sentential part of the two tests were significantly different with Test B lower than Test A at all age levels. This led O'Brien and Shapiro to conclude that:

> Children of the same age ... experience great difficulty in testing the logical necessity of a conclusion, and they show slow growth in this ability. (ibid., 537)

It is interesting to note that the performance on the 33 open items in Test B was, at all grade levels, below chance.

O'Brien and Shapiro conclude that contrary to Suppes' claim: "... hypothetical-deductive ability cannot at all be taken for granted in children of this age." (ibid., 539) Like Hill, there were no significant differences due to sex or to ageXsex interactions.

4. Shapiro and O'Brien (1970)

In this study, the same tests given in O'Brien and Shapiro (1968) were administered to subjects over a wider age-grade range (grades 1 through 8, ages 6 through 13) in an upper-middle-class suburban school. There were 48 randomly selected subjects at each grade level. 24 (12 males and 12 females) were randomly assigned to receive Test A and the rest were assigned Test B.

Once again, analyses of Test A confirmed many of Hill's original findings for grades 1 through 3. Scheffe's 'a posteriori' tests showed that statistically significant differences (0.05 level) existed between: the grade 1 mean and that of each of the other grades and the grade 4 mean and those of grades 6 and 8. Although within each of these stages no statistically significant differences between grade-age means existed, every one of the older age level means were significantly higher than any of the means at the prior age stages.

Further analysis found that for every age-grade level the total test means for Tests A and B differed significantly. In each case, Test B total means were quite a bit lower than Test A means. Analysis of covariance on the 33 altered items and the 67 unaltered items showed that these test differences were primarily due to the differential performance on the the 33 altered items. With respect to the 33 altered items, statistically significant differences appeared at each of the grade-age levels and in every case the Test B mean was "substantially lower" than the Test A mean. These results confirm the findings in the O'Brien and Shapiro (1968) study.

Shapiro and O'Brien examined the responses given for the 33 open items in Test B and found that an intriquing pattern emerged. Out of all of the items in the sentential logic subtest, 16 were of the "if-then" conditional form. Shapiro and O'Brien found that a great many subjects "consistently" interpreted the "if-then" conditional statement as the biconditional statement, "if and only if." Using this biconditional interpretation, they rescored the subject's responses. The results are given in Table 2 and are taken from Shapiro and O'Brien (1970, 828).

TABLE 2

FREQUENCY OF CORRECT RESPONSES "IF-THEN" OPEN ITEMS IN TEST B
(Shapiro and O'Brien, 1970)

	Key	
Grade/Age	"If-Then"	"If and Only If"
1/6	12	320
2/7	14	316
3/8	26	310
4/9	55	289
5/10	58	298
6/11	70	298
7/12	90	272
8/13	137	244

Chi-squared tests of the frequencies are not appropriate here because of the lack of independence. However, the considerable differences are quite apparent. Note that the number of biconditional responses decreases with increasing age and that the number of "if-then" responses increases with increasing age. Shapiro and O'Brien remark:

> This pattern suggests that as children get older, they tend to switch from a child's logic interpretation to one which is consistent with mathematical logic. (ibid., 828)

Since the authors believe that their tests (and Hill's) measure hypothetical-deductive thinking, they conclude that their results:

> ... suggest that hypothetical deductive thinking - at least that which is consistent with mathematical logic - cannot at all be taken for granted in students of elementary-school age. (ibid., 829)

It is possible, Shapiro and O'Brien concede, that children 6 to 13 years of age follow a "consistent 'child logic'" (a material conditional interpretation of the conditional statement) which differs from "'adult' mathematical logic" (a biconditional interpretation). This is a suggestion that will gain in importance as this literature survey proceeds.

5. Ennis and Paulus (1965)

a. The Experiment

This study (hereafter, Ennis), massive in its scope and precise in its objectives, paved the way and provided a paradigm for a number of subsequent investigations. The study focused upon the "natural-cultural" development of logic mastery and the "readiness" for mastery of logic of students in grades 5 through 12 (10 years to 18 years of age). Since a version of this test is used in my study, a thorough examination is presented here.

The readiness question was concerned with the ages at which and the extent to which youngsters were ready to master logic. This involved the deliberate instruction in various logical principles. The natural-cultural developmental question was concerned with the ages at which and extent to which youngsters mastered logic without the benefit of instruction in logic. Ennis delineates twelve principles of conditional logic arguments and twelve principles of class logic arguments for testing. We are interested in only the first four principles of conditional logic: Forward Conditional, Inversion, Conversion and Contraposition.

Conditional logic was tested for at grade levels 5, 7, 9, and 11 in two schools. LTD stands for "logic deliberately taught" at school 1. LNDT-1 and LNDT-2, stand for "logic not deliberately taught" in schools 1 and 2, respectively.

Ennis devised The Cornell Conditional-Reasoniong Test, Form X which contained 72 items in 12 item groups (corresponding to 12 conditional logic principles) of 6 items each. The six items of any one group were scattered throughout the paper and pencil test. Getting at least 5 of the 6 items in an item group correct was deemed a sufficient condition for probable mastery of the principle associated with the item group. Getting at least 4 correct was judged to be a probable necessary condition of mastery. See Ennis (1964) for his theory of "operational definition."

The Cornell Conditional Reasoning Test, Form X (hereafter, "test") includes three content components for each item group: the concrete familiar (4 items), the symbolic (1 item), and the suggestive component (1 item). The concrete familiar (CF) items are things that are not abstract and are reasonably familiar to the test population:

> [T]here is no reason to believe that a subject will, because of background factual knowledge, accept or reject the conclusion of concrete familiar items. (Ennis, 1965, VIII-10)

Example 1 is a concrete familiar item for Principle 2 (Inversion).

Suppose you know that
If the automobile in the parking lot belongs to Mr. Brown, then it is black.
The automobile in the parking lot doesn't belong to Mr. Brown.
Then would this be true?

The automobile isn't black.

A. YES
B. NO
C. MAYBE

Example 1. Concrete Familiar Example for Principle 2.
("MAYBE," as explained in the test directions, means "not enough clues to answer yes or no).

Symbolic component (SY) items use symbols (not words) to refer to objects in the sentences. See Example 2. Suggestive component (SU) items are such that the truth status of the conclusion of the argument differs from the validity status of the argument. See Example 3.

Suppose you know that
If there is an X, then there is a Y.
There is not an X.
Then would this be true?

There is not a Y

A. YES
B. NO
C. MAYBE

Example 2. Symbolic Component Example of Principle 2.

Suppose you know that
If whales are birds, then they can fly.
Whales aren't birds.
Then would it be true?

Whales can't fly

A. YES
B. NO
C. MAYBE

Example 3. Suggestive Component Example of Principle 2.

The test directions ask whether the conclusions would be true on the assumption that nothing but the premises are known. The three possible answers are explained in Table 3. Table A-1 exposes the logical form of the test items. Subjects are assumed by Ennis to have mastered the principle of double negation, i.e., two negatives result in a positive statement.

TABLE 3

POSSIBLE TEST ANSWERS

A. YES:	It must be true.
B. NO:	It can't be true.
C. MAYBE:	It may be true or it may not be true. You weren't told enough to be certain whether it is "YES" or "NO."

b. Natural-Cultural Results

Let us now consider Ennis's results concerning the natural-cultural development of the ability to understand conditional logic. Ennis isolates four major features of Piaget's theory with respect to the formal operational period:

> ... possession of the truth-validity characteristic, ability to operate within the framework of a combinatorial system, ability to control variables, and ability to do propositional logic .... (ibid., V-8)

Ennis's study deals with the first and fourth of these features because they, according to Ennis, are "clearly" related to the ability to do deductive logic. His results can be summarized in five parts.

1. Table A-2 together with Tables A-3 and A-4 show that there is development in logical ability with age.

2. Ennis was unable to definitively answer the question of whether there existed stages in the learning of the logical principles. His results were too inconclusive.

3. Consider the question: is conditional logic mastered by age 11-12 (grades 5 and 6)? Seventh graders clearly did not master the fallacy principles (2 and 3) and do not show great mastery of the validity principles (1 and 4). This can be seen in Table A-4. When we consider all 12 principles, grades 5 and 7 have mean pretest scores of 42 and 52, respectively, out of a possible 99. It appears that on the whole, the 12 principles tested for by Ennis were not mastered by subjects at this grade level. The mean difficulty indices for all 12 items, indicate that conditional reasoning is not mastered by age 11-12. Despite a large improvement from grade 5 to 11, the latter group seems not to have mastered conditional reasoning according to the Ennis criteria.

4. Ennis answers, "partly," to the question of whether the truth-validity characteristic is achieved by age 11-12. This characteristic is represented by the suggestive component items where the truth status of the conclusion is different from the validity status of the argument.

For grades 5 and 7, the mean difficulty indices for the SU items are 41.3 and 53.4, respectively. For grades 11 and 12, the indices are 60 and 70, respectively. Thus, although the truth-validity characteristic is partially achieved at age 11-12 and to a greater extent by age 18, it is not completely achieved at any age.

5. Is the acquistion of conditional logic differentially developed? This question has a number of interesting aspects. First I look at the logical principles and then at the three kinds of item components.

The <u>fallacy principles</u> (Principles 2 and 3) are clearly differentiated from the <u>validity principles</u> (Principles 1 and 4) by the raw scores, the difficulty indices and the necessary and sufficient condition percentages. In fact, the fallacy groups have the four lowest mean difficulty indices. The same holds true for the necessary and sufficient condition percentages with but one minor exception. Some of the largest improvements across grades are made for the fallacy groups.

A statistical comparison (Tukey test and t-tests) between the mean difficulty indices of the three <u>item components</u> show the greatest development is on the suggestive component, followed by the concrete familiar and lastly the symbolic. None of the differences are statistically significant. Ennis speculates that the non-significant differences between the concrete familiar and suggestive components might be because the suggestive component is a part of all conditional items, due to their "iffy" nature:

> What the subject is asked to suppose in the conditional statement on the conditional reasoning test is not simply that something is the case but rather that, on <u>another supposition</u> (the if-clause) something would be the case.
> ....
> But with the concrete familiar conditional statements, what he is working with, though he might well believe it to be true, is the <u>implication</u> of another <u>supposition</u>. Thus he is forced to think in terms of what is implied, rather than what is true. (ibid., V-39)

Ennis found no significant differences between CF and SY components. He postulates that this may be due to the way the symbols (letters) were used. The test statements refer to the existence or non-existence of letters and this is presumably no more difficult than dealing with familiar categories. Finally, the order of difficulty is the same as in Wilkins (1928), in ascending order: CF, SU and SY.

c. Direct Teaching of Logic Results

The second part of Ennis's study dealt with the mastery of propositional logic (strictly, Ennis measured "readiness" to master logic). In short, the LDT's were taught the conditional logic principles for 15 days, one school period of 50 minutes each day and were given the post-test (identical to the pre-test) six weeks later. Ennis found that 5th and 7th graders were not ready to master the logical principles and that the 11th graders made great improvements. These improvements were quite evident among the fallacy principles and occurred amongst the validity principles as well. Let us now look at just what conditional logic the subjects learned.

<u>Total Scores</u>. Ennis adjusted the means by analysis of covariance, partialling out IQ and pre-test scores, and found that there was no significant difference for grades 5 and 7 (on all 12 principles) but that the LNDT-1 and LDT show significant differences for 9th and 11th grade results. The 9th grade results favor the LNDT's whereas the 11th grade results favor the LDT's.

Component Scores: The 11th grade LDT's did significantly better on all three components scores than the LNDT's but there was a significant difference favoring the LNDT-1's for the concrete familiar component.

Item Group: For grade 11, the fallacy items and the Contraposition item significantly favor the LDT's. There is a lack of any great improvement for the Forward Conditional and for grade 5, it is only the fallacy item, Conversion, that shows a significant difference for the LDT's.

Ennis concludes, that on the basis of his study:

> [T]here is not much point in trying to teach conditional logic in elementary and lower secondary [schools]. Furthermore, these results suggest that the things that can be taught are the fallacies, contrapostion to some extent, and perhaps the validity of affirmation of the antecedent, which might be partly teachable, but also seems to develop on its own without deliberate teaching. (ibid., VI-24)

Ennis suggests that for the fallacy items, knowledge of logic was not improved but that vocabulary was changed. He posits that perhaps "if-then" was interpreted as the biconditional and only later was the proper terminology learned through teaching (the "if-then" conditional was only one of the twelve principles taught). Ennis does not elaborate upon this idea and only notes that it "bears further investigation."

6. Ennis (1969)

a. The Experiment

Here, Ennis attempts to determine whether first, second and third graders have mastered and whether they are ready to master four basic principles of conditional logic: Inversion, Conversion, Contraposition and Transitivity. A basic understanding of the Forward Conditional is assumed and checked for throughout the study. Parts of this study are reported in Ennis (1971). Three principles are considered here.

In this study, the components of the logical principles have been narrowed to suppositional and factual because the paper and pencil test has been abandoned in favor of a test demanding some apparatus. The component restrictions also serve as a way to examine Piaget's "claim":

> [T]he child cannot reason from premises without believing in them. Or even if he reasons implicitly from assumptions which he makes his own, he cannot do so from those which are proposed to him. (Piaget, 1928, 252)

For example, the child is shown a model of a house and is told that if a handle on the house is in the up position, the bell on the house works, (if p, then q). This is the Forward Conditional. The child is asked to suppose that the handle (now hidden) is up, (p), and then is asked whether the bell would work, (q). The child is then asked to tell why he answers the way he does. Success in this manner, Ennis claims, counts as evidence toward a refutation of Piaget's above "claim" and is an example of a suppositional conditional. After the child responds to the suppositional conditional, he is shown the handle in an up position (p) and is asked whether the bell works (q) and why. This is termed a factual conditional.

Ennis's developmental question is:

> To what extent have 6-9 year olds of various sorts already mastered particular basic principles of conditional logic, as a result of natural-cultural forces? (Ennis, 1969, 8)

Mastery was given an operational definition similar to the one in the previous study (Ennis, 1965).

Total scores simply reflected the number of correct answers because no credit was given unless both a correct answer and a good justification were given. Mastery of suppositional ability was not sought. The readiness question investigated was:

> To what extent are a variety of children ready to learn the basic features of conditional logic? (ibid., 13)

The instruction involved a series of 15 audio-tutorial lessons designed to impart a knowledge of conditional logic via a science content. Three elementary schools (urban, rural and suburban) contributed subjects to the study. Ten students in grades one, two and three in each school were randomly chosen to be in the control group and an equal number were similarly selected to be in the experimental groups. In all but the urban school, the experimental and control subjects were drawn from different classrooms. There were thus 30 subjects for each grade level.

The test. The study design was a post-test-only control group (Campbell and Stanley, 1963). Because of motivational and reading problems endemic to primary school children, the individual interview technique and concrete apparatus test materials were considered appropriate. The test is divided into two parts using different test material. The house part involves a model house and the chemical part involves chemical reactions.

The model house has two knife switches (one large and the other small) and a push-button switch on the outside. A doorbell and a light (visible when lit and seen through an open window) are inside the house. The child is shown the two handle-switches but is not allowed to know the function of the small one. He is allowed to find out that the bell does not always ring when the button is pushed. Two possibilities are left open for the child: 1) the small handle being up implies that the bell works and 2) the bell works only when the large handle is up. Table 4 (from Ennis, 1969, 39) shows the conditional statements that the subjects are asked to reason about on the house part of the test.

TABLE 4

HOUSE PART CONDITIONAL STATEMENTS

1. If the big handle is up, then the bell works.
2. If the light is on, then the big handle is up.
3. If the bell does not work, then the big handle is down.
4. If the big handle is down, then the light is not on.

The subjects are shown, and thus assume, that the large handle being up is a sufficient condition for the bell to work. The existence of the second switch, of course, is to provide the child with the knowledge of the existence of a specific alternative to consider.

Table 5 (from Ennis, 1969, 40) shows the conditional statements from which the subjects are asked to reason from on the chemical part of the test.

TABLE 5

CHEMICAL PART CONDITIONAL STATEMENTS

1. If a white powder is soda, then it bubbles when vinegar is added.
2. If a white powder is sugar, then vinegar added to it turns white.
3. If a liquid is vinegar, then it makes soda bubble.
4. If a liquid makes soda bubble, then it turns litmus paper red.

Immediately preceding the chemical part of the test, the subjects are shown a number of different reactions with vinegar and unidentified white powders. This is done so that they realize that some white powders turn vinegar milky and other white powders make vinegar bubble. In presenting each question, the tester first teaches or reviews the major premise(s) using materials and appropriate pictures. If the subject demonstrates a recall of the major premise(s) she is:

1. for the suppositional questions, asked to suppose that one part of the conditional statement is true (or false).

2. for the factual questions, shown that one part of the conditional statement is true (or false).

The subject is then given three answers from which to choose:

1. affirmation of that part of the conditional statement that is under consideration,

2. denial of that part of the conditional statement that is under consideration,

3. neither affirmation nor denial of that part of the conditional statement under consideration.

After the subject responds, the tester asks for a justification of the answer. The suppositional form is always asked first.

Test Structure: Out of a total of 24 items, there are 6 items for each of the four principles (each group of 6 items corresponding to a principle is an item group). Within each item group, 3 items are suppositional and 3 are factual. Table A-5 describes each item. Subscores for the four principles, the fallacy and validity principles, suppositional and factual forms and the house and chemical parts (as well as total scores) are computed. Probable mastery of a principle is judged for a score of 5 or 6, probable non-mastery for a score of 3 or less and judgement is withheld for a score of 4.

b. Developmental Results

The results of the control group are shown in Tables A-6 , A-7 and A-8 and can be summarized as follows.

Both Tables A- 6 and A- 7 show that there is considerable variation from principle to principle. Table A-8 shows the results of significance tests done on the differences among principle scores, between suppositional and factual items and between validity and fallacy principles. All of the differences are statistically significant at the 0.05 level. There is a marked difference between the validity and fallacy principles with subjects doing much better on the validity principles than on the fallacy princples.

These three tables show that the subjects did consistently better on the factual items than on the suppositional items. One explanation has parallels in the test-retest situation: students improve even without instruction. The subjects were first asked to suppose the minor premise(s) and then after answering the test item, were shown the facutal counterpart. Another explanation parallels Piaget's claim that it is more difficult for young children to work with possibilities than with actual situations. Whatever the explanation, suppositional ability was demonstrated by a large number of subjects.

The Pearson Product Moment correlations for sex, age, SES, IQ and dwelling area with the test scores for both the control and experimental groups show no significant correlations and almost no correlation between sex and anything else. The chronological age correlation with total test scores was low and might be due to a plateau in the logical development within the age ranges studies, according to Ennis. There appears to be a positive relationship between SES and logical ability with higher correlations in the control group than in the experimental group. Perhaps, as Ennis suggests, the logic instruction made up in part, for the SES differences, although the overall effect of instruction was nil (see "Readiness Results", below).

Out of all the correlations, the ones between IQ and conditional logic scores were the highest, with verbal IQ the highest. Taking the grades separately, these correlations were even higher for second and third grades. Finally, no significant difference was found between the logic test scores and dwelling area.

c. Readiness Results

Ennis concludes that, on the basis of the data:

> ... the methods which we used to try to bring about the mastery of the four basic principles of conditional logic were not successful. (ibid., 64)

Only one out of 30 comparisons between control and experimental groups was statistically significant and that one favored the control group (suburban subjects on the Contraposition principle).

Why was the instruction unsuccessful? Ennis offers four possible explanations, rejects the first two and is unable to decide between the latter two.

1. Logical abilities cannot be hastened, they unfold on their own. Ennis notes that in his earlier study (Ennis, 1965), upper secondary school children did learn logic.

2. Children of the age range tested (6 to 9 years) cannot learn conditional logic principles. However, Tables A-6 , A-7 and A-8 show that many children have learned conditional logic.

3. At this age level, conditional logic cannot be taught, even though children might acquire it on their own and although older children apparently can be successfully taught.

4. The teaching materials used in the study were inadequate.

While explanations 1 and 2 can safely be rejected based upon the data, there exists no sure way of choosing between 3 and 4. The negative results of the readiness study do not necessarily show that children are not ready to master something. Positive results on the other hand, show that the children are ready for mastery. Thus, the question of whether 6-9 year olds are ready to learn more conditional logic than they already know remains unanswered.

C. Analysis-I

The studies examined so far present a somewhat confusing view of the conditional reasoning capabilities of young children, particularly those in the concrete operational stage of development. These studies do, however, give us some evidence for an alternative interpretation of the standards against which the subjects were tested. The Shapiro and O'Brien (1970) study and aspects of both of the Ennis (1965; 1969) studies lend some support for a biconditional-like interpretation of the conditional statement by children. There is some Piagetian theory to support this alternative interpretation.

According to Piaget's theory of ontogenesis of cognitive processes and structures, the child moves from the concrete operational stage of cognitive development to the formal operations stage via reflective abstraction. The formal operational child can now operate under a "complete combinatorial system" whereas the concrete operational child is relegated to mathematical groupings, an incomplete combinatorial system. A formal operational child when confronted with a conditional assertion, "if p then q," can form a combinatorial system of p and $\bar{p}$ conjoined with q and $\bar{q}$:

$$\begin{aligned} &1.\ p.q, \\ &2.\ \bar{p}.\bar{q}, \\ &3.\ \bar{p}.q, \\ &4.\ p.q, \end{aligned} \tag{1}$$

and can consider the possibility of 1, 2 and 3 as true. Note that "if p then q" is equivalent logically to

$$(p.q)v(\bar{p}.q)v(\bar{p}.\bar{q}) \tag{2}$$

and

$$\neg(p.\bar{q}). \tag{3}$$

The three conjunctions in (2) together with the negation of (3) form the combinatorial system of formal operations. Since the concrete operational child is unable to form this system, he is unable to consider p.q, $\bar{p}.\bar{q}$ and $\bar{p}$.q as true, given the "if p then q" conditional statement. But <u>which</u>, if any of the conjunctions in (1) would the concrete operational child consider? <u>Transductive reasoning</u> provides a clue.

Transductive reasoning is neither true inductive nor true deductive reasoning. It is associative rather than implicative and causal. For the young child, elements visually close tend to be juxtaposed in her reasoning. The child moves from particular to particular. As Flavell puts it:

> Centering on one salient element of an event, the child proceeds irreversibly to draw as conclusion from it some other perceptually compelling happening. Piaget makes the important point that the factual correctness of the child's conclusion... is by itself no guarantee that the mechanism for arriving at it was logical rather than transductive. (Flavell, 1963, 160)

The child who reasons transductively tends to juxtapose elements, thus making associative "and" connections.

> [J]uxtaposition is after all the sign of the complete absence of necessity from the thought of the child. The child knows nothing either of physical necessity (the fact that nature obeys laws) nor the logical necessity (the fact that such a proposition necessarily involves such another). For him everything is connected with everything else, which comes to exactly the same thing as nothing is connected to anything else. (Piaget, 1928, 60)

According to Piagetian theory, transductive reasoning is not fully extinguished and replaced with logical necessity until the formal operations stage is reached (Piaget, 1926; 1928).

Thus, the O'Brien and Shapiro (1968) sample test item can be translated into "transductive logic" terms. This translation and the next are taken from Knifong (1974):

Room 9 and the fourth grade co-occur.
This is Room 9.
Is this the fourth grade?

a. YES b. NO

Transductive logic tells us that the "correct" response is "YES." "NOT ENOUGH CLUES" has no transductive interpretation and thus may be meaningless for the subject. Thus, if transductive logic is operative on the original question, the child answers correctly according to the standard mathematical logic interpretation, but for the wrong reasons. Note that all of the studies surveyed thus far show a notable success on the Forward Conditional principle (modus ponens).

Here is how the Inversion question is translated using transductive reasoning:

Room 9 and the fourth grade co-occur.
This is not Room 9.
Is it the fourth grade?

a. YES b. NO

Using transductive reasoning, the "correct" response would be "NO," which is of course incorrect under the material conditional interpretation of the conditional statement. Similar cases can be made for Conversion and Contraposition.

This transductive anlaysis shows that the child may very well be utilizing a biconditional-like interpretation of the conditional statement in her reasoning. The Shapiro and O'Brien (1970) study gives evidence for this interpretation and the poor showing on the fallacy items and the good showing on the validity items in the Ennis studies (1965; 1969) also lends some support for a transductive explanation.

Knifong (1974) reexamines some of Ennis's (1969) raw data. See Table 6. Since the "MAYBE" (similar to "NOT ENOUGH CLUES") was so infrequent, Knifong excluded them from his statistical analysis. It is not clear just why Knifong did not include the Inversion principle in his table. Table 6 strongly supports the transductive interpretation of the conditional statement. Note that for Principle III (Contraposition), the incorrect answer "MAYBE" was chosen about as often as when it represented the correct response in the other two principles (for each grade level). Knifong believes that this pattern implies:

> ... that these children may have been responding from an "I don't know" rather than the more sophisticated "I don't know and it is impossible to know" understanding of the question. (Knifong, 1974, 81)

Thus, the success on Contraposition that Ennis found could very well be attributable to a transductive interpretation. This notion bears further examination, for as Flavell so aptly points out:

> ... the factual correctness of the child's conclusion is by itself no guarantee that the mechanism for arriving at it was logical rather than transductive. (Flavell, 1963, 160)

D. <u>The Research-II</u>

1. Peel (1967)

This study lends support for a transductive biconditional-like interpretation for children's reasoning using conditional (Peel calls them "implication") statements and arguments. This study is important because it utilizes a wholly different methodological approach than Ennis, Hill, and O'Brien and Shapiro.

TABLE 6

PERCENTAGES OF EACH OF THE THREE RESPONSES REFLECTING VARIOUS KEYS ON THE SMITH-STURGEON CONDITIONAL REASONING TEST (Knifong, 1974)

	Principle I: p-->q[a] not p			Principle II: p-->q[a] q			Principle III: p-->q[a] not q		
	Question:			Question:			Question:		
	(q)	(not q)	(maybe)	(not p)	(p)	(maybe)	(p)	(not p)	(maybe)
	Other	Transductive Logic	Math Logic	Other	Transductive Logic	Math Logic	Other	Transductive and Math Logic	Other
First Grade (N = 53)	1	70	29	4	73	24	2	72	26
Second Grade (N = 58)	1	56	43	3	82	15	1	74	25
Third Grade (N = 57)	0	68	32	3	85	12	4	80	16
Average for all groups (N = 168)	1	64[b]	35[b]	3	80[b]	17[b]	2	70[b]	22[b]

Source - Ennis et al., 1969

[a] The letters p, q, r, and s are used to represent logical statements, and --> is used to represent the logical conditional.

[b] For Principles I and II the percentages of responses reflecting the transductive and math logic interpretations were not significantly different from their respective averages for all age groups. For Principle III the percentage of responses for the (not p) and (maybe) responses were not significantly different from their respective averages for all all groups.

Peel, by using a *gaming* technique avoids the pitfalls of purely formal implication, where the propositions are not necessarily naturally or causually related and avoids quasi-natural settings such as the Ennis house and chemicals test. Peel creates a novel situation, unaffected by previously learned relations (e.g., a parent's admonition: "if you eat your supper you'll get a story," smacks of a biconditional admonition) and independent of an understanding of propositions masked in the guise of ordinary English sentences.

The game, played between experimenter and child, is extraordinarily simple:

> The experimenter has before him a collection of many beads in a tin, there being at least 10 each of the four colours, blue, yellow, red and green. In another shallow tin is a similar collection of counters, also made up of some 10 each of the same colours. The experimenter and the subject face each other with their beads and counters before each respectively, and an empty box is placed on the table between them into which they make their play. The game consists essentially of the experimenter drawing a certain bead, and then the subject drawing a counter depending upon the rule of the particular game....
> 1. Implication: "We may pick any colour we like, but in this game, if and whenever I draw a red bead, you are to draw a red counter." (Peel, 1967, 83)

Ten pairs of draws (with replacement) are made with each of 150 subjects (only 58 subjects were used for implication, the other subjects were tested on the other logical relationships) in random order consisting for implication of 6 reds and 4 non-reds by the experimenter. All of the experimenter-subject draws (i.e., conjunctions) are recorded.

Let "R" stand for the initial *bead* draw by the experimenter and "r" stand for the responding *counter* draw by the subject. "$\overline{R}$" and "$\overline{r}$" stand for non-red bead and non-red counter draws, respectively. Table 7 gives the theoretically expected proportions of conjunctions according to the material conditional interpretation of the conditional statement and equivalence (biconditional) interpretation. This table also gives the experimental results.

TABLE 7

PEEL'S THEORETICAL AND EXPERIMENTAL RESULTS FOR THE FORWARD GAME

Initial bead draw by experimenter:	R		$\bar{R}$	
Responding counter draw by subject:	r	$\bar{r}$	r	$\bar{r}$
Binary Conjunction:	Rr	$R\bar{r}$	$\bar{R}r$	$\bar{R}\bar{r}$
Truth value according to standard meaning of implication:	true	false	true	true
Truth value according to standard meaning of equivalence:	true	false	false	true
Theoretically expected proportion of implication:	1.0	0	0.5	0.5
Theoretically expected proportion of equivalence:	1.0	0	0	1.0
<u>EXPERIMENTAL RESULTS</u>				
Age group: 5+	0.92	0.07	0.05	0.95
8+	0.99	0.01	0.02	0.98
11+	0.99	0.01	0.03	0.97

Note that the subjects appear to be operating under a biconditional interpretation of the conditional rule when they produce a consequence to the antecedent draw made by the experimenter. Peel cites a study by Matalon (1962) and claims:

> [T]he results of the implication game link closely with those obtained by Matalon (1962) who found that implication is predominantly read as equivalence. (ibid., 90)

The version of the experiment described above is called the <u>forward</u> game. Peel played a <u>reverse</u> game with the same subjects. In the reverse game, the initial bead draw is made by the subject and the experimenter responds with a counter draw. The child must decide whether this counter draw is correct ("acceptable") or incorrect ("non-acceptable"), for the given implication rule. The theoretically expected proportions for implication and equivalence and the exprimental results are given in Table 8.

TABLE 8
PEEL'S THEORETICAL AND EXPERIMENTAL RESULTS
FOR THE REVERSE GAME

Initial bead draw by subject:	R				$\bar{R}$			
Bead-counter conjunction:	Rr		$R\bar{r}$		$\bar{R}r$		$\bar{R}\bar{r}$	
Judgment of subject:	acc.	non-acc.	acc.	non-acc.	acc.	non-acc.	acc.	non-acc.
Theoretically expected proportions of:								
Implications:	1	0	0	1	1	0	1	0
Equivalence:	1	0	0	1	0	1	1	0
<u>EXPERIMENTAL RESULTS</u>								
Age group: 5+	1.0	0	0.03	0.97	0.24	0.76	0.93	0.07
8+	1.0	0	0.01	0.99	0.58	0.42	1.0	0
11+	1.0	0	0	1.0	0.82	0.18	0.98	0.02

Note that the $\bar{R}r$ conjunction proportion change over age has the effect of transforming the equivalence interpretation of the game rule into the material conditional interpretation. It is not until 11+ that this transformation is fairly complete.

Peel's gaming technique allows us to test inferences from the antecedent and enables us to get at the principle of Inversion. Peel's results show us that poor results on the Inversion principle, when assessed by material conditional standards, may be due to a biconditional interpretation of the conditional statement. Improved results are to be expected for older age groups (given the results in Table 8) although forward game evidence does not support this trend.

Peel's gaming technique allows us to fit a theoretical model to empirical results. His study lends support to the biconditional interpretation of the conditional statment ("rule" for Peel) in the concrete operational child and lends credence to a transductive interpretation theory of conditional logic reasoning ability in young children.

2. Taplin, et. al. (1974)

In this study (hereafter, Taplin) the findings support the notion that logical performance improves with age. Taplin's ingeneous analysis suggests that youngsters initially treat the "if-then" conditional as a conjunction, then as equivalence and finally as a material conditional statement.

Taplin shows, using abstract material (as opposed to the more concrete material of Peel and Matalon), that his subjects showed evidence of reasoning truth-functionally, but used alternative interpretations of the conditional statement.

Eight forms of the conditional argument were presented (see Table 9 from Taplin, 1974) with 12 replications of each of the eight forms. The arguments were contained on a slide and letters were used to stand for the propositions.

For example:

A. If there is a Y, then there is an H.
B. There is a Y.
C. There is an H.

The subjects were asked to assume that the premises were true and to evaluate the conclusion as being always correct, sometimes correct or

TABLE 9

TAPLIN'S EIGHT FORMS OF CONDITIONAL ARGUMENTS
(Taplin, 1974)

Name of argument	Conclusion	First Premise	Second Premise	Conclusion to be Evaluated
Affirming the antecedent	Affirmative	If p, then q	p	$\underline{q}$
	Negative	If p, then q	p	q
Denying the antecedent	Affirmative	If p, then q	$\overline{\underline{p}}$	$\underline{q}$
	Negative	If p, then q	p	q
Affirming the consequent	Affirmative	If p, then q	q	$\underline{p}$
	Negative	If p, then q	q	p
Denying the consequent	Affirmative	If p, then q	$\overline{\underline{q}}$	$\underline{p}$
	Negative	If p, then q	q	p

never correct. Using the material conditional interpretation of the conditional statement, the "correct" responses were tallied. Taplin's analysis showed a statistically significant effect of grade level, $F(4,291)=30.45$, $p<0.001$ and a significant effect of type of argument, $F(7,2037)=198.92$, $p<0.001$. The analysis also showed a significant GradeX Argument interaction, $F(28,2037)=3.00$, $p<0.001$.

> [P]erformance was found to improve around 6 years to college level, with the most dramatic change taking place between 11 and 15 years. (Taplin, 1974, 363)

If Taplin had stopped at this point his results would have merely confirmed the results of a number of other studies. However, Taplin pushed on, trying to find the very basis for this observed change and to infer a process or function which might be able to account for these changes. To do this Taplin makes an inference for each individual subject, based upon the subject's responses to eight conditional argument forms, of any truth functional interpretation consistent with the responses. This analysis is what sets Taplin's study apart from the others considered thus far.

First, Taplin determines whether the child was "statistically consistent" on each form of the conditional arguments, using a binomial test (0.05 level). A subject is classified as statistically consistent only if he had a preference for a certain response above the criterion level (0.05). If the subject did not reach the criterion level on any argument he was categorized as statistically inconsistent. Thus, a number of "inconsistent" subjects were eliminated from further analysis.

Next, Taplin infered a truth function for the conditional statement "if p then q" from the responses made by the child to the arguments that she was statistically consistent on:

> Basically this step of the analysis involved the inference of truth values for the truth-table representing the meaning of the connective that were compatible with the particular evaluation of the argument obtained. Different inferences apply to different arguments. (ibid., 365)

The inferences that Taplin made concerning the truth table values for each of the arguments are shown in Table 10 (from Taplin, 1974).

Taplin's criterion for constructing these truth functions was quite strict. Note from Table 10 that four estimates are made of the value for each of the truth table contingencies. If all four estimates agreed then they were deemed the truth-function for the conditional statement.

TABLE 10

INFERENCES TO BE DRAWN ABOUT THE TRUTH-TABLE VALUES
FOR A CONDITIONAL SENTENCE FROM RESPONSES MADE TO EIGHT CONDITIONAL ARGUMENTS
(Taplin, 1974)

		S's Responses to Conclusion		
Name of Argument	Conclusion	Always Correct	Sometimes Correct	Never Correct
Affirming the antecedent	Affirmative	pq(T) and $p\bar{q}$(F)	pq(T) and $p\bar{q}$(F)	either pq(F) and $p\bar{q}$(T) or pq(F) and $p\bar{q}$(F)
	Negative	pq(F) and $p\bar{q}$(T)	pq(T) and $p\bar{q}$(T)	either pq(T) and $p\bar{q}$(F) or pq(F) and $p\bar{q}$(F)
Denying the antecedent	Affirmative	$\bar{p}q$(T) and $\bar{p}\bar{q}$(F)	$\bar{p}q$(T) and $\bar{p}\bar{q}$(T)	either $\bar{p}q$(F) and $\bar{p}\bar{q}$(T) or $\bar{p}q$(F) and $\bar{p}\bar{q}$(F)
	Negative	$\bar{p}q$(F) and $\bar{p}\bar{q}$(T)	$\bar{p}q$(T) and $\bar{p}\bar{q}$(T)	either $\bar{p}q$(T) and $\bar{p}\bar{q}$(F) or $\bar{p}q$(F) and $\bar{p}\bar{q}$(F)
Affirming the consequent	Affirmative	pq(T) and $\bar{p}q$(F)	pq(T) and $\bar{p}q$(T)	either pq(F) and $\bar{p}q$(T) or pq(F) and $\bar{p}q$(F)
	Negative	pq(F) and $\bar{p}q$(T)	pq(T) and pq(T)	either pq(T) and $\bar{p}q$(F) or pq(F) and pq(F)
Denying the consequent	Affirmative	$p\bar{q}$(T) and $\bar{p}\bar{q}$(F)	$p\bar{q}$(T) and $\bar{p}\bar{q}$(T)	either $p\bar{q}$(F) and $\bar{p}\bar{q}$(T) or $p\bar{q}$(F) and pq(F)
	Negative	$p\bar{q}$(F) and $\bar{p}\bar{q}$(T)	$p\bar{q}$(T) and $\bar{p}\bar{q}$(T)	either $p\bar{q}$(T) and $\bar{p}\bar{q}$(F) or $p\bar{q}$(F) and $\bar{p}\bar{q}$(F)

If, however, there were any disagreements, the child was categorized as contradictory. This once again served to narrow down the range of subjects. The results of the above analysis are shown in Table 11 (from Taplin, 1974) and are quite revealing.

TABLE 11

PERCENTAGE OF Ss IN EACH GRADE IN EACH CLASSIFICATION
(Taplin, 1974)

Statistically consistent on all arguments

Grade	Conjunction	Biconditional	Conditional	Contradictory	Total
3	19.6	19.6	---	17.6	56.8
5	3.7	43.9	---	13.4	61.0
7	---	41.0	---	9.6	50.6
9	---	32.5	10.0	7.5	50.0
11	---	23.7	13.1	15.8	52.6

Statistically inconsistent on at least one argument

Grade	Conjunction	Biconditional	Conditional	Contradictory	Other	Total
3	7.8	9.8	---	5.9	19.6	43.1
5	2.4	14.6	---	9.8	12.2	39.0
7	3.6	9.6	2.4	14.5	19.3	49.4
9	---	10.0	12.5	12.5	15.0	50.0
11	---	10.5	13.2	15.8	7.9	47.4

Here is what an analysis of Table 11 tells us:

1. In Grade 3, no material conditional truth functions were inferred, but, equal portions of subjects interpreted the conditional argument in a conjunctive and biconditional manner.

2. In Grade 5, the conjunctive and biconditional interpretations decreased and increased significantly, respectively ($\chi^2(2)=14.09$, $p<0.001$). There were still no material conditional interpretations.

3. The Grade 7 distribution was not significantly different from the Grade 5 distribution, but the only Grade 7 conjuctive interpretation remaining was 3.6% of the group initially classified as statistically inconsistent. Aside from the 2.4% who were inferred to prefer the material conditional truth function the only truth function inferred in Grade 7 was the biconditional.

4. The Grade 9 distribution was significantly different from the Grade 7 distribution ($\chi^2(2)=8.99$, $p<0.05$). Still, the biconditional interpretation was the highest even though its actual percentage has decreased from that of Grade 7. Note that the material conditional interpretation shows 10% and 12.5% for the consistent and inconsistent groups, respectively.

5. Grade 11 shows a further decrease in the biconditional percentage and a slight increase in the percentage of material conditional interpretations. These changes are not statistically significant.

Taplin's inferential analysis seems to show that the truth functional interpretation changes with age in a manner that appears to account for observed performance changes. This intepretive progression (from conjunctive and biconditional interpretations by the third and fifth grade students, to a biconditional interpretation for the seventh graders, to a biconditional and material conditional interpretation for the ninth and eleventh grade subjects) suggests the following interpretation expressed succinctly by Staudenmayer and Bourne (1977):

> When faced with <u>if p then q</u>, the younger children link the perceptual properties, formally p and q, in an associative manner (juxtaposition according to Piaget, 1928). They cannot comprehend the possibility of relations between properties that are absent but implied ($\bar{p}$ and $\bar{q}$). As children get older they understand that the absence of propostions may also have specific relations. Thus, the first nonconjunctive interpretation that is achieved is one consistent with the [biconditional], where the absence of one property implies the absence of the other. Only with adolescence does a person achieve the capacity to understand the asymmetric truth relation between p and q implied by the [conditional statement], where the second property (q) may sometimes occur without the first (p) but the first may not occur without the second. (Staudenmayer and Bourne, 1977, 617)

Such an interpretation of Taplin's results posits a "limited capacity" Piaget-like model of cognitive development where children can only utilize a specific interpretation if they have achieved the requisite cognitive level. Thus, the theoretical distinction between formal and concrete operations may be reflected in the different interpretive strategies employed by the subject's in Taplin's study.

As for transduction, we could say that the conjunctive interpretation reflects simple transduction and that the biconditional interpretation reflects a more sophisticated transduction, symmetrical transduction. The asymmetrical nature of the material conditional interpretation of conditional statements of the form "if p then q" would seem to preclude transduction and call for a kind of reasoning structure that is qualitatively different from transduction.

E. Analysis-II

Peel's and Taplin's studies add strong evidence for alternative interpretations of the child's conditional reasoning. Taplin's (1974) study lends strong support to an explanatory theory which includes the Piagetian notion of transductive logic. Piaget (1926; 1928) indicates that he found that implicative relations were rarely used prior to age 7-8 and hardly at all before 11-12. When they are used, they juxtapose rather than causually link events. Thus, Piaget, as early as 1926 seems to suggest that the meaning of these kinds of connectives change with cognitive ontogenesis.

Consider the conjunction relation. Here one event (or proposition) cannot occur without the other also being present. This is the most primitive transductive stage. A transductive advance occurs when the link between the two events is such that not only do they occur mutually, but whenever one of them is absent, so is the other. This stage conforms to a biconditional interpretation of the conditional statement. Transduction becomes the material conditional interpretation when the second event (however defined) at times can occur in the absence of the first, but it cannot be the case that the first can occur without the second.

Thus, in one sense, the developmental changes observed by Taplin can be considered linguistic and not logical, although logic is always present. In another sense, the changes can be considered logical because the truth-functional meanings of the conditional connective does seem to change with age. We can argue that the material conditional interpetation is the more abstract, for as Taplin puts it:

> Now one must consider not only the form of the statements but also their order. One must determine not only whether the conclusion follows from the premises but also whether it is the only conclusion that can follow. Quite obviously there is a very close relationship here between linguistic and cognitive development. To say whether the linguistic change gives rise to the cognitive, or visa versa, would be merely speculation based upon the present findings. (Taplin, 1974, 372)

Staudenmayer and Bourne (1977) explore the lingusitic and cognitive explanatory alternatives using a variant of Taplin's experiment. Piaget's theory of cognitive development (and the transductive interpretation raised here) posits that the more complex conditional interpretations require the formal operations period of cognitive development. However, conventional linguistic development theory (see Brown, 1973) suggests that the progression from conjunctive to biconditional to material conditional interpretation depends upon their frequency of occurrence in the experience of the child and upon the semantic complexity of the propositions.

These two views appear to be exclusive in the sense that the linguistic interpretation of the premises does not involve the inferences needed in the evaluation of the conditional argument's conclusion. Is the linguistic interpretation of the connective primary in development or are the cognitive operations central? Staudenmayer and Bourne suggest that this formulation of the problem is incorrect.

To eliminate the assumption that two separate processes are involved in reasoning ("one to interpret the premises and the other to guide reasoning"), Staudenmayer and Bourne pose the question in this way: "What cognitive operations are necessary to make a particular interpretation?" It is their belief that in the evaluation of certain arguments, the inferences are entailed in the interpretation of the premises. All of the above reduces to this:

> [W]e are concerned with clarifying the developmental progression in the understanding of the conditional sentence. Does the progression reflect a change in preference for some interpretations over others with experience? Or, is there some lack of cognitive ability that precludes understanding the more complex relations and implications of rules like the [material conditional interpretation], even when experienced repeatedly in a learning task? (Stuadenmayer and Bourne, 1977, 617)

Staudenmayer and Bourne's experiment was the same as Taplin's (1974) except that a correct response feedback, consistent with particular interpretations (i.e., conjunctive, biconditional, material conditional), was provided for each trial. Subjects were chosen from grades 3, 6 and 9 with no control groups. Can subjects learn interpretations that are different from those that were shown to be prefered in the Taplin (1974) study? Note that no teaching of logical principles is involved here, only the exposure to "correct" interpretations after response trials. Staudenmayer and Bourne found that the youngest children matched the propositions or properties that appeared in the first premise to the properties that appeared in the second premise and conclusion. This type of juxtaposition was found to decrease with increasing age.

As a result of their study, Staudenmayer and Bourne concluded that:

> It seems likely that the youngest children used some form of an associative or rote strategy in the reasoning task, similar to Evan's (1972) matching bias, making it impossible to keep the asymmetry between the propostions straight for the [material conditional interpretation].
> On the other hand, some of the older adolescents in the reasoning task were probably using a different strategy characteristic of formal, inferential reasoning since some of them learned the [material conditional interpretation]. (ibid., 622)

Their study and that of Taplin's supports a transductive interpretation of the conditional statement at the pre-adolescent age and adds to the notion of a cognitive developmentally significant factor in differential interpretations of conditional statements.

F. The Research-III

1. Paris (1973)

This study, although not quite as definitive as the previous one, adds further evidence to the biconditional interpretation of the conditional statement and argument. Paris's sample are equal numbers of white middle-class male and female elementary and secondary school students and college freshmen (40 subjects in grades 2, 5, 8, 11 and college).

Eight linguistic connectives were examined: "and," "but," "both-and," "neither-nor," "or," "either-or," "if-then," "if and only if-then." We are concerned only with the latter two connectives. For each connective, a descriptive propositional sentence was read and a slide showing pairs of propositions was simultaneously displayed. These slides represented the attributes or truth forms of the verbal propositional sentences. See Table 12.

TABLE 12

LINGUISTIC CONNECTIVES, LOGICAL RELATIONSHIPS
AND ATTRIBUTES
(Partial from Paris, 1973)

Linguistic Connective	Logical Relationship	Attributes (truth forms) AB	$A\bar{B}$	$\bar{A}B$	$\bar{A}\bar{B}$
If-then	Conditional	T	F	T	T
If and only if-then	Biconditional	T	F	F	T

The propositional test sentences consist of compound sentences formed from two of eight simple sentences used for the testing, and a connective. Each picture proposition also had its complement. For instance, one picture might show a boy riding a bicycle. Its complement showed a boy standing next to his bicycle. There were four test slide-sentence pairs for each connective and the subjects had to judge the test sentence as true or false depending upon the value of its components determined by the slide. Table 13 shows the "error" patterns that are associated with each of the truth attributes, by grade for the conditional connective.

TABLE 13

PERCENT ERRORS ON THE CONDITIONAL CONNECTIVE IF-THEN
ACCORDING TO TRUTH FORM AND GRADE
(Paris, 1973)

Grade	TT	TF	FT	FF
2	0	17.5	75	92.5
5	0	2.5	95	72.5
8	5	0	100	42.5
11	17.5	0	100	40
C	17.5	2.5	95	45

Paris found no age related differences in these error scores and states:

> The errors suggest that Ss consistently employed erroneous, nonlogical processing strategies. (Paris, 1973, 286)

Let us look at Table 13 from a different perspective. Since the subjects were required to answer true or false to the test sentence, the TT column of Table 13 indicates the percentage of those subjects who answered false when they should have answered true. Column TF indicates those subjects who answered true when they should have answered false. Column FT are those who answered false when they should have answered true and column FF shows the percentage of those subjects who answered false when they should have answered true. All this according to the material conditional interpretation of the conditional statement.

From the data, I have determined the percentage of correct responses under the biconditional interpretation of the conditional connective. Table 14 shows the percentage in Table 13 that answered "correctly" according to the biconditional interpretation.

TABLE 14

PERCENTAGE FROM TABLE 13 ANSWERING CORRECTLY ACCORDING TO A BICONDITIONAL INTERPRETATION OF THE CONDITIONAL IF-THEN STATEMENT

Grade	TT	TF	FT	FF
2	100	82.5	75	7.5
5	100	97.5	95	27.5
8	95	100	100	57.5
11	92.5	100	100	60
C	82.5	97.5	95	55
Total	94	95.5	93	41.5

The key column, FT, that distinguishes the biconditional from the material conditional appears to have been answered in an increasingly correct manner progressing from second to eleventh grade. Column FF remains inexplicitly modest but increasing in correct responses through the eleventh grade. Although no significant test has been performed upon Table 14 a trend is obvious: the biconditional interpretation grows stronger with age. Table 15 shows the percentage of errors on the Paris biconditional connective by grade and truth attribute form.

TABLE 15

PERCENT OF ERRORS AND CORRECT ANSWERS (IN PARENTHESES) ON THE BICONDITIONAL CONNECTIVE (Paris, 1973)

Grade	TT	TF	FT	FF
2	10 (90)	5 (95)	17.5 (82.5)	90 (10)
5	5 (95)	7.5 (92.5)	7.5 (92.5)	72.5 (27.5)
8	15 (85)	2.5 (97.5)	10 (90)	50 (50)
11	10 (90)	2.5 (97.5)	0 (100)	42.5 (57.5)
C	20 (80)	0 (100)	0 (100)	50 (50)
Total	12 (88)	3.5 (96.5)	7.5 (92.5)	61 (39)

Paris's analysis of the errors indicates that a trend for above-chance errors declines to chance levels with increasing age and is visible for the FF items. It can be seen by comparing Tables 14 and 15 that the subjects responded similarly to the conditional and biconditional statements. There is some evidence that indicates that a biconditional interpretation of the conditional statement increases somewhat with age.

Paris notes that younger children "consistently" responded correctly to TT statements and "false" to all others. He attributes this in part to a "mismatch" strategy which posits that a mismatch between propositions always yields a "false" response. This explanation, however, does not account for the FF responses, although Paris notes that while younger children viewed FF as false, older subjects gave responses at chance levels.

This interpretation of FF seems to support some research into adult conditional reasoning. In that research, the FF form was judged to be irrelevant or ambiguous for the comprehension of the conditional statement (Wason and Johnson-Laird, 1972).

Finally, Paris attempts to explain the lack of correspondence of his results to the material conditional interpretation of the conditional statement by a causal-semantic notion. Here, responses to both the conditional and biconditional propositions are to be interpreted as judgments of "logical consequence," i.e., a necessary relationship exists between the propositions' antecedent and consequence. This relationship is determined by the semantics of the sentence. Thus:

> The truth-functional interpretation of material implication simply does not adequately represent Ss' comprehension process because it explicitly ignores the semantic relationship betwen the elements within the proposition. (Paris, 1973, 290)

This Paris claims, counts against the explanatory value of formal logical models which "deny" the necessity of meaningful relationships between components of propositions.

2. Roberge and Paulus (1971), Roberge (1970)

This and the subsequent three studies are patterned after or replicate Ennis's (1965; 1969) work. This is a study that uses class and conditional reasoning tests with 263 students from fourth, sixth, eighth and tenth grades. I examine only the results of the conditional reasoning test which consisted of six conditional principles and three content dimensions. The first five principles and three content dimensions are the same as Ennis (1965). All of the items are in the form:

Suppose you know that
    Premise-1
    Premise-2
Then would this be true?
    Conclusion

A. YES   B. NO   C. MAYBE

The responses had the same meaning as Ennis (1965). There were 12 items for each principle and total test and principle subtest scores were calculated by the number of correct responses, according to the material conditional interpretation of the conditional statement. Mastery of a principle was defined as correctly answering at least eight of the twelve principle items. This was deemed to be a sufficient condition for concluding that a subject had mastered that particular principle.

The results of an analysis of variance for grade level and principles confirm Ennis's earlier results. There were significant differences ($F=487.81$, $p<0.001$) among the mean principle subtest scores and a significant interaction ($F=12.59$, $p<0.001$) between grade level and principles.

At each grade and for all grades combined, the Forward Conditional was significantly easier than each of the other principles. The only exceptions were in the eighth and tenth grades where the Forward Conditional was easier but not significantly easier than the Transitivity principle ($p>.05$). The fallacy principle means and the validity principle means were significantly different ($p<.01$) with the fallacy principles more difficult in all cases.

Finally, no significant differences between the principle subtest means existed between the fourth and sixth grades. The validity principles did significantly differ ($p<.01$) between eighth and tenth grades but there was no significant difference ($p>.05$) between the fallacy principles. However, between the eighth and tenth grades, the fallacy principles differed significantly ($p<.01$) favoring the tenth graders. Thus, growth in validity principle understanding reached its "peak" at about the eighth grade and growth in fallacy principle understanding showed a "notable spurt" between the eighth and the tenth grades; all this without any formal instruction in the principles of conditional logic.

The study found that the concrete familiar dimension was the easiest of the three content dimensions and that the abstract and suggestive dimensions were equally difficult. The means of the concrete and each of the other two dimensions were significantly different ($p<.01$). Grade level and content dimension interaction were not signficant ($p>.05$). There were no significant differences found for sex and its interaction with the other factors. In general, the findings of this study tended to confirm the results of the Ennis (1965) study.

3. Berzonsky and Ondrako (1974)

The authors conducted two studies to analyse the role of cognitive style, measured by the Matching Familiar Figures test (MFF) developed by Kagan (1964), and deductive reasoning.

> Cognitive style or reflectivity-impulsivity (R-I) ... refers to an individuals' tendency to pause and reflect over possible solutions in a problem solving situation. (Berzonsky and Ondrako, 1974)

The measure of deductive ability used was Ennis's (1969) chemical test, which was administered to sixty-nine 6 and 7 year olds. Only three logical principles were tested for: Conversion, Inversion and Contrapositon. One of the objectives of the study was to replicate Ennis's findings.

Item content, suppositional and factual dimensions showed no significant differences (t=.98, df, 67) which is at variance with Ennis's findings that the suppositional items were more difficult. There was, however, a highly significant principle effect (F=45.70, df, 2136, p<.01). All of the mean principle differences were statistically significant at the .01 level: Conversion mean=1.40; Inversion mean=2.07; Contraposition mean=3.42. This is in agreement with Ennis's (1969) findings. Also, no sex differences were found. The study did not show a significant difference between R-I cognitive styles with respect to the logial principles.

The second study was similar to the first but used the Paulus-Roberge Conditional Reasoning Test (Roberge and Paulus, 1971) to test for six logical principles and three content dimensions (concrete, abstract and suggestive). The subjects were 39 eleven year old students with mean age 11 years 5 months. This test has already been described.

This study found no sex differences, which confirmed the results of Roberge and Paulus (1971). It also found that content dimension accounted for significant differences (F=4.22, df, 2, 76, p<.05) in the subject's logic test scores. The suggestive items were found to be significantly more difficult than the concrete or the abstract ones. This result is at variance with Roberge and Paulus's general conclusion but is consistent with the sixth graders considered alone. Also, the Berzonsky and Ondrako study shows that the logical principles accounted for significant differences (F=16.05, df, 5, 190, p<.01) in the test scores.

The Forward Conditional principle was the easiest. The fallacy principles (Conversion and Inversion) were more difficult than the others. There were no significant differences found between the other two principles. Ennis (1965) found that Conversion was more difficult than Inversion. This study and Roberge (1971) found this not to be the case. The results of the two studies reported here tend, for the most part, to confirm previous experimental results. In addition, this second study of Berzonsky and Ondrako did not find a significant difference between R-I cognitive styles with respect to the logical principles.

4. Eisenberg and McGinty (1974)

This study tried to determine whether maturation enables a person to reason "more logically" from worded premises. A test of 30 sentential logic items was constructed to measure five logical principles (6 items per principle). We will examine only the three conditional principles: Forward Conditional (I), Conversion (II) and Contrapositon (III). The test was similar to Ennis's (1965) test. The form of the items was:

Premise-1
Premise-2
Conclusion?

A. YES B. NO C. MAYBE

Premise-1 is an if-then conditional statement. Two of the six questions for each principle have both antecedent and consequent affirmatively worded (A). Two questions have one part affirmatively worded and the other part negatively worded (A&N). The remaining two questions have both antecedent and consequent negatively worded (N). Examples appear in Table 16.

TABLE 16

EXAMPLE QUESTIONS

Principle	Type	Example
I	A	If John is big, then Jane is big. John is big. Is Jane big?
II	A&N	If it is a black cat, then it is not an old cat. It is not an old cat. Is it a black cat?
III	N	If there is not a red wagon, there is not a blue bike. There is a blue bike. Is there a red wagon?

The subjects were 50 second grade, 50 third grade children, 89 college elementary education majors in the first (MA 150) of a two course sequence in mathematics and 65 elementary education majors in the second sequence (MA 151) of the course. A unit in logic was not part of this two sequence course. All test questions were read to the subjects who had a copy of the test in front of them. Table 17 summarizes the significant differences between the school children and the prospective teachers on the three logial principles tested.

TABLE 17

SIGNIFICANT DIFFERENCES BETWEEN CHILDREN AND COLLEGE STUDENTS ON DIFFERENT LOGICAL PRINCIPLES

	Third Grade	MA 150	MA 151
Second Grade	----	I, II, III	I, II, III
Third Grade		I, III	I, II, III
MA 150			----

Note: (----) means not statistically significant at the .05 level. The presence of a principle indicates significance at the .05 level.

All significant differences favored the pre-teachers. Note that there is no significant difference between the grade school children and the pre-teachers on the Conversion principle (a fallacy item). Table 18 shows the results of the semantic form of the test items.

TABLE 18

PERCENT CORRECT RESPONSES BY PRINCIPLE, SEMANTIC FORM AND GRADE

Principle

	I				II				III			
	Grade 2	Grade 3	MA 150	MA 151	Grade 2	Grade 3	MA 150	MA 151	Grade 2	Grade 3	MA 150	MA 151
A	48	38	88	91	16	19	22	32	36	52	69	65
A&N	63	56	95	97	23	31	30	37	39	42	67	72
N	34	49	90	89	14	22	35	41	26	35	49	52

Maturation seems to be important only in Items I and III. Both grade schoolers and pre-teaches tend to make similar kinds of logical errors <u>within</u> different categories.

Here and in the previous three studies, the material conditional interpretation was utilized as the standard measure of correct reasoning. It is not possible, given the data reported in these studies, to perform an alternative analysis using the biconditional as the standard measure. It should be noted, however, that the fallacy items seem to be more difficult than the validity items, and on the whole, tend to elicit improved performance with age.

5. Howell (1967)

This study examined subject recognition of selected inference patterns of conditional reasoning in the absence of explicit training in logic. Five of the ten patterns correspond to the inference patterns we have previously examined: Forward Conditional, Inversion, Conversion, and Contraposition. Howell's sample consisted of 164 secondary level students (grades 7, 8 and 9), enrolled in an accelerated college prepatory mathematics program who had no prior training or instruction in logic. He created a measure to test whether his subjects could "recognize" the correct use of certain patterns of inferential reasoning and whether subjects in one grade "out-perform" those in any other grade.

The test was divided into three parts. Parts I and II tested familiar and abstract content, respectively. Part III used abstract symbols (as in Part II), but substituted De'Morgan's Law equivalences for "and" and "or" propositions. The types of arguments we are interested in were unaffected. Two items for each inference pattern per test part gave a total of 6 items per pattern of inference. Mastery of an inference pattern required 5 or 6 correct answers. The conclusions of each pattern were randomly chosen to follow logically or not follow logically. No attempt was made to vary the negation of premises. Students in the study could answer "YES" or answer "NO" when asked whether the conclusion followed from the premises (which were to be assumed true).

Howell found that over 90% of the subjects in each grade mastered the Forward Conditional principle; 39.6, 40.4 and 77.8 percent mastered

Contraposition in grades 7, 8 and 9, respectively; no one mastered the fallacy principles in any of the grades.

If we use the criteria that if two-thirds of any one grade mastered a pattern, we see that Howell's results showed that students in each grade understood Transitivity and the Forward Conditional and only grade nine understood Contraposition as well. From his results, Howell concluded that:

> The students were much more successful at recognizing valid inference pattern than invalid ones. In fact, not one student correctly responded to five or six of the six occurrences of either of the invalid patterns. (Howell, 1967, 82)

Finally, Howell found that the ability to infer valid conclusions increased slightly by grade and that there was no significant effect of sex.

6. Gardiner (1966)

This study tested 277 subjects in grades 4 through 13, on 14 rules of logic comprising 6 different content areas. The subjects had no previous logic instruction and no demographic information was reported. We are concerned here with only 4 logic principles which were worded in a fashion similar to Ennis (1965). Mastery ("understanding") of a logical principle is defined operationally as scoring 5 or 6 correct out of the 6 items per principle. A subject had a "misunderstanding" of a principle when she makes the same response to at least 5 out of 6 items within a set that is keyed incorrect. Table 19 shows the percent of subjects in each grade group who have mastered each principle. Grades were collapsed in pairs to increase the number of subjects within each group.

TABLE 19

PERCENT UNDERSTANDING AND NON-UNDERSTANDING OF PRINCIPLES

	Conditional Principle											
Grade	Forward Conditional			Inversion			Conversion			Contraposition		
	A	B	C	A	B	C	A	B	C	A	B	C
4&5	59	0	2	4	24	2	43	0	0	8	16	0
6&7	79	0	2	0	21	2	45	0	0	4	23	0
8&9	89	0	0	2	39	4	63	0	4	0	48	0
10&11	99	0	0	0	28	20	43	0	30	0	47	9
12&13	100	0	0	0	19	31	26	0	38	0	45	5
Mean Correct	86					12			15		36	

Note: A="YES," B="NO," C="MAYBE" answer. Underline indicates correct understanding. Numbers not underlined indicate misunderstandings

Note the great difficulty that the subjects have in mastering the fallacy principles. If we use the biconditional as our criterion standard, then columns B and A under Inversion and Conversion, respectively, represent "understanding." Now, with the exception of grades 12 and 13, the biconditional interpretation score is greater than the material conditional interpretation score and decreases with increasing grade after grades 8 and 9.

7. Paulus (1967)

This study measures differential performances between deducing (drawing) and assessing conclusions of conditional arguments. We look at the assessing test, which is similar to Ennis (1965) and where mastery is set at at least 7 items correct (out of 10) for each principle.

A total of 165 subjects in grades 5, 7, 9 and 11 took part in the study. None had any previous training in logic. Paulus distinguished between four different kinds of content: concrete familiar, abstract (a combination of unfamiliar and symbolic), suggestive and negative. All but negative are as in Ennis (1965). Negative content simply refers to negatively worded premises. Subjects could answer either "YES," "NO" or "MAYBE" when assessing the conclusions of the conditional test arguments.

While Paulus did not tablulate mastery by grade level, he did find that subjects on the whole did better on the validity principles than on the fallacy principles. Scores on the validity principles were highly correlated among themselves and near zero or negatively correlated with the fallacy principles. In addition, the fallacy principles were found to be higly correlated with each other. Paulus concludes that scores on validity and fallacy principles were "relatively independent" of one another. Even when the "MAYBE" scores were statistically partialled out of each principle test score, the fallacy principles still correlated highly with each other although they no longer tended to correlate negatively with the validity principles.

When grades were combined, chronological age correlated positively and significantly with principle subtest scores (with the exception of Contraposition). The type of content interacted significantly with mental age, and negative terms in the premise did not significantly affect scores.

Paulus's data warrants the rejection of his null hypothesis that there is no change of deductive reasoning abilities as children grow older. Mental age, not chronological age is his measure.

When grades were combined, chronological age correlated positively and significantly with principle subtest scores, with the exception of Contrapositon.

8. Martens (1967)

This study used as subjects high school seniors in college prepatory courses. Ten inference patterns were examined of which only four are of interest here. None of the premises in the test arguments contained artificial vocabulary or nonsense syllables. None were contrary to fact and valid conclusions did not contradict common sense or previous school instruction. Finally, no controversial materials were employed in the test items.

There were four questions for each logical principle. The subject could choose from amongst three answers: the valid conclusion, the negation of the valid conclusion (if the correct answer is not a "MAYBE" response), and the equivalent of the "MAYBE" response, stating that neither the first nor the second answers are valid conclusions drawn from the premises. Significant differences were found to exist ($p<.01$) among the mean scores of the principles. Inversion differed significantly from Contraposition, favoring the latter, and Conversion differed significantly from Contraposition, favoring the latter principle. In addition, mean scores for each valid inference pattern (considering all 10 principles) differ significantly from the mean score of each of the fallacy principles.

The results indicated that only 26% and 15% of the subjects mastered (three or four correct) Inversion and Conversion, respectively. However, 87% mastered Contraposition.

In analyzing the incorrect responses for the fallacy principles, Martens found that a majority of the subjects completed the fallacious inferences of denying the antecedent and affirming the consequence. For Inversion, 26% chose the correct response 3 or more times and 59% chose an incorrect response three or more times. Of those who chose incorrect responses, they chose the response that denied the consequent in a ratio of 5 to 1. For the Conversion principle, only 15% chose the correct response 3 or more times, whereas, 70% chose an incorrect response 3 or more times. Of those choosing incorrect responses, they chose the response that affirmed the antecedent in a ratio of 35 to 2. Martens did not test for the Forward Conditional principle.

9. Miller (1968)

In this study, 100 8th, 10th and 12th graders were tested on 6 principles of inference including the four principles of conditional logic that we are interested in. Of the twenty items, 5 were in symbolic form and 15 were in verbal form. Of the 15 verbal items, 5 had premises agreeing with physical world situations, 5 had premises that violated physical world situations and 5 had premises that were nonsense statements. All of the statements involved in each item were deemed unambiguous, unemotional, short and neutral, and were in a positive form.

The subjects had three choices for answers. One choice was the valid conclusion in the valid patterns or the usually accepted conclusions (i.e., the conclusions that completed the fallacies of affirming the consequence and denying the antecedent). Another choice was the negation of the above choice and the third choice was that neither of the above choices "necessarily follows." Table 20 shows mean scores for principles by grade.

TABLE 20

MEANS AND STANDARD DEVIATION (PARENTHESES)
BY GRADE AND PRINCIPLE

	Principle			
Grade	Forward Conditional	Inversion	Conversion	Contra-position
8	16.25 (4.48)	1.80 (3.08)	1.96 (2.90)	14.84 (4.57)
10	18.17 (3.08)	1.56 (3.16)	2.25 (4.20)	16.50 (4.25)
12	18.14 (2.82)	2.61 (4.66)	3.39 (5.20)	15.70 (4.54)

For the fallacy items, Miller found that the

> ... results of the analysis of the correct responses to invalid patterns indicate that for the vast majority of the students, the analysis was performed on a random error response. (Miller, 1968, 137)

Miller reports the means and standard deviations for the fallacy items which were responded to in such a fashion as to commit the appropriate fallacies. These results are shown in Table 21. Note that these results, together with the results for the validity principles (Table 20) constitute strong evidence for a biconditional interpretation of the conditional statement.

TABLE 21

MEANS AND STANDARD DEVIATIONS (IN PARENTHESES)
BY GRADE

	Fallacies	
Grade	Affirming the Consequence	Denying the Antecedent
8	15.99 (3.84)	15.43 (4.51)
10	16.45 (4.47)	16.53 (4.24)
12	15.35 (5.55)	15.93 (5.03)

These results lead Miller to conclude that:

> Students in the eighth, tenth and twelth grades select the usual error responses to the two invalid patterns to about the same extent as they select the valid responses to the valid patterns.... (ibid., 137)

In conclusion, Miller states that:

> ... the vast majority of the students accepted the invalid patterns as valid. This happened because the students were using an inadequate criterion for judging the validity of a pattern. (ibid., 144)

While this criterion may be "inadequate" in the context of the material conditional interpretation of the conditional statement, it may be appropriate in a biconditional interpretation context.

10. McAloon (1969)

This study explores the role of logic in elementary school mathematics. It examines the teaching of certain principles of "logical thinking" and their relationship to learning certain aspects of mathematics in grades three and six.

McAloon utilized four treatment groups:

> LM, in which logic was taught interwoven with mathematics; L, in which logic was taught separated from any mathematical application; I, in which teachers of this group received in-service instruction in mathematics; and C, a control group which received no instruction in logic. Teachers in LM and L received in-service instruction in logic. Treatment I was designed to help isolate the factor of in-service training. Teachers in C received no in-service training and students in this group were taught no logic. (McAloon, 1969, 159)

The subjects were 511 students in 13 third grade classes and 589 students in 13 sixth grade classes in white middle-class schools. The logic instruction consisted of class and conditional logic, particularly Forward Conditional (Modus Ponens), Inversion and Negation. All experimental subjects received 16 hours of instruction. McAloon states that not much "implication" was taught to the third graders and only a little more was taught to the sixth graders. A modified version of the Ennis (1965) test was used along with a measure of mathematics achievement and a final logic and logic-retention test. We examine the modified conditional reasoning test results.

McAloon used the first seven principles in Ennis's test but chose only three items for each principle (mixed between the content dimensions) and found that the fallacy principles were considerably more difficult than the validity principles (unfortunately, percentages correct by grade, principle and experimental group were not reported). McAloon found that the groups taught logic scored significantly higher ($p<.01$) than the other groups on both the conditional and class reasoning tests using corresponding pretests or IQ scores as covariates. Although the LM group scored higher than the L group in the third grade and although the L group scored higher than the LM group in the sixth grade, on both reasoning tests, the differences were not significant using pretest IQ scores as covariates.

McAloon found that there was a significant difference ($p<.01$) between the scores of the third graders and sixth graders on both class and conditional pre and post tests. Also, no sex differences were found. Clearly, both grade levels found the validity principles much easier than the fallacy principles (given a material conditional scoring scheme).

11. Carroll (1971)

Carroll examined the effects of teaching logic on the conditional reasoning abilities of 9th grade low-achievers in mathematics (below the 15th percentile in mathematics and above the 25th percentile in reading). The conditional principles tested were: Forward Conditional, Conversion, Inversion, and Conversion. They were tested along four content dimensions: concrete familiar, symbolic, misleading and removed from reality. The first two dimensions are described in Ennis (1965). In the misleading dimension, the correct response is contrary to fact and in the removed from reality dimension, the subject is asked to imagine himself or herself in a fictitious situation:

> The stranger says: "If horses are animals, the horses are green."
> Horses are not green.
> Then: ...... (Carroll, 1971, 26)

Subjects were presented with three answers: a conclusion, the conclusion's negation, and "You aren't told enough to be certain whether or not ......" The blank is filled in by the first conclusion, above. The experimental group consisted of 48 9th grade students who were given six lessons in conditional logic. An alternate treatment group (of 23 subjects) received six lessons of instruction in probability. The control group of 119 students received no special instruction. There were twelve questions for each principle and three questions for each of the four content dimensions.

Carroll's results show that on the overall test scores, the experimental group where logic was taught showed no more improvement between pre and post tests than did the other groups, although average changes were higher for the experimental group (the percent of subjects showing improvement was subjected to Chi-squared tests of significance, .01 level). The same tests of improvements were performed for the conditional principles with only the Conversion principle yielding significant results, although the trends favored the experimental group on Forward Conditional and on Inversion. The control group was favored on the Contraposition principle.

The trend on the content dimensions favored the experimental group, although only contrasts on the misleading category were significant. Carroll's subjects had much more difficulty on the fallacy than on the validity items.

Except for the males on the Contraposition principle, 9th graders have great difficulty with the fallacy as opposed to the validity principles. One explanation offered by Carroll for this apparent anomaly is:

> [S]tudents, in accepting the invalidity of the inverse, become more reluctant to accept the contrapositive as valid .... (ibid., 114)

Recall that Ennis (1965) gave a similar explanation for his similar results.

12. Ryoti (1973)

Ryoti makes a distinction between the "standard" conditional and the "generalized" conditional. The standard conditional is of the form, "if p, then q," where "p" and "q" are particular propositions. The generalized conditional is of the form, "if p(x), then q(x)," where "p(x)" and "q(x)" are class statements (or propositional functions). Compare Examples 4 and 5 from Ryoti (1973, 8).

If Pat is Joan's friend then she is going to the museum today.	(if p then q)
Pat is Joan's friend.	(p)
Then, would this be true?	
Pat is going to the museum today.	(q)

Example 4. Standard Conditional Argument

If a person is Joan's friend, then he is going to the museum today.	(if p(x), then q(x)
Pat is Joan's friend.	(p(m))
Then, would this be true?	
Pat is going to the museum today.	(q(m))

Example 5. Generalized Conditional Argument

Ryoti compared 12 principles of class reasoning with 12 isomorphic principles of conditional reasoning. The subjects were 81 fourth graders and 67 ninth graders and his test consisted of 6 items for each principle. His results for the generalized conditional principles of

Forward Conditional (Modus Ponens), Inversion, Conversion and Contraposition showed that both grade levels had more difficulty with the fallacy principles than with the validity principles. Subjects did better on Inversion than on Conversion and better on the Forward Conditional than on Contraposition.

13. Flener (1974)

Let us call the standard conditional, SS, for single statement and call the generalized conditional (which contains one generalized and one standardized statement), GS. Ryoti's (1973) study examined the GS type argument. Flener adds a third category of conditional, one where both premises are in generalized form, GG. The test items all have familiar content (Ennis, 1965). Subjects can answwer "YES," "NO" or "MAYBE" to the stated conclusions. These conclusions are of two types: an affirmative conclusion or a negative conclusion. Examples 4 and 5 illustrate the differences between the two types. Flener tests for Modus Ponens, Contraposition, Inversion and Conversion.

$p \longrightarrow q$

p

$\therefore q$

Example 4. Modus Ponens With Affirmative Conclusion

$p \longrightarrow q$

p

$\therefore \bar{q}$ is false

Example 5. Modus Ponens With Negated Conclusion

Essentially, Flener is interested in three variables comprising the validity and fallacy principles:

1. The number of valid inferences correctly recognized,
2. The number of fallacies correctly recognized.
3. The total number of items correctly recognized.

Three test items were constructed for each of the four validity principles tested and two items each for the four fallacy principle inference patterns were constructed. Thus, for each of the three categories (SS, GS and GG) a test of 20 questions is constructed. The three tests were then administered to 306 fifth, 305 seventh and 305 ninth grade students. Over the three conditional argument categories, the Inversion patterns were the most difficult. Of the 27 correlations between pattern subscores, 23 correlations between valid and invalid inferences were negative. All of these in the fifth grade were significantly negative. Only 3 correlations were significantly negative in the seventh grade and none of the correlations between valid and invalid subscores were significant in the ninth grade. Flener states that:

> The results imply that in the fifth grade most of the subjects who were above the mean for the subscores measuring valid inferences tended to be below the mean for the subscores measuring invalid inferences. In the other two grades the near-zero correlation indicated that this trend still continued for many of the students, although it was not as prevalent. (Flener, 1974, 69)

An analysis of the total number of correct responses for each of the submeasures revealed that there was no apparant pattern to the subject's responses to the fallacy items. Flener notes that if the students were randomly guessing, the distribution of the number of correct responses would be binomial. He concludes that:

> There are apparently other psychological factors which influence the student's choices. An investigation into reasons for the student's choices might give an explanation for this behavior, and it might also explain why the correlations between the valid and invalid submeasures are negative. (ibid., 76)

Flener's study is a source of great enlightenment, although it is limited by the small number of questions for each principle. Flener found that there were significant differences for grade levels and test categories for all three of the variables discussed above, although none of the interactive effects were statistically significant. The results showed that for each of the three variables, the ninth graders scored significantly higher than the seventh or fifth graders and the seventh graders scored significantly higher than the fifth graders. Of particular interest, however, was that the absolute increase was much greater between the fifth and seventh grades than between the seventh and ninth grades for the validity items subscore (1.10 as compared with 0.44). There was also a much higher increase between seventh and ninth grade scores than between the fifth and seventh grade scores on the fallacy item subscores (1.03 as compared with 0.32).

For all of the three variables, the scores on the GG test were the lowest and the scores on the GS test were the highest. These results were statistically significant and were observed at each grade level. Considering only SS and GG, it was found that subjects were

> ... better able to recognize both valid and invalid inferences where the conditional premise was a generalization rather than a singular statement. Furthermore, when comparing absolute differences, the difference between Test GG and Test SS was greater than the difference between Test SS and Test GS when measuring the valid inferences (0.66 as compared with 0.09), but the difference between Test SS and Test GG was greater than the difference between Test SS and Test GG for the subscores using the invalid inferences (0.32 as compared with 0.13). (ibid., 136-7)

Thus, for the valid inferences, there was little difference between the singular or the general conditional principles as long as the second premise in the item was singular. When the second premise was a general proposition, however, the subject's found the items to be much more difficult. For the fallacy patterns, the GG subtest was the most difficult and the GS subtest was the easiest for all three grades.

Flener was able to show that mental ability was not a factor in the subjects' responses to the valid and invalid arguments in the fifth grade. Mental ability was a factor in the seventh grade for the validity principles with the upper ability group having significantly higher scores than the lower and middle groups. The middle group had significantly higher scores than the lower group.

In the ninth grade, for each of the three variables, there were significant differences between each ability level, with the upper ability group having the highest mean and the low ability group having the lowest mean. Thus, by the ninth grade, there appeared a significant relationship between subjects' abilities and awareness of valid and invalid arguments.

Flener notes two trends in his results:

> For the valid inferences, although each was above the low ability group there was no significant difference between the middle and high ability groups in the fifth grade. However, the three groups were significantly different in both the seventh and ninth grades. For the invalid inferences there were no significant differences in the fifth grade between any of the groups. By seventh grade the two upper level groups had higher scores than the low ability group using the invalid inferences. (ibid., 141)

14. <u>Kodrof and Roberge (1975)</u>

Twelve children (six males and six females) each from grades 1, 2, and 3 (36 subjects total) were given a twelve item test containing six Forward Conditional items and six Contraposition items. The subjects were all white lower-middle-class pupils in a suburban New Jersey public school system. Grades 1, 2 and 3 had mean Lorge-Thorndike IQ scores of 111.73, 107.50 and 109.44, respectively. These means were not significantly different.

The object of the study was to determine the effects of the content-relatedness of test questions and the mode of test presentation on the subjects' scores on the two valid inference principles. This was accomplished by analyzing their correct judgments and explanations of their answers.

For each of the two inference principles there were three related-content items (i.e., "obvious real-life association between the nouns in the antecedent and the consequent of the hypothetical premeise") and three unrelated content items (i.e., "minimal real-life associations between the nouns in the two parts of the hypothetical premise"). Explanations for correct judgments were sorted into categories (see Table 22) and the subjects were informed about the correctness of each item they answered.

Subjects were tested using a concrete materials version of the test (a stimulus box) and a verbal form of the test (orally). These two test versions were separated by a two week interval and randomly mixed in order of presentation. A split-plot repeated measures statistical analysis design was employed.

None of the means of the principles by related-content type by mode of presentation were above 3.0 and most were below 2.5 (out of a possible 6). Using Ennis's (1965) criterion for mastery, these pupils did not do very well even with these validity principles. Significant main effects were found for grade level ($F(2,33)=3.54$, $p<.05$) and post hoc comparisons showed a significant difference ($p<.05$) between the means of the first and third graders. Significant main effects were found for mode of presentation ($F(1,33)=9.22$, $p<.01$) which favored the concrete presentation form. The subjects found the Modus Ponens items

TABLE 22

EXPLANATION CATEGORIES
(condensed from Kodrof and Roberge, 1975)

1. Verbalization deficits, in which the subject demonstrates an inability to describe the thinking which led to his response.

2. Memory deficits, in which the subject erroneously repeats a given premise or loses his train of thought.

3. Arbitrary explanations, in which reasoning is not based on facts given in the problem.

4. Structural deficits, in which the subject usually gives evidence that he is aware, or becoming aware, of the structure of conditional logic, but his response is incomplete and cannot be classified as adequate conditional reasoning.

5. Conditional reasoning, in which an understanding of conditional reasoning is suggested by the repetition of the hypothetical premise or both premises. The subject's response must include "if" or "if ... then."

6. Patterened explanations, in which the subject attempts to organize the information presented to him by establishing a pattern. The subject must refer to at least three terms in the test item for a response to be coded as patterned.

significantly easier than Modus Tollens items ($F(1,33)=19.12$, $p<.01$).

An examination of the simple main effects for the interaction between grade level and principles showed that third graders scored significantly higher ($p<.01$) than first graders on Contraposition items and that first graders found the Forward Conditional items significantly easier ($p<.01$) than the Contraposition items.

The data suggested that exposure to the concrete materials test prior to the verbal test improved the latter scores. But taking the verbal test first did not improve subsequent concrete test scores.

An examination of the frequencies of explanation type for each of the logical principles reveals some interesting results. Even though many of the children made correct judgments only a few of their explanations indicated a "conditional reasoning" explanation category. In fact, in all of the principles by grades (except Forward Conditional, grade 2), the conditional reasoning explanation frequency was the lowest. In all of the principle by grades the "patterned explanation" frequencies were by far the highest. This is an important result and can be interpreted to support a transductive reasoning type explanation for logical performance on this measure. Also, there were "appreciable" increases in "structural defects" frequencies for Contraposition items by grade. Finally, there were negligible differences in types of explanation by the related and unrelated items.

Another important finding was that there were no significant differences between the related and unrelated content means and the similarities in the types of explanations for each type of content. Other logical principles were not tested for.

15. Antonok and Roberge (1978)

The authors sought to determine the effects of inference principles and type of content upon the ability of educable mentally retarded (EMR) children to judge conclusions of conditional arguments. This, and another study by Tucker (1971) are the only ones that I know of that have tested for conditional reasoning ability in EMR children. Tucker (1971) showed no growth in propositional reasoning ability across ages 6 and 8 when administered Hill's (1961) test orally.

Here, a pictorial conditional reasoning test was used on 6 male and 6 female subjects for each of three age groups (mean chronological ages of 11.72, 13.94 and 16.85 years with corresponding mental ages of 8.46, 9.60 and 12.08 years, respectively), testing for Forward Conditional, Contraposition and Conversion along four content types.

The test items were presented in pictorial format (simple line drawings) to control for language and memory deficits. Answers of "YES," "NO" or "MAYBE" were required as well as an explanation of the judgment. These explanations were sorted into the categories in Kodroff and Roberge (1975). The results showed that each of the validity principles were easier than the fallacy principle but that the means were not significantly different. Mastery of a principle was judged for at least 6 correct out of a total of 8.

The percent of subjects who were found to have mastered the Contraposition principle in each age group is similar to reported results for chronologically younger normal children (Ennis, 1965,; Gardiner, 1966; Kodroff and Roberge, 1975). Also, mastery of Conversion follows from previous studies (Ennis, 1965; Gardiner, 1966; Miller, 1968; Roberge, 1970), although most of these studies used paper and pencil tests, not pictures. There is reason to believe that picture test items might raise the number of correct answers (Kuhn, 1977).

There was a lack of statistical significance in the content dimensions which was an unexpected result since previous research with normal children showed that concrete content was easier than abstract content (Carroll, 1975; Gardiner, 1966; Miller, 1968; Roberge and Paulus, 1971). For the lower two ages, the Forward Conditional principle had the highest frequency of correct responses and Conversion the highest frequency of errors.

16. Wildman and Fletcher (1979)

This experiment sets out to directly test the "biconditional misinterpretation hypothesis," which states that the "if-then" conditional statement is misinterpreted as a biconditional "if and only if, then" statement (Wildman and Fletcher, 1977). To do this, 237 students in grades 8, 10, 12 and 14 were administered a paper and pencil test consisting of questions reflecting all 16 unique syllogistic forms that are generated by the orthagonal arrangement of:

> (a) stating the major premise in each of the four combinations of positive and negative instances of the antecedent or consequent, and (b) stating the second premise in each of the four cases of affirming or denying the antecedent or consequent. (Wildman and Fletcher, 1979, 367)

These sixteen unique syllogistic forms were generated for both conditionals and biconditionals. Subjects were divided randomly into three groups. Group 1 was administered a 16-item multiple-choice test of conditional syllogisms. Group 2 received a 16-item test of biconditional syllogisms and Group 3 got a 32-item test containing both biconditional and conditional syllogisms. This scheme permitted within subject and within experiment comparisons of problem solving behavior with respect to conditional and biconditional statements. Item content consisted of assigning "brown" and "square" or "black" and "circle" as the antecedent and consequent in the major premise.

Groups 1 and 2 results could be combined and they were analyzed using both biconditional and conditional criteria for correctness. An error analysis was performed to test whether the biconditional misinterpretation hypothesis was adequate in explaining the "apparent logical errors" made on conditionally stated test problems. Of course, only the fallacy test items could be used in this analysis since the validity items have identical "correct" answers for both conditional and biconditional interpretations.

The data gathered from the fallacy items of the sixteen syllogistic forms are summerized in an aggregate fashion in Table 23. The authors claim that the biconditional misinterpretation hypothesis is supported since the biconditionally consistent responses for both the conditional and biconditional major premises were most frequently chosen across all grade levels.

TABLE 23
PERCENTAGE OF CONDITIONALLY CORRECT (C), BICONDITIONALLY CORRECT (B), OR OTHER (O) RESPONSES ON ALL DENY ANTECEDENT AND AFFIRM CONSEQUENT PROBLEMS WITH BICONDITIONALLY OR CONDITIONALLY STATED MAJOR PREMISES
(Wildman and Fletcher, 1979)

		Major premise					
		Biconditionally stated			Conditionally stated		
Grade	N	C (incorrect)	B (correct)	O (incorrect)	C (correct)	B (incorrect)	O (incorrect)
8	49	10	64	27	9	62	29
10	58	10	71	20	19	71	10
12	71	14	72	14	28	58	13
14	59	25	62	12	47	47	6

The eight fallacy items were examined for Group 3. For each of the items, subjects were identified who selected identical biconditional responses for both of the biconditionally and conditionally stated items. In addition, subjects were identified who responed correctly to the biconditionally and conditionally stated items. The data are shown in percentages in Table 24.

TABLE 24

PERCENTAGE OF SUBJECTS[a] (GRADES 8, 10, 12, AND 14) SELECTING (1) THE BICONDITIONALLY CONSISTENT RESPONSE ON BOTH CONDITIONALLY STATED AND BICONDITIONALLY STATED PROBLEMS AND (2) THE CORRECT RESPONSE ON BOTH CONDITIONALLY STATED AND BICONDITIONALLY STATED PROBLEMS (Wildman and Fletcher, 1979)

	Biconditional responding on both conditionally and biconditionally stated problems				Correct responding on both conditionally and biconditionally stated problems			
	Grade				Grade			
Problem[b]	8	10	12	14	8	10	12	14
2	38	90	59	61	0	10	18	22
3	63	90	59	56	6	0	18	28
6	38	55	50	28	0	15	27	44
7	44	65	68	61	6	0	9	11
10	44	55	18	11	0	5	9	6
11	56	50	41	17	6	5	18	50
14	25	35	27	17	0	5	27	11
15	50	80	55	50	0	5	14	28
All	45	65	47	37	2	6	18	25

[a] Percentages are based on sample sizes of 18, 16, 20, and 22, respectively, across grades 8, 10, 12, and 14.

[b] The problem numbers correspond to the cell numbers shown in Tables 1 and 2

[Wildman and Fletcher, 1979]

Table 24 indicates that:

> The tendency of a given subject to select the same biconditionally consistent response to problems presented either conditionally or biconditionally varied somewhat across these eight problems. But the magnitude of the numbers throughout the table attests to the strong tendency of all subjects to give the singular biconditional interpretation to the two logically different sets of problems. (ibid., 378)

For Wildman and Fletcher, the biconditional misinterpretation hypothesis is empirically verified. This is especially true for the younger subjects who display little functional/performance difference on the two different forms of syllogisms. Considering the overall performance picture on the conditional and biconditional problems, the subjects "simply respond in a generalized way" to problems that contain the words "if" and "then". For Wildman and Fletcher, their study shows just how "developmentally complex and puzzling" deductive reasoning ability is.

G. Analysis-III

In these last few studies we see additional evidence for the prevalence of the logical fallacies in the reasoning of the experimental subjects. Paris postulates a "matching scheme" to help explain some of his results. This explanation fits nicely with the notion of transductive reasoning and with Evan's matching bias theory (below).

The replications of Ennis's tests by Berzonsky and Ondrako, and by McAloon, and the Ennis-like studies by Roberge and Paulus, Howell, and Gardiner, all appear to confirm that the validity principles are easier than the fallacy principles and that the concrete familiar content items are easier than the other content dimensions. The studies that considered maturation generally showed a weak correlation of "correct" responses with age. In addition, the studies by Kodrof and Roberge, and Antonok and Roberge seem to indicate a cognitive developmental progression suggesting a transduction explanation.

McAloon found that logic instruction significantly affected test scores and that grade level also had a significant effect. These results were contradicted by Carroll as well as Ennis (1965; 1969). Ryoti and Flener showed that while forms of generalized conditionals were more difficult than standard conditionals, fallacy principles were more difficult than validity principles. Wildman and Fletcher provide a direct experimental test of the biconditional misinterpretation hypothesis. This is in marked contrast to the usual post hoc explanation of past studies.

Overall, the studies on conditional reasoning reported in this chapter show that, using the material conditional interpretation of the conditional statement as the criterion measure of correctness, subjects commit the logical fallacies quite consistently and succeed for the most part on the validity items. There appears to be some improvement with age and grade level. Most of the studies show that it is easier to reason "correctly" using concrete familiar content; that there are no sex and little socioeconomic differences affecting conditional reasoning; that logical reasoning improves somewhat with age and that there is only very limited success in teaching the principles of conditional logic.[4]

Kuhn (1977) suggests that complicating factors in the above types of studies may cause subjects to perform under their "ideal competence" in conditional reasoning. Such factors as length and tediousness of paper and pencil tests might cause subjects to revert to a "less taxing" way of responding such as transductive reasoning or guessing. Also, extra-logical factors might influence the interpretation of a concrete testing situation. For example, in Ennis (1969), subjects might look for some "comprehensible way to interpret" the house apparatus. Kuhn's study showed that children given only 10 test items consisting of highly concrete and simplified pictures yielded a higher percentage of correct answers (using the material conditional interpretation) than other studies have reported.

Kuhn was also interested in determining whether children needed to have mastered class inclusion, a hallmark of the concrete operational stage of cognitive development, before they could master conditional reasoning. In confirming this hypothesis, Kuhn reported that the children in her study initially interpreted the "if-then" conditional statement in the transductive sense and then as the concrete operational stage emerged, recognized that the implication allows for the absence of one proposition in the presence of the other. Later on in the concrete operations stage, the children seem to master the material conditional interpretation of the conditional statement. She also reported that it was the method of presentation and not the smaller number of test items that was responsible for the children's superior performance.

Finally, Kuhn (1977) reported that with a particular test (the "bug" test, which is really a propositional function test) children's performance on conditional reasoning measures was significantly influenced by two essential characteristics of the formal operations stage of cognitive development: structured whole and combinatorial system as measured by the isolation of variables in a problem.

Kuhn's results and some of the studies examined in this chapter provide evidence that young children "naturally" reason in a biconditional-like manner and developmentally progress to the material conditional interpretation of the conditional statement. But, certainly not all progress to this state.

Hidi and Bereiter (1981) conclude that the younger children in their study demonstrated that their reasoning was guided by "linking" together propositions belonging to the same general schema and not by a deductive process. In another study, they characterize as "intuitive responding" leading to the right answers, that which appears to be propositional reasoning. (Bereiter and Hidi, 1977).

Many of the studies considered in this chapter seem to indicate that this biconditional-like reasoning may be due to a transductive mental structure that is opertive. Ennis (1976) is correct in pointing out that we should not confuse transduction with biconditional reasoning. One of the defining characteristics of transduction is that it is not deduction. Biconditional reasoning is deduction.

Ennis also warns that to attribute the presence of transduction in concrete and even formal operational children, due to what Piaget terms "vertical decalage," may be stretching Piaget's theory a bit too far. It is not clear, however, just how the presence of transduction over three developmental stages affects the theory, if at all. This question must await further testing and theory synthesis.

Since, by its very nature, a formal rule system cannot be said to underlie transductive reasoning, I have called its manifestation in conditional reasoning "biconditional-like." Kuhn's suggestion that certain tests or testing situations trigger a regression to transduction (and hence a biconditional-like interpretation of the conditional statement) is an intriquing thought since there is some evidence that cognitive state development is explanatorily more appropriate than a linguistic explanation for the development of conditional reasoning abilities (see Staudenmayer and Bourne, 1977).

H. Conclusion

It seems clear that while for many skills, establishing criteria for successful performance is rather straightforward, utilizing "appropriate" canons of formal logic as criteria for conditional reasoning performance is not as simple a matter. The empirical evidence reported in this chapter lends support to the use of a biconditional-like (or "material equivalence"), rather than a material conditional, interpretation as the standard of logical competence with regard to "if-then" conditional statements expressed in ordinary language.

There appears to be a strong tendency, over a wide age range, for persons to commit the fallacies of conversion and inversion which suggests that the conditional statement is thought to imply its converse and its obverse, respectively. Thus, someone who asserts the conditional statement "If you eat your lunch, then you can play" is committed to the truth of $p \dashrightarrow q$ as well as to the truth of $\bar{p} \dashrightarrow \bar{q}$ (where p and q stand for the antecedent and consequent of the assertion, respectively). We can say that the assertion promises $p \dashrightarrow q$ and suggests or invites the inference, $\bar{p} \dashrightarrow \bar{q}$.

In an even more radical view, we can reject the bivalent nature of the propositional calculus when applied to ordinary discourse. Quine points out that the ordinary conditional statement is really nothing more than a conditional assertion (rather than the assertion of a conditional) and that should the antecedent turn out to be false, it is as if the assertion was never made (Quine, 1972). This view leads to an entirely different truth table for the conditional statement. See Table 25.

TABLE 25

DEFECTIVE LOGIC INTERPRETATION OF THE CONDITIONAL STATEMENT

P	Q	P ----> Q
T	T	T
T	F	F
F	T	Irrelevant
F	F	Irrelevant

A fascinating series of studies by Wason and his colleagues lends empirical support to the altered truth table shown in Table 25. These studies are reviewed and analyzed in Wason and Johnson-Laird (1972). Although a great number of variations of the primary study have been created, the basic experiment remains the same:

> Subjects are given a rule that concerns a set of cards that have letters written on one side and numbers on the other. The rule is: "If a card has a vowel on one side, then it has an even number on the other side." The subject is shown four cards that display on their visible sides a vowel (p), a consonant ($\bar{p}$), an even number (q) and an odd number ($\bar{q}$). The subject is asked which card or cards they need to turn over in order to test the truth or falsity of the rule.

Of course, under the material conditional interpretation of the rule, only the cards showing p and $\bar{q}$ need to be turned over. However, it was found (and subsequently verified in a variety of repliciations) that the subjects tended to select p and q and failed to select $\bar{q}$. Wason explains the source of this error in the failure to appreciate the importance of falsification as opposed to verification. The subjects have a verification bias since their choices provide more chance of getting the confirming disjunct, p.q, than the falsifying disjunct, $p.\bar{q}$.

In an extension to Wason's work, Legrenzi (1970) found that when the situation is strictly binary (e.g., letters are either "A" or "B" and numbers are either "4" or "7"), the conditional rule tended to be evaluated as a material equivalence. While the Wason and Legrenzi results are interesting in and of themselves, Evans (1972a) has discovered a powerful operational (or task) variable which appears to explain Wason's results without resorting to the verification bias theory. Evan's explanation is important because, like transductive reasoning, it is non-logical, and actually parallels the transductively induced biconditional-like interpretation of the conditional statement.

Evans (1972b) attempted to elicit "psychological truth tables" by asking subjects to construct verifying and falsifying cases of conditional rules via the Wason card game. Evans systematically varied the presence and the absence of negative components in these rules and found that his subjects tended to choose values which "matched rather than altered" the values named in the rules themselves, irrespective of the presence of negatives. Evans noted:

> The construction of cases containing a falsified antecedent, argued by Wason (1966) to be psychologically irrelevant to the truth of a conditional rule, was almost entirely determined by matching. (Evans, 1972b, 195)

In another experiment, Evans and Lynch (1973) found that the selection of all of the four logical alternative answers were "largely determined by matching." There is some correspondence between these results and those of Paris (1973).

Evan's work shows that his subjects' reasoning capabilities, with respect to the card problem, are extra-logical. This provides an added warning against classifying reasoning behavior as "correct" or "incorrect" according to some rule of formal logic. Evidence of a "matching bias," lends support to a non-logical transductively induced biconditional-like interpretation of conditional statements. The Wason, Evans, and Legrenzi experiments were performed upon adult subjects which lends support to the idea of vertical decalage with respect to the appearance of transductive reasoning.

It is true, of course, that in ordinary discourse, conditional connectives usually function to establish semantic relations between sentential components. Thus, while conditional connectives may relate truth-values of propositions, they may violate rules which govern cohesion of discourse. For instance, in ordinary usage, the "if-then" conditional serves both to express material conditional and equivalence meanings: "If this is a cat, then it is an animal" and "If she is 18 or older, then she is entitled to vote," respectively.

The propositional calculus merely analyses statements according to their logical forms and ignores their specific content. It is not within the realm of this review to consider the effects of content other than some of the content dimensions reported in this chapter. It is interesting to note that Evans and Newstead (1977) presents evidence that the temporal ordering of antecedent and consequence in the conditional statement significantly affects the correctness of response for adult college level reasoners.

This report has examined considerable evidence supporting a biconditional-like interpretation of the conditional statement which could reflect a transductive reasoning paradigm cutting across developmental stages in an extinguishing fashion and which may reappear due to a number of conjectured factors. It appears that any attempt to teach conditional logic principles runs up against a non-logical structure, yet, the empirical evidence shows that some logic is understood _logically_ and some logic has been taught.[5] Clearly, the _criteria_ used to judge correct and incorrect utilization of the principles of conditional logic is critical.

NOTES

1. Much of the empirical research in this area has been heavily influenced by Jean Piaget's theories and models of cognitive development. Piaget's psychological theories are embedded within a genetic epistemological framework that seeks to account for, amongst other things, the ontogenesis of logical reasoning from birth to adulthood. His theories also attempt to account for non-logical "reasoning" (e.g., transductive reasoning). Although Piaget does not explicitly discuss the child's verbal or written ability to utilize correctly the principles of conditional logic, some researchers have interpreted his work to suggest that propositional reasoning is not possible until age 11 or 12.

   Piaget's theory as it applies to the ontogenesis of deductive necessity and self-evidence forms a theoretical background for much of the research done in the area of human reasoning ability. Unfortunately, a complete account of it is not readily accessible in any one primary or secondary published source. It is scattered throughout Piaget's writings and addressed piecemeal by some of his interpreters. Such an account is, however, attempted in Seidman (1986).

2. I assume that the reader is familiar with the terminology of propositional logic and with the notational shorthand: "p → q" for "if p then q".

3. Mason (1979) identifies and contrasts three different approaches to the study of logical thinking. They are: the logical structure approach, the Piagetian approach, and the information processing/memory approach.

4. See Hardar and Henkin (1978) and Hardar (1977, 1978) for an example of the considerations and limitations in developing a measurement instrument (e.g., the role of negation, content effects and the relations between symbolic logic and the process of psychological reasoning). Also, see Mason (1979) for an example of an open-ended assessment tool.

NOTES

5. Could it be that teaching method and/or medium might be the key to successfully learning logical principles? Conventional wisdom has it that learning computer programming, with its logical structure and "if-then" conditional branch statements, might constitute indirect instruction in logical principles.

According to Piaget's theory of cognitive development, logical necessity and self-evidence develop ontogenetically and there exists a crucial difference between the concrete operational and formal operations stages of development. The transition between these two stages (at about 11-12 years) signals the substitution of deduction for experience. Deduction is now able to deal with pure hypotheses, whereas in the concrete operational stage it was limited by concrete objects and sense experience. The formal operational child is able to reason deductively about abstract propositions. There is some conjecture that the transition between the two stages is aided by certain kinds of "problem circumstances" that help to trigger the process of reflective abstraction which is the driving mechanism in the transitional stage. Hunt (1961), calls this the problem of the "match".

Could computer programming provide such a "problem circumstance"? See Seidman (1981) for an experimental study addressing this question for the LOGO programming language. Ennals (1984) and Ennals and Briggs (1985) investigate the teaching of an artificial intelligence Fifth Generation computer language to young children as an aid to logical thinking.

BIBLIOGRAPHY

Antonak, R. F. & Roberge, J. (1978, Fall). Characteristics of the conditional reasoning of educable retardates. American Educational Research Journal, 15 (4), 519-528.

Barker, J. A. (1969). A formal analysis of conditionals (3rd ed.). Southern Illinois University Monograph.

Bereiter, C. & Hidi, S. (1977). Biconditional versus factual reasoning in children. Biennial Meeting of the Society for Research in Child Development, New Orleans. (ERIC ED 138351)

Berzonsky, M. D. and Ondrako, M. (1974, Fall). Cognitive style and logical deductive reasoning. Journal of Child Psychology. 43(1).

Bonser, F. G. (1910). The reasoning ability of children of the fourth, fifth, and sixth grades. New York: Columbia University.

Brown, R. A. (1973). A first language. Cambridge: Harvard University Press.

Burt, C. (1919). The development of reasoning in school children (I, II). Journal of Experimental Pedagogy, 5(2&3).

Campbell, D. T. & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally and Co.

Carrol, C. A. (1970). Low achiever's understanding of four logical inference forms: an analysis of difficulties and of the effect of instruction. Unpublished Doctoral Dissertation. Columbia University.

Donaldson, M. (1963). A study of children's thinking. London: Tavistock.

Eisenberg, T. A. & McGinty, R. L. (1974, Nov.). On comparing error patterns and the effect of maturation in a unit on sentential logic. Journal for Research in Mathematics Education, 5(4).

Ennals, R. & Briggs, J. (1985). Fifth generation computing: introducing micro-PROLOG into the classroom. Journal of Educational Computing Research, 1(1).

Ennals, J. R. (1984). Beginning micro-PROLOG (2nd revised ed.). London: Ellis Horwood and Heinemann, Harper Row.

Ennis, R. H. (1962). A concept of critical thinking: a proposed basis for research in the teaching and evaluation of critical thinking ability. Harvard Education Review, 32, 81-111.

Ennis, R. H. (1964). Operational definitions. American Educational Research Journal, (1), 183-201.

Ennis, R. H. (1967). Readiness to master a principle. Psychological Concepts in Education. (Edited by B. P. Komisar and C. J. B. Macmillan., Chicago: Rand-McNally).

Ennis, R. H. (1970). Conditional logic and primary children. Annual Meeting: American Educational Research Association. (Minneapolis).

Ennis, R. H. (1971). Conditional logic and primary school children: a developmental study. Interchange, 2(2).

Ennis, R. H. (1975). Children's ability to handle Piaget's propositional logic: a conceptual critique. Review of Educational Research, 45(1), 1-41.

Ennis, R. H. (1976). An alternative to Piaget's conceptualization of logical competence. Child Development, 47, 903-919.

Ennis, R. H., Finkelstein, M. R.; Smith, E. L. & Wilson, N. H. (1969). Conditional logic and children. Critical Thinking Project. (Cornell University).

Ennis, R. H. & Paulus, D. (1965). Critical thinking readiness in grades 1-12 (Phase I: deductive reasoning in adolescence). Critical Thinking Project. (Cornell University).

Evans, J. St. B. (1972a). On the problem of interpreting reasoning data: logical and psychological approaches. Cognition, 1(4), 373-384.

Evans, J. St. B. (1972b). Interpretation and 'matching bias' in a reasoning task. Quarterly Journal of Experimental Psychology, 24, 193-199.

Evans, J. St. B. & Lynch, J. S. (1973). Matching bias in the selection task. British Journal of Psychology, 64, 391-398.

Evans, J. St. B. & Newstead, S. E. (1977). Language and reasoning: a study of temporal factors. Cognition, 5, 265-283.

Falmagne, R. J., ed. (1975). Reasoning: representation and process in children and adults. New Jersey: Lawrence Erlbaum Associates.

Flavell, J. H. (1963). The developmental psychology of Jean Piaget. New York: VanNostrand & Reinhold.

Flener, F. O. (1973). A comparison of reasoning with general and singular propositions by 5th, 7th, and 9th grade students. Unpublished Doctoral Dissertation. (University of Illinois).

Gardiner, W. L. (1965). An investigation of understanding of the meaning of the logical operators in propositional reasoning. Unpublished Doctoral Dissertation. (Cornell University).

Hadar, N. (1977). Children's conditional reasoning. Educational Studies in Mathematics, 8, 413-438.

Hadar, N., & Henkin, L. (1978). Children's conditional reasoning Part II: towards a reliable test of conditional reasoning ability. Educational Studies in Mathematics, 9, 97-114.

Hadar, N. (1978). Children's conditional reasoning Part III: a design for research on children's learning of conditional reasoning and research findings. Educational Studies in Mathematics, 9, 115-140.

Hamlyn, D. W. (1978). Experience and the growth of understanding. London: Routledge and Kegan Paul.

Hidi, S. & Bereiter, C. (1981). Plausibility versus logical necessity in children's development. Biennial Meeting of the Society for Research in Child Development, Boston. (ERIC ED 200303)

Hill, S. (1961). A study of logical abilities in children. Unpublished Doctoral Dissertation. (Stanford University).

Howell, E. N. (1965). Recognition of selected inference patterns by secondary school mathematics students. Unpublished Doctoral Dissertation. (University of Wisconsin).

Hunt, J. McV. (1961). Intelligence and experience. New York: Ronald.

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

Inhelder, B. & Piaget, J. (1964). The early growth of logic in the child: classification and seriation. New York: Harper and Row.

Kagan, J., Rosman, P. L., Day, D., Albert, J., Philips, W. (1964). Information processing in the child. Psychological Monograph, 78(1), Whole No. 578.

Knifong, J. D. (1974, March). Logical abilities of young children--two styles of approach. Child Development, 45(1), 78-83.

Kodroff, J. K. & Roberge, J. J. (1975). Developmental analysis of the conditional reasoning abilities of primary grade children. Developmental Psychology, 11(1), 21-28.

Kuhn, D. (1977). Conditional reasoning in children. Developmental Psychology. 13(4), 342-353.

Legrenzi, P. (1970). Relations between language and reasoning about deductive rules. Advances in Psycholinguistics. (Edited by G. B. Flores, and W. J. M. Levelt). Holland: North-Holland.

Lunzer, E. A. (1965). Problems of formal reasoning in test situations. European Research in Cognitive Development: Monographs of the Society for Research in Child Development. (Edited by P. H. Mussen), 30(2).

McAloon, M. D. (1969). An exploratory study in teaching units in logic to grades three and six. Unpublished Doctoral Dissertation. (University of Michigan).

Martens, M. A. (1967). Use of selected rules of logical inference and of logical fallacies by senior high students. Unpublished Doctoral Dissertation. (University of Wisconsin).

Mason, E. J. (1979). The development of logical thinking in children. Report to the Netherlands Ministry for Pure Science Research (ZWO), Catholic University, Nijmegen, Netherlands.

Matalon, B. (1962). 'Etude genetiqua de l'implication' in Etudes d'espiste-mologie genetique: XVI, Implication, formalization, et logique naturelle. Geneva, 69-95.

Miller, W. (1968). The acceptance and recognition of six logical inference patterns by secondary students. Unpublished Doctoral Dissertation. (University of Wisconsin).

Miller, W. A. (1969). A unit of sentenial logic for junior high school students, involving both valid and invalid inference patterns. School Science and Mathematics, 69(6), 548-552.

Moore, T. V. (1929). The reasoning ability of children in the first years of school life. Studies in Psychology and Psychiatry, 2(2), 1-34.

Morgan, W. J. & Morgan, A. B. (1953). Logical reasoning: with and without training. Journal of Applied Psychology, 37(5), 399-401.

O'Brien, T. C. & Shapiro, B. J. (1968). The development of logical thinking in children. American Educational Research Journal, 5, 531-542.

Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. Journal of Experimental Child Psychology, 16, 278-291.

Paulus, D. H. (1967). A study of children's abilities to deduce and judge deductions. Unpublished Doctoral Dissertation, (Cornell University).

Peel, E. A. (1967, May). A method for investigating children's understanding of certain logical connectives used in binary propositional thinking. British Journal of Mathematical and Statistical Psychology, 20(Part I), 81-97.

Piaget, J. (1926). The language and thought of the child. New York: Harcourt, Brace.

Piaget, J. (1928). Judgement and reasoning in the child. New York: Harcourt, Brace.

Quine, W. V. (1972). Methods of logic(3rd ed.). New York: Holt, Rinehart and Winston.

Roberge, J. J. (1970). A study of children's abilities to reason with basic principles of deductive reasoning. American Educational Research Journal, 7, 583-595.

Roberge, J. J. (1971, Dec.). Some effects of negation on adult's conditional reasoning abilities. Psychological Reports, 29(3), 839-844.

Roberge, J. J. (1972). Recent research on the development of children's comprehension of deductive reasoning schemes. School Science and Mathematics, 72, 197-200.

Roberge, J. J. & Paulus, D. H. (1971). Developmental patterns for children's class and conditional reasoning abilities. Developmental Psychology, 4, 191-200.

Robitaille, H. J. (1943). The development of a test for reasoning ability and its application to students majoring in science, classics and commercial courses. Studies in Psychology and Psychiatry, 2(6).

Ryoti, D. E. (1972). Student responses to equivalent inference schemes in class and conditional logic. Unpublished Doctoral Dissertation, (University of Illinois).

Seidman, R. H. (1981). The effects of learning a computer programming language on the logical reasoning of school children. American Educational Research Association, Annual Meeting, Los Angeles. (ERIC ED 205206)

Seidman, R. H. (1986). Deductive necessity and the logical structures of reasoning: Piaget's psycho-logic models. Research Report 86-1, New Hampshire College Graduate School, Manchester, NH, 92 pages.

Shapiro, B. J. & O'Brien, T. C. (1970). Logical thinking in children ages six through thirteen. Child Development, 41, 823-829.

Smith, B. O. (1957, Oct.). Logic, thinking and teaching. Educational Theory, 7(4).

Smith, B. O. & Ennis, R. H., eds. (1961). Language and concepts in education. Chicago: Rand-McNally.

Smith, G. (1966). Commentary upon Suppes-Binford report of teaching mathematical-logic to 5th and 6th grade pupils. The Arithmetic Teacher, 13(8), 640-643.

Staudenmayer, H. & Bourne, L. E., Jr. (1977). Learning to interpret conditional sentences: a developmental study. Developmental Psychology, 1(13b), 616-623.

Sternberg, R. J. (1982). Reasoning, problem solving, and intelligence. Handbook of Human Intelligence, (In R. J. Sternberg Ed.), New York: Cambridge University Press.

Suppes, P. (1964). The ability of elementary school children to learn the new mathematics. Theory into Practice, 3, 57-61.

Suppes, P. (1970). On the behavioral foundation of mathematical concepts. Cognitive Development in Children. Illinois: University of Chicago Press.

Suppes, P. & Binford, F. (1965). Experimental teaching of mathematical logic in the elementary school. The Arithmetic Teacher, March.

Suppes, P. & Feldman, S. (1971). Young children's comprehension of logical connectives. Journal of Experimental Child Psychology, 12, 304-317.

Suppes, P. & Ginsberg, R. (1962, April). Experimental studies of mathematical concept formation in young children. Science Education, 46(3), 230-240.

Taplin, J. E. (1971a). An experimental study of human reasoning and conceptual behavior. Unpublished Doctoral Dissertation, (University of Adelaide, South Austrailia).

Taplin, J. E. (1971b). Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behavior, 10, 219-225.

Taplin, J. E. & Staudenmayer, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behavior, 12(5), 530-542.

Taplin, J. E., Staudenmayer, H. & Taddonio, J. L. (1974). Development changes in conditional reasoning: linguistic or logical? Journal of Experimental Child Psychology, 17, 360-373.

Tucker, R. C. (1970). A study of the logical abilities of children as a function of mental and chronological age. Unpublished Doctoral Dissertation, (St. Louis University).

Wason, P. C. (1966). Reasoning. New Horizons in Psychology-I. (Edited by B. Foss), Baltimore: Penguin.

Wason, P. C. & Johnson-Laird, P. N. (1972). The psychology of reasoning: structure and content. Cambridge: Harvard University Press.

Wildman, T. M. & Fletcher, H. J. (1977). Developmental increases and decreases in solutions of conditional syllogism problems. Developmental Psychology, 13(6), 630-636.

Wildman, T. M. & Fletcher, H. J. (1979). Processing errors in conditional and biconditional problem solving behavior. Contemporary Educational Psychology, 4, 366-380.

Wilkins, M. C. (1928). The effect of changed material on ability to do formal syllogistic reasoning. Archives of Psychology, 16(102), 1-83.

Wilmoth, J. N. (1973). Factors that may affect deductive reasoning in adult students enrolled in educational research classes: a review of empirical and expository documents from non-textbook literature between 1920-1973. School of Education, Auburn University. (ERIC ED 087700)

Winch, W. H. (1921). Children's reasoning: experimental studies of reasoning in school children. Journal of Experimental Pedagogy, 6, 121-141, 199-212, 275-287.

Woodcock, L. P. (1941). Life and ways of the two-year-old. New York: E. P. Dutton & Co., Inc.

[illegible] J. [illegible] & Tasdemir [illegible] (19[illegible]). [illegible] conditional [illegible] Linguistic [illegible] Journal of [illegible] Psychology [illegible]

[illegible] (1970). A [illegible] of [illegible] [illegible] Dissertation [illegible]

[illegible] The Nonsense [illegible] [illegible]

[illegible] (1971). The psychology of [illegible] Cambridge [illegible]

[illegible] (1977). Development [illegible] conditional [illegible] process [illegible] Journal of [illegible] Psychology [illegible]

[illegible] & [illegible] (1979). Processing [illegible] conditional [illegible] reasoning behavior. Journal of [illegible] Psychology [illegible]

[illegible] The effect of changed [illegible] Archives of Psychology [illegible]

[illegible] conditional [illegible] educational [illegible]

[illegible] Children's reasoning [illegible] Journal of [illegible] Psychology [illegible]

[illegible]

APPENDIX A

TABLE A-1

LOGICAL FORM OF ITEMS IN THE CORNELL CONDITIONAL REASONING TEST, FORM X
(from Ennis, 1965, IV-4)

Item Group	Princi-ple(s)	Basic Form which was used for the first two concrete familiar items (CF1 and CF2), the symbolic item (SY) and the suggestive item (SU)	Answer for Basic Form	Form of CF3	Answer to CF3	Form of CF4 (Same answer as basic form)
1	1	If p, then q p : q	Yes	If p, then q p : not q	No	not p If not p, then q : q
2	2	If p, then q not p : not q	Maybe	Same as basic form	Maybe	p If not p, then q : not q
3	3	If p, then q q : p	Maybe	Same as basic form	Maybe	q If not p, then q : not p
4	4	If p, then q not q :p	No	If p, then q not q : not p	Yes	not q If not p, then q : not p

TABLE A- 2

MEAN DIFFICULTY AND DISCRIMINATION INDICES FOR
THE CORNELL CONDITIONAL REASONING TEST
BASED UPON PRE-TEST OF LDT'S, LNDT-1'S AND LNDT-2'S COMBINED
(Ennis, 1965, IV-30)

	Mean Difficulty Indices				Mean Discrimination Indices			
Grade	5	7	9	11	5	7	9	11
N =	102	99	80	78	102	99	80	78
Item Group								
1	71.4	74.7	77.3	78.8	37.0	32.1	18.9	34.9
2	22.7	27.4	24.8	35.3	-17.9	16.7	5.3	27.8
3	17.7	26.8	31.5	35.7	-13.0	6.2	3.8	29.4
4	55.7	68.7	59.6	65.0	50.0	33.9	24.2	25.4
CF	48.9	55.4	55.2	61.8	23.2	26.3	19.4	32.0
SY	48.1	55.8	53.4	59.5	29.0	26.5	21.2	19.1
SU	41.3	53.4	53.3	59.9	34.3	26.5	22.3	31.8
Mean over all items	47.5	55.8	54.6	61.5	28.1	27.0	19.8	29.1

Mean Discrimination Index for All Grades on Total Test: 26.0
Total N = 359

TABLE A-3

MEAN CHRONOLOGICAL AGE; IQ; ESTIMATED MENTAL AGE;
AND TOTAL, COMPONENT, AND ITEM GROUP CONDITIONAL REASONING
PRE-TEST SCORES; BY GRADE FOR ALL SUBJECTS GROUPED TOGETHER
(adapted from Ennis, 1965, V-10)

	Conditional Reasoning			
Grade	05	07	09	11
N =	102	99	80	78
Chronological Age (mos.)	129	153	184	203
IQ	108	117	110	109
Estimated Mental Age (mos.) (CA x IQ/100 before rounding)	139	179	201	220
Total Score*	42.4	51.7	55.3	56.6
Component**				
CF	23.3	27.1	29.0	29.5
SY	5.8	6.7	6.5	7.2
SU	4.6	6.0	6.4	6.4
Item Group**				
1	4.3	4.5	4.9	4.7
2	1.4	1.7	2.1	2.2
3	1.2	1.6	2.1	2.0
4	3.3	4.1	3.8	3.9

*Total score (for all 12 item groups) was calculated using scoring formula: R-W/2 + 27.
**Component and Item Group scores are number of right answers.

TABLE A-5

PERCENT MEETING THE SUFFICIENT CONDITION AND THE PERCENT FAILING TO MEET THE NECESSARY CONDITION FOR MASTERY OF EACH PRINCIPLE AT EACH GRADE LEVEL ON THE CONDITIONAL REASONING TEST (from Ennis, 1965, V-18)

	Percent Meeting the Sufficient Condition				Percent Failing to Meet the Necessary Condition			
Grades	05	07	09	11	05	07	09	11
N =	102	99	80	78	102	99	80	78
Item Group								
1	51	56	66	62	30	26	21	22
2	3	6	5	12	92	80	90	73
3	2	3	4	3	94	92	89	85
4	30	41	35	35	54	36	41	40

TABLE A-5

DESCRIPTION OF ITEMS ON THE SMITH-STURGEON CONDITIONAL REASONING TEST
(Ennis, 1969, 41a)

Principle	Logical Form of Items[1]	Alternative Answers	Content of Items: Item	Content of Items: Part of Test	Content of Items: Form
Inversion	If p, then q	q	5	Chemicals	Suppositional
	<u>not p</u>	not q	6	Chemicals	Factual
	∴ ?	<u>maybe</u>	11	Chemicals	Suppositional
			12	Chemicals	Factual
			17	House	Suppositional
			18	House	Factual
Conversion	If p, then q	p	1	Chemicals	Suppositional
	<u>q</u>	not p	2	Chemicals	Factual
	∴ ?	<u>maybe</u>	7	Chemicals	Suppositional
			8	Chemicals	Factual
			15	House	Suppositional
			16	House	Factual
Contraposition	If p, then q	p	3	Chemicals	Suppositional
	<u>q</u>	<u>not p</u>	4	Chemicals	Factual
	∴ ?	maybe	9	Chemicals	Suppositional
			10	Chemicals	Factual
			19	House	Suppositional
			20	House	Factual

[1]The three alternatives presented to the child are given and the correct one is underlined.

Robert H. Seidman

TABLE A-6

MASTERY OF FOUR BASIC PRINCIPLES OF CONDITIONAL LOGIC
(Ennis, 1969, 54a)

		Inversion				Conversion				Contraposition			
		Mastery (5-6)	? (4)	Non-Mastery (2-3)	Non-Mastery (0-1)	Mastery (5-6)	? (4)	Non-Mastery (2-3)	Non-Mastery (0-1)	Mastery (5-6)	? (4)	Non-Mastery (2-3)	Non-Mastery (0-1)
Grade 1	C N = 30	6	6	7	11	0	7	8	15	12	7	7	4
	E N = 30	4	6	8	12	0	5	9	16	12	9	6	3
	Combined N = 60	10	12	15	23	0	12	17	31	24	16	13	7
Grade 2	C N = 28	12	5	1	10	3	9	4	12	18	8	2	0
	E N = 30	5	4	10	11	0	6	7	17	14	7	7	2
	Combined N = 58	17	9	11	21	3	15	11	29	32	15	9	2
Grade 3	C N = 29	9	6	6	8	2	8	7	12	18	10	1	0
	E N = 30	8	10	5	7	2	10	9	9	17	8	4	1
	Combined N = 59	17	16	11	15	4	18	16	21	35	18	5	1
All Grades Combined	C N = 87	27 (31%)	17 (18%)	43 (51%)		5 (6%)	24 (26%)	58 (68%)		48 (55%)	25 (29%)	14 (16%)	
	E N = 90	17 (19%)	20 (22%)	53 (59%)		2 (2%)	21 (23%)	67 (75%)		43 (48%)	24 (27%)	23 (25%)	
	Total N = 177	44 (25%)	37 (20%)	96 (55%)		7 (4%)	45 (25%)	125 (61%)		91 (51%)	49 (28%)	37 (21%)	

Note: A score of 5 or 6 resulted in classification under mastery; a score of 4 was judged borderline; a score of 3 or below resulted in classification under non-mastery. See Chapter 4 for explanation of method of obtaining score.

TABLE A-7

MEAN SCORES AND STANDARD DEVIATIONS ON FOUR BASIC PRINCIPLES,
ON SUPPOSITIONAL AND FACTUAL ASPECTS, AND ON TOTAL TEST
(Ennis, 1969, 54b)

			Inversion (6 items)			Conversion (6 items)			Contraposition (6 items)			Total (24 items)		
Grade	Group & No.	Statistic	S*	F*	S+F	S	F	S+F	S	F	S+F	S	F	S+F
1	C*	Mean	1.1	1.5	2.6	.80	.90	1.7	1.8	2.1	3.9	5.1	6.0	10.7
	N=30	S.D.*	1.1	1.2	2.1	.81	1.9	1.6	1.0	.88	1.8	3.2	2.5	5.1
	E*	Mean	1.0	1.2	2.2	.77	.73	1.5	1.7	2.2	3.9	4.9	5.9	10.8
	N=30	S.D.	1.1	1.0	2.0	.73	.83	1.5	.88	.76	1.5	2.6	2.4	4.9
	C+E	Mean	1.1	1.4	2.4	.78	.82	1.6	1.7	2.2	3.9	5.0	5.9	10.7
	N=60	S.D.	1.1	1.1	2.0	.76	.85	1.5	.94	.82	1.6	2.9	2.5	4.9
2	C	Mean	1.5	1.9	3.5	1.1	1.3	2.5	2.2	2.7	4.9	6.6	7.7	14.3
	N=28	S.D.	1.3	1.1	2.3	1.1	1.3	2.5	.69	.46	1.0	3.2	2.4	5.5
	E	Mean	1.0	1.5	2.5	.60	.87	1.5	1.8	2.4	4.2	5.3	6.6	11.8
	N=30	S.D.	1.1	1.0	2.0	.86	.94	1.7	1.1	.76	1.6	2.9	2.4	5.0
	C+E	Mean	1.3	1.7	3.0	.88	1.1	2.0	2.0	2.5	4.5	5.9	7.1	13.0
	N=58	S.D.	1.2	1.1	2.2	.99	1.0	1.9	.92	.65	1.4	3.0	2.4	5.3
3	C	Mean	1.5	1.6	3.0	1.1	1.0	2.0	2.5	2.6	5.1	7.0	7.0	14.1
	N=29	S.D.	1.2	1.2	2.3	.98	.96	1.9	.57	.50	1.0	2.8	2.7	5.4
	E	Mean	1.5	1.8	3.3	1.2	1.2	2.4	2.3	2.5	4.7	6.8	7.3	14.2
	N=30	S.D	1.1	1.0	2.0	.90	.92	1.8	.78	.78	1.4	2.6	2.5	4.9
	C+E	Mean	1.5	1.7	3.3	1.2	1.1	2.3	2.4	2.5	4.9	7.0	7.2	14.2
	N=59	S.D.	1.1	1.1	2.1	.93	.94	1.8	.70	.65	1.2	2.6	2.5	5.0
All 3 Combined	C	Mean	1.4	1.7	3.0	1.0	1.1	2.0	2.2	2.4	4.6	6.2	6.9	13.0
	N=87	S.D.	1.2	1.2	2.3	.95	.99	1.9	.83	.69	1.4	3.1	2.6	5.5
	E	Mean	1.2	1.5	2.7	.87	.93	1.8	1.9	2.3	4.3	5.7	6.6	12.3
	N=90	S.D.	1.1	1.0	2.0	.86	.91	1.7	.94	.77	1.5	2.8	2.5	5.1
	C+E	Mean	1.3	1.6	2.9	.94	1.0	1.9	2.0	2.4	4.4	6.0	6.8	12.7
	N=187	S.D.	1.2	1.1	2.1	.90	.94	2.9	.89	.73	1.5	3.0	2.5	5.3

Note: 'S' = 'suppositional'; 'F' = 'factual'; 'S.D.' = 'estimated population standard deviation; 'C' = 'control'; 'E' = 'experimental'

TABLE A-8

COMPARISONS OF CONTROL GROUP PART SCORES
N=87 (d.f.=86) (Ennis, 1969, 55a)

		Comparisons	Mean	Difference	t	Superiority indicated (all differences are significant)
1.	A.	Fallacy Principles	5.06	2.87	6.8	
	B.	Validity Principles	7.93			X (Validity Principles)
2.	A.	Suppositional Items	6.21	0.68	4.1	
	B.	Factual Items	6.89			X (Factual Items)
3.	A.	Inversion	3.01	0.97	5.6	X (Inversion)
	B.	Conversion	2.04			
4.	A.	Inversion	3.01	1.60	6.5	
	B.	Contraposition	4.61			X (Contraposition)
6.	A.	Conversion	2.04	2.57	12.2	
	B.	Contraposition	4.61			X (Contraposition)

Critical Values for Two-Tailed t Tests

Degrees of Freedom	Levels of Significance .05	.01
60	2.00	2.66
120	1.98	2.62