Learning to Solve Complex Propositions: Does knowledge of truth-values bootstrap modal operators?

Bradley J. Morris (bjmorris@pitt.edu)
University of Pittsburgh, LRDC, 3939 O'Hara St.
Pittsburgh, PA 15260 USA

David Klahr (klahr@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University
Pittsburgh, PA 15213 USA

Abstract

Evaluating complex propositions requires evaluating truth-values and assigning modal operators. Previous research suggested that evaluating truth-values may be the key to assigning modal operators. This study placed 111 third and fifth grade children in one of three training conditions: no training, training truth-value assignment, and training truth-value and modal operator assignment. The results indicate that truth-value assignment training is sufficient to significantly improve children's evaluations of complex propositions.

Reasoning with complex propositions (statements using AND, OR, NOT, IF) forms the basis of much higher-order thinking. There are three reasoning classes associated with processing propositions: evaluating a proposition as true or false (truth-values), evaluating whether a conclusion follows from the premises (validity), and judgments about possibility and necessity (modal operators). While much research has focused on judgments about validity (for a recent review see Markovits & Barrouillet, 2002), we will focus on a less-researched area: evaluating truth-values and assigning modal operators.

The assignment of truth-values entails determining the truth or falsity of a statement (Johnson-Laird, 1983). The complexity of assigning truth-values depends on the number of elements being evaluated (i.e., how many items need to be evaluated) and the number of states under evaluation (the number of combinations and their associated truth-values). To better illustrate this point we will provide examples of two representative tasks. The first is sentence verification. In a sentence verification task, subjects typically are given a simple proposition (e.g., the star is white) to evaluate either with their existing knowledge or with some reference materials (e.g., picture of a white star). There are two possible values for each proposition: true or false. Because there is only one element under consideration, the evaluation is based on semantic properties (Roberts, Wood, & Gilmore 1994). The second type of task, evaluating complex propositions such as conjunctions and disjunctions, is more complicated than sentence verification because it requires the evaluation of two elements and four possible states. For example, when evaluating a conjunction (e.g., the star is white and the circle is blue), each single proposition has its own truth-value (white star; blue circle). Additionally, the statement as a whole is only true if both single propositions are true, thus there is only one of the four possible resulting combinations that results in an assignment of "true" for the entire statement.

Assigning modal operators is determining when a statement is possible or necessary (Johnson-Laird, 1983). Like the assignment of truth-values, the assignment of modal operators differs in complexity depending on the nature of the task. Modal operators can be assigned on statements such as "A brother is a boy" in which by definition the statement is necessarily true (Miller, Custer, & Nassau, 2000). In this task the assignment is based on purely semantic factors. A more difficult task is assigning modal operators for complex propositions such as contradictions and tautologies. To determine that a tautology is always true (possible) and that a contradiction is always false (impossible) requires evaluating the semantics and syntax of the statement. That is, one must consider the truth-value of the connective and whether the semantic elements match any of the possible truth-values. Thus, a contradiction will always be false because one of the two propositions will always be false and an AND statement requires both elements to be true for the entire statement to be true.

Development of Truth-values and Modal Operators

Very little attention has been given to how children coordinate assigning truth-values and assigning modal operators. That is, are these processes related and if so, how? Perhaps gains on one phase do not correspond to gains on the other, thus we will call this possibility the *separate phase hypothesis*. Much previous research on either process has focused on a single process without examining the other typically reporting performance in one without respect to changes in the other (Ruffman, 1998; Braine & Rumain, 1981; Osherson & Markman, 1976; Paris, 1974). Thus, perhaps the two are not related.

The only theoretical position that has examined both processes, mental logic, states that the two processes are part of a single inferential schema that is acquired with language (Braine & O'Brien, 1997). Once activated, these schemas fire a series of inferential rules that produce a correct conclusion for all types of inferences (truth-values and modal operators). Thus this position suggests that children should either err on assigning truth-values or modal operators or produce correct evaluations for both, but there should be no inconsistency. We will call this approach the *compiled phase hypothesis*. Previous research demonstrates differences in young children's ability to assign truth-values and modal operators (Morris & Sloutsky, 2002; Miller, Custer, & Nassau, 2000; Braine & Rumain, 1981; Osherson & Markman, 1976), however, most of this research did not look at the consistency of responses within individual children. Two studies that examined within-participant consistency (Morris & Sloutsky, 2002; Morris & Klahr, in review) found large differences in children's performance specifically, that children produced correct responses on assigning truth-values before modal operators. These differences may suggest that such coordination in these types of inferences may occur late in development (see Morris & Sloutsky, 2002, for a discussion).

It is also possible that the two types of inferences are related and that knowledge about one is related to the development of the other. We will call this the *dependent phase hypothesis*. Morris & Sloutsky (2002) and Morris and Klahr (in review) examined children's assignment of truth-values and modal operators and observed that many children erred on one while giving correct responses on the other, providing evidence against the compiled phase hypothesis. Further, in each case, children who correctly assigned modal operators also assigned correct truth-values while the converse was not true suggesting that the ability to assign correct truth-values *preceded* the ability to assign correct modal operators. The *dependent phase hypothesis* suggests that children must first learn to assign correct truth-values before they can assign modal operators because the mappings that allow truth-value provide sufficient information to infer modal operators.

The present study was designed to compare the three approaches and to test the dependent-phase hypotheses. A training study was conducted in which children were given instruction on assigning truth-values and modal operators. Children were randomly assigned to one of three conditions. The first instructional condition ("mapping") provided explicit instruction on evidence evaluation only. The second instructional condition ("necessity") gave children explicit instruction on evidence evaluation and evidence requests. The third condition ("control") gave children no instruction.

Our first prediction is that mere exposure to logical statements is insufficient to improve performance on any processing phase. If all conditions show equal improvement on the post-test, then this supports the compiled phase hypothesis because learning was not required, only familiarity with the type of problem. The second prediction is that training is necessary for improvements in performance. If the mapping and necessity conditions show significant improvement from pre- to post-test and children's performance in these conditions are significantly better than the performance of children in the control condition, this provides evidence against the compiled phase hypothesis (because training was required to improve performance) and supports the separate and related phases hypotheses. The third prediction is that improvement is specific to the phases on which training has been given. If children in the mapping and necessity condition are not be significantly different in their post-test performance levels then this supports the dependent phase hypothesis, however, if the mapping and necessity conditions are different on the post-test, then this supports the separate phase hypothesis.

Method

Participants

The participants were 111 children: 60 third (mean age 8,8; 29 boys, 31 girls) and 51 fifth (mean age 10,8; 25 boys, 26 girls) graders enrolled in two public or two private schools located in Pittsburgh, Pennsylvania. Participants were chosen on the basis of returning a parental consent form. Third and fifth grade children were selected because cross-sectional studies (e.g., Morris & Sloutsky, 2002) have demonstrated significant improvements in logical reasoning between these ages. Children in each grade were randomly assigned to one of the three conditions.

Design

We used a 3 (condition) x 2 (grade) x 4 (session) design with session as within-participants measure. The three conditions (control, mapping, necessity) differed in the amount of explicit instruction on how to evaluate different types of logical statements. The control group was given a series of 12 problems formally identical to those in the pretest in both training sessions without any instruction. Children in the mapping condition were given explicit instruction about rules used to evaluate evidence in logical statements. Children in the necessity condition were given the same evaluation rules as those in the mapping condition and were given explicit instruction about when evidence was necessary and unnecessary.

Each child participated in four sessions: a pre-test, training 1, training 2, and a post-test. In the pre-test, children were asked to evaluate the truth-status of a series of 16 statements. In the second and third sessions children were given one of the three training conditions. In the fourth session each child was given a post-test formally identical to the pre-test.

Procedure

The procedure was divided into four sessions over four days: (1) Day 1- Pre-test, (2) Day 2- Training 1, (3) Day 3- Training 2, (4) Day 4- Post-test. Each session was separated by approximately one week (M= 8.2 days). Each child was interviewed individually in a quiet location in his or her school. The interviewer recorded each child's responses for all four sessions.

Session 1- Pre-test

All instructions for the experimental segment were read to each participant and repeated if requested. The pre-test consistent of 4 warm-up statements and 16 actual statements. Two cards were placed in front of each child: a statement card (face up) and an evidence card (face down). The order of presentation of statement and evidence pairs was counterbalanced across participants. Each child was presented a total of 16 statements corresponding to 4 of each of the following types: conjunctions, disjunctions, tautologies, and contradictions. The child then read the statement card aloud. After the child read the statement, Question phase 1 (*a priori* evaluation) was then asked ("Is the sentence true, not true, or can't you tell?"). After the answer was recorded, Question phase 2 was asked (evidence request: "Do you need to see the picture to find out?"). If evidence was requested, the evidence card was turned over (all were pictures) and placed in front of the participant. Question phase 3 was then asked (evidence evaluation) in which the child was asked to evaluate the initial statement using the evidence requested ("Now that you have seen the picture, was the sentence true, not true, or can't you tell?"; Asked only if evidence was requested). The task took approximately 15 minutes.

Session 2- Training 1

Control Condition. The procedure used in the control condition was identical to the procedure used in the pre-test. Children in the control condition were given a set of 12 statements structurally identical to those in the pre-test but with altered content. As in the pre-test children were asked up to three questions for each statement corresponding to the three processing phases. Children were given no feedback or instructions after their responses. This procedure lasted 10 minutes.

Training Warm-up Segment. Note the training warm-up was used only with the mapping and necessity conditions. The training warm-up segment was a brief session in which children were given basic rules for evaluating evidence with the connectives AND or OR. The training consisted of three parts: (1) explanation of sentence parts, (2) explanation of evaluation rules, and (3) rule use/feedback. Children were first given an explanation of the "parts" of each sentence in the task. Each sentence was divided into two parts (clauses separated by the connective) and an "important word" (the logical connective). Children were told that the important word indicated which rule was used to evaluate each type of statement.

Children were then asked to demonstrate the parts and important words on a new sentence with feedback provided for errors. Next, children were given brief instruction on how to evaluate evidence with each type of important word. Children were given simple rules for each connective. Children were told that the evidence had to match both parts for an AND sentence to be true, otherwise it is false. Children were told that evidence had to match only one part of the statement for an OR to be true otherwise the statement is false. The decision to explain an OR as an exclusive OR was made to make a clean conceptual distinction between OR and AND.

Once the rules for important words were explained, children were given a statement and a series of evidence cards, presented one at a time, and asked to identify the parts and important word then to evaluate the statement as a whole. In total, two statements (1 OR, 1 AND) and four evidence cards (1 true, 1 false for each statement type) were given over the training warm-up. Feedback was provided for incorrect responses giving the correct answer and re-explaining the evaluation rule. This procedure took 5 minutes. The warm-up was only given before Training 1.

Mapping Condition. The mapping condition provided explicit instruction about evaluating evidence with statements. Each child was given four statements (one of each type) and three different pieces of evidence for each statement. The mapping condition provided instruction for each statement in three stages over which the scaffolding provided by the interviewer was gradually reduced: (1) explicit instruction, interviewer-led solution, (2) probe questions, scaffolded solution and (3) probe questions, child-led solution.

In the first stage of the instruction, the interviewer placed a statement card on the table and asked the child to evaluate the statement as true, not true, or can't tell. The interviewer then placed the first evidence card on the table and asked the child to evaluate the statement based on the evidence card. The interviewer then provided explicit instruction about the parts and important words of each statement and the correct conclusion regardless of the child's response. The second evidence card was then placed on the table beside the first evidence card. The interviewer asked the child to repeat the rule for the important word, match evidence to each part of the sentence, and then to evaluate the sentence as a whole. If any evaluation was incorrect, children were given immediate feedback. The third evidence card was placed on the table beside the first two cards. The interviewer then asked the child evaluate each part of the sentence and evaluate the sentence as a whole. Feedback was provided only if the child answered incorrectly. This procedure was repeated for each statement type and took approximately 10 minutes per child.

Necessity Condition. The procedure in the necessity condition was identical to the mapping condition with the addition of two probe segments after evaluating the second and third evidence cards. After the second evidence card, the two evidence cards (and the truth-values associated with them) were reviewed to determine if their contents changed the truth-values of the statement as a whole. After the third evidence card, all cards were reviewed. The child was then asked whether the evidence changed the statement's truth-value. If the evidence was necessary, then the child was told that it was necessary to first see evidence before evaluating "*this type*" of statement. If evidence was unnecessary, then the child was told that no evidence would change the truth-value of "*this type*" of statement because, for example, the two clauses in a contradiction can never match the same evidence (in this case the same picture).

Session 3: Training 2

The training 2 was identical to that used in Training 1.

Session 4: Post-test

The post-test procedure was identical to the pre-test.

Materials

Pre- and Post-Test Materials. Pre- and post-test materials were logically identical but had slightly different content. Each session required the evaluation of four statement types: tautology, contradiction, disjunction, and conjunction. Each child saw four instances of each statement type. The materials consisted of 16 unlined 3 x 9" cards with one statement and 16 3 x 5" cards with a corresponding piece of evidence. An additional six cards (3 statement, 3 evidence) were used as materials for the warm-up items.

Control Condition. Two sets of materials (one for each training session) were created for the control condition. Each set contained 12 statements (3 of each type) printed on 3 x 9" cards and 12 3 x 5" evidence cards. Each set also included three statement cards and three evidence cards for the warm-up segment.

Training Warm-up Condition. The training warm-up (for mapping and necessity conditions only) included two statement cards used to demonstrate the parts of the sentence, plus two statement cards and four evidence cards for articulating the AND and OR rules.

Mapping and Necessity Conditions. The mapping condition required two sets of materials, one for each training session. Each set of materials included four statement cards (one of each statement type) and twelve evidence cards, three evidence cards corresponding to each statement card.

Results

The results will be presented in two sections: (1) aggregated analyses of the effects of training across each processing phase separately and (2) individual strategy analyses comparing changes in the consistency of response patterns across processing phases. The first set of analyses will separately examines training effects on a priori evaluations, evidence requests, and evidence evaluations. The second analysis examines changes in inter-phase consistency before and after training.

Aggregated analyses

Question Phase 1: A Priori Evaluations (AP). For all *a priori* evaluation phases, possible responses were "true," "not true," or "can't tell." Correct responses were coded as follows: contradictions- not true; tautologies-true; conjunctions and disjunctions-can't tell. Correct responses were scored as 1 while incorrect responses were coded as 0.

To determine the effectiveness of training on the number of correct a priori evaluations, a 3 (condition) x 2 (age) x 2 (pre-test vs. post-test) ANOVA was performed with session as a within-subjects variable. The analysis reveals a main effect for condition, F (2, 110) = 6.1, p <.003, and age F (1, 110) = 8.9, p <.003 and no interaction between condition and age F (2, 110) = .42, p >.65. Children in the mapping and necessity conditions gave significantly more correct responses in the post-test than in the pre-test. The performance of children in the control condition did not differ significantly from pre-to post-test. Fifth graders gave significantly more correct requests in the post-test than third graders.

Question Phase 2: Evidence Requests (ER). For the evidence request phase, possible responses were "yes" or "no." Correct responses were coded as follows: contradictions-and tautologies-No; conjunctions and disjunctions-Yes. Correct responses were scored as 1 while incorrect responses were coded as 0.

To determine the effectiveness of training on the number of correct evidence requests, a 3 (condition) x 2 (age) s 2 (pre-test vs. post-test) ANOVA was performed with session as a within-subjects variable. The analysis reveals a main effect for condition, F (2, 110) = 3.9, p <.01, and age F (1, 110) = 10.6, p <.001 and no interaction between condition and age F (2, 110) = .42, p >.65. Children in the mapping and necessity conditions gave significantly more correct responses in the post-test than in the pre-test. The performance of children in the control condition did not differ significantly from pre-to post-test. Fifth graders gave significantly more correct requests in the post-test than third graders.

Question Phase 3: Evidence Evaluations (EE). For the evidence evaluation phase, possible responses were "true," "not true," or "can't tell." Correct responses were coded as follows: contradictions-not true; tautologies-true; conjunctions (2 true, 2 false) and disjunctions (2 true, 2 false). Correct responses were scored as 1 and incorrect responses were coded as 0.

To determine the effectiveness of training on the number of correct evidence evaluations, a 3 (condition) x 2 (age) x 2 (pre-test vs. post-test) ANOVA was performed with session as a within-subjects variable. The analysis reveals a main effect for condition, F (2, 110) = 20.8, p <.001, and age F (1, 110) = 9.2, p <.003 and no interaction between condition and age F (2, 110) = .25, p >.77. Children in the mapping and necessity conditions gave significantly more correct responses in the post-test than in the pre-test while children in the control condition did not differ significantly from pre- to post-test. As in previous conditions, fifth graders gave significantly more correct requests in the post-test than third graders.

The aggregated analysis demonstrated that experience with statements was not sufficient to improve performance, at least not in the limited exposure provided during the training period. Training effectively improved performance for children in the mapping and necessity conditions. The next series of analyses will examine the structure of change.

Individual Analysis

The individual analysis examined the consistency of a child's correct response patterns within each processing phase and across all processing phases. For example, although the aggregated data demonstrates that fifth graders generally outperformed third graders, these data do not indicate the extent to which an individual fifth grader produced correct answers for the evidence request phase or for all question phases. We considered a pattern in which 75% of responses were correct as *consistently correct*. A pattern below 75% was considered inconsistent. Tables 1 and 2 display the number of children coded as consistently producing correct responses for each processing phase in the pre- and post-tests.

Table 1- Percentage of Third Grade Children Giving Consistent, Correct Responses Within Each Processing Phase by Condition

Condition	AP	ER	EE
Control	0 (10)	0 (19)	19 (19)
Mapping	5 (30)	8 (40)	5 (85)
Necessity	5 (37)	5 (63)	11 (79)

Note. Posttest scores are presented in parentheses.

Table 2- Percentage of Fifth Grade Children Giving Consistent, Correct Responses Within Each Processing Phase by Condition

Condition	AP	ER	EE
Control	10 (14)	19 (29)	14 (10)
Mapping	15 (59)	10 (60)	25 (80)
Necessity	26 (62)	37 (79)	26 (95)

Note. Posttest scores are presented in parentheses.

There was a significant difference between the number of consistent, correct responses in pretests and in the post-tests for third and fifth graders. Third graders produced significantly more consistent evidence evaluations $\chi 2$ (2, 40.8, $p < .001$) in the post-test than the pre-test. Fifth graders produced significantly more evidence requests $\chi 2$ (2, 10.5, $p < .003$) and significantly more consistent evidence evaluations $\chi 2$ (2, 14.5, $p < .001$) in the post-test than the pre-test.

These data were used to code the child's overall response pattern (i.e., consistency measure on each of the three processing phases) as either consistently correct (consistent, correct responses on all phases) or mixed (inconsistent on one or more phases). The results are displayed in Table 3. Children in both training conditions produced significantly more consistent correct responses than children in the control condition, third graders $\chi 2$ (2, 4.2, $p < .03$), fifth graders $\chi 2$ (2, 15.1, $p < .001$).

Table 3- Percentage of Children Giving Consistent Correct Responses by Grade and Condition

	3rd		5th	
Condition	Pre	Post	Pre	Post
Control	0	9.5	13	13
Mapping	0	20	0	59
Necessity	0	37	11	61

Discussion

The results support the dependent phase hypothesis in which changes in performance are due to gains on knowledge of assigning truth-values, which then lead to improvements in assigning modal operators. We tested three predictions with a training study in which children were given varying levels of instruction. These data indicate that (1) contra the compiled phase hypothesis, a small amount of experience with logical statements is not effective for improving performance. Exposure to the control condition was not related to increases in performance. Training was related to significant increases in performance. Specifically, (2) both training conditions were effective in improving performance but (3) training on assigning truth-values was sufficient to produce consistent, correct responses at levels roughly equal to those of the necessity condition (in which both truth-values and modal operators were trained).

Morris and Klahr (in review) examined the order in which children make consistent correct responses on each processing phase. In all cases, correct evidence evaluation always preceded correct evidence requests and *a priori* evaluations. The authors suggested that correctly evaluating evidence provided sufficient knowledge from which children could make further inferences about problem classes.

The results can be explained by extending mental models theory. In Mental Models theory (Johnson-Laird, 1983), for each problem, models are created and searched for possible solutions. Thus, to derive a valid conclusion, a set of tokens is created and searched for possible and impossible solution states. This constitutes a solution for a single instance. As currently formulated, Mental Models makes no provisions for within-subject change in processing due to learning (e.g., a child's performance on trial 100 should be different than their performance on trial 1). The *dependent phase hypothesis* suggests that this type of change should be demonstrated by improvements on each phase and in the structure of change between processing phases. As currently formulated, mental models suggests that developmental change is a function of increases in working memory. Although this is likely important, such a factor would not explain improvement derived from experience.

We suggest two possible mechanisms that may facilitate changes in performance: *discrimination* and *compilation*. After seeing a sufficient number of instances, a reasoner should form expectations about a class of similar types (e.g., determinate statements). In this way, a reasoner does not approach each instance as a new problem type. Rather, increased reasoning efficiency results from (1) eliminating redundant or unnecessary steps (Crowley, Shrager, & Siegler, 1997) and (2) making additional inferences about the class of similar problem types (Eisenstadt & Simon, 1997; Morris & Klahr, in review). For example, when reasoning about two propositional classes: indeterminate and determinate formal types, children initially confuse the two, treating both as indeterminate (Fay & Klahr, 1996; Morris & Sloutsky, 2002). By sixth grade, children begin to distinguish the two forms, but often fail to correctly determine when evidence is necessary and when unnecessary (Morris & Sloutsky, 2002).

On first exposure to a formal type, a reasoner may make errors at all phases. In evaluating a contradiction, for example, a reasoner may assert that they cannot assign a truth-value to the statement *a priori*, may request evidence and then fail to evaluate this evidence correctly. Once evidence is correctly evaluated and after several correct conclusions have been drawn, our hypothetical reasoner may then be able to infer that no evidence will change the truth-value of this particular statement. Once this inference is drawn, the reasoner may then also assert that the statement is false *a priori*. The processes described may illustrate (1) compilation, or eliminating a redundant processing step (evidence request), and (2) discrimination, or creating a new conceptual category for "statements that do not require evidence." Once this inference is made about one statement, the reasoner may generalize this to other propositions of this type: a statement of the form (A & ~A) is false.

Acknowledgements

This work was supported in part by a postdoctoral fellowship from NICHHD (HD08550) to the first author and grant HD25211 to the second author. Special thanks to Jen Schnakenberg, Anne Siegel, and Jolene Watson for data collection and coding and Corrine Zimmerman and four anonymous reviewers for editorial suggestions.

References

Braine, M., & O'Brien, D. (1991). A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review, 98*, 182-203.

Braine, M. & O'Brien, D. (Eds.). (1998). *Mental Logic*. Mahwah, NJ: Erlbaum.

Braine, M., & Rumain, B. (1981). Development of comprehension of "Or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology, 31,* 46-70.

Cheng, P., & Holyoak, K. (1985). Pragmatic Reasoning Schemas. *Cognitive Psychology, 17*, 391-416.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187-316.

Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review, 17,* 462-489.

Eisenstadt, S. A., & Simon, H. A. (1997). Logic and thought. *Minds and Machines, 7,* 365-385.

Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development, 67*(2), 689-716.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1999). Deductive Reasoning. In J. Spence (Ed.), *Annual Review of Psychology,* Vol. 50 (pp.109-135). Palo Alto, CA: Annual Reviews.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22*, 5-36.

Miller, S. A., Custer, W., & Nassau, G. (2000). Children's understanding of the necessity of logically necessary truths. *Cognitive Development, 15*, 383-403.

Morris, B.J., & Klahr, D. (in review). From Empirical to Formal: The Role of Evidence in Children's Reasoning Strategies

Morris, B. J., & Sloutsky, V. (2002). Children's solutions of logical versus empirical problems: What's missing and what develops? *Cognitive Development, 76*, 1-22.

Osherson, D., & Markman, E. (1975). Language and the ability to evaluate contradictions and tautologies. *Cognition, 3*(3), 213-226.

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

Roberts, M.J., Wood, D.J., & Gilmore, D.J. (1994). The sentence-picture verification task: Methodological and theoretical difficulties. *British Journal of Psychology, 85*, 413-432.

Ruffman, T. (1999). Children's understanding of logical inconsistency. *Child Development, 70* (4), 872-886.