

1986

# Conditional And Biconditional Interpretations Of If-then Sentences: The Role Of Content And Context

Nancy Linda Digdon

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

---

## Recommended Citation

Digdon, Nancy Linda, "Conditional And Biconditional Interpretations Of If-then Sentences: The Role Of Content And Context" (1986). *Digitized Theses*. 1492.  
<https://ir.lib.uwo.ca/digitizedtheses/1492>

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact [tadam@uwo.ca](mailto:tadam@uwo.ca), [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

The author of this thesis has granted The University of Western Ontario a non-exclusive license to reproduce and distribute copies of this thesis to users of Western Libraries. Copyright remains with the author.

Electronic theses and dissertations available in The University of Western Ontario's institutional repository (Scholarship@Western) are solely for the purpose of private study and research. They may not be copied or reproduced, except as permitted by copyright laws, without written authority of the copyright owner. Any commercial use or publication is strictly prohibited.

The original copyright license attesting to these terms and signed by the author of this thesis may be found in the original print version of the thesis, held by Western Libraries.

The thesis approval page signed by the examining committee may also be found in the original print version of the thesis held in Western Libraries.

Please contact Western Libraries for further information:

E-mail: [libadmin@uwo.ca](mailto:libadmin@uwo.ca)

Telephone: (519) 661-2111 Ext. 84796

Web site: <http://www.lib.uwo.ca/>



National Library  
of Canada

Bibliothèque nationale  
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada  
K1A 0N4

## CANADIAN THESES

## THÈSES CANADIENNES

### NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in-part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION  
HAS BEEN MICROFILMED  
EXACTLY AS RECEIVED**

### AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ  
MICROFILMÉE TELLE QUE  
NOUS L'AVONS REÇUE**

CONDITIONAL AND BICONDITIONAL INTERPRETATIONS  
OF IF-THEN SENTENCES:  
THE ROLE OF CONTENT AND CONTEXT

by

Nancy Linda Digdon

Department of Psychology

Submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

Faculty of Graduate Studies  
The University of Western Ontario  
London, Ontario  
April, 1986

© Nancy Linda Digdon 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-29492-2

## ABSTRACT

Three experiments examined content and context effects on the interpretation of if-then sentences as either conditionals or biconditionals. In Experiment 1, participants were from four grade levels, ranging from kindergarten to grade 12. All grade levels gave biconditional interpretations to sentences in which the antecedent was necessary to the consequent but gave conditional interpretations to sentences in which the antecedent was only sufficient to the consequent. In Experiment 2, subjects were from three grade levels, ranging from grade 4 to university. All grade levels interpreted the same abstract if-then sentences as conditionals in one context but as biconditionals in another context. However, university students gave more conditional interpretations than did the other participants when the abstract sentences were presented without context. In Experiment 3, university students' interpretations of if-then sentences were correlated with their ratings of the necessity of the antecedent to the consequent. In addition, conditional and biconditional if-then sentences led to different equivalence judgments and paraphrases. Considered together, the results of the three experiments provided consistent evidence that content and context are crucial factors in the interpretation of if-then sentences. These findings are discussed in relation to (a) theories of if-then reasoning, (b) age differences in reasoning, and

(c) the design of reasoning instruction programs.

## ACKNOWLEDEMENT

My progress throughout all the stages involved in completing this dissertation has been helped by the guidance and encouragement of my advisor, Dr. Allan Paivio. I especially thank him for giving me the freedom to work independently, but yet at the same time giving me support when it was needed. My committee members, Dr. Peter Denny and Dr. Michael Pressley have each been helpful in his own way. Peter has given freely of his time and has exposed me to new ideas and new literature. I appreciate his interest and contribution. I thank Mike for his critical treatment of my statistics at my departmental oral. He has certainly broadened my statistical knowledge. I would also like to thank Dr. Albert Katz and Dr. Barb Morrangiello who participated at my departmental oral. I appreciate their insightful comments and questions, as well as their support and encouragement.

I would like to acknowledge the individuals who participated as subjects in my research. I am particularly grateful to the London Board of Education and Dr. R. Stennett for allowing me to conduct my studies in the public schools. Their cooperation is greatly appreciated.

Finally, I would like to thank my husband, Michael Dawson. I appreciate him for reasons too numerous to list.



## TABLE OF CONTENTS

	Page
CERTIFICATE OF EXAMINATION .....	ii
ABSTRACT .....	iii
ACKNOWLEDGEMENTS .....	v
TABLE OF CONTENTS .....	vi
LIST OF TABLES .....	viii
LIST OF APPENDICES .....	ix
INTRODUCTION .....	1
Literature on Adults' Reasoning with If-then Sentences .....	14
Literature on Children's Reasoning with If-then Sentences .....	21
EXPERIMENT 1 .....	30
METHOD .....	32
Subjects .....	32
Materials .....	34
Procedure .....	37
RESULTS .....	39
Performance on the Logic Problems .....	40
effect of bias .....	45
effect of type of problem .....	47
Ratings of Problem Difficulty .....	51
effect of type of problem .....	53
effect of response type .....	53
DISCUSSION .....	55
EXPERIMENT 2 .....	57
METHOD .....	62
Subjects .....	62
Materials .....	63
Procedure .....	65
RESULTS AND DISCUSSION .....	66
effect of bias .....	71
effect of grade .....	75
effect of type of problem .....	78
EXPERIMENT 3 .....	80
METHOD .....	81
Subjects .....	81
Materials .....	82
Procedure .....	86
RESULTS AND DISCUSSION .....	88
Performance on the Logic Problems and Ratings of the Necessity of Antecedents to Consequents .....	88

Equivalence Judgments .....	92
effect of bias .....	92
effect of type of statement .....	95
Paraphrase Task .....	96
GENERAL DISCUSSION .....	100
Theories of If-then Reasoning .....	101
Truth Table Accounts .....	101
Inference Rule Schemata Accounts .....	104
Possible Worlds Semantic Approach .....	108
Pragmatic Reasoning Schemas .....	112
Conclusions About the Theories .....	114
Are There Age Differences in Reasoning? .....	121
Implications for Instruction .....	125
Generality of Content and .....	126
Context Effects	
REFERENCES .....	128
FOOTNOTES .....	140
APPENDICES .....	143
VITA .....	148

LIST OF TABLES.

Table	Description	Page
1	Items used in Experiment 1	35
2	Percent Frequencies of Response Alternatives to the Logic Problems Used in Experiment 1	41
3	Mean Conditional Performance on the Problems in Experiment 1	44
4	Means used in Tests of Problem and Bias in Experiment 1	48
5	Effects of Type of Problem in Experiment 1	50
6	Ratings of Problem Difficulty in Experiment 1	52
7	Percent Frequencies of Response Alternatives to the Logic Problems Used in Experiment 2	67
8	Mean Conditional Performance on the Problems in Experiment 2	70
9	Analysis of Variance for Performance on Problems in Experiment 2	72
10	Items used in Experiment 3	83
11	Averages Used in the Correlations Between Problem Type and Rating of Necessity in Experiment 3	90
12	Experiment 3 Equivalence Judgments	93
13	Responses Scored as Biconditional Paraphrases in Experiment 3	97

## LIST OF APPENDICES

Appendix	Description	Page
I	Instructions for Experiment 1	142
II	Samples of Materials used in Experiment 2	143
III	Instructions for Experiment 2	146

## INTRODUCTION

The focus of this dissertation is on reasoning with sentences that contain the connective if-then, as in, "If you mow my lawn, then I will give you \$10.00." The types of inferences that people make when reasoning with these sentences were of particular interest, together with the factors that might predict when people will draw one or another type of inference. Understanding the inferences people generate from if-then sentences is important because if-then is a central aspect of reasoning in such areas as science (If we observe M, then theory K is supported), advertising (If you use brand X, then your clothes will be clean and white), legal proceedings (If the suspect was out of town on July 8th, then he is not the murderer), and several informal everyday situations (If you sleep in, then you will be late for school).

The current investigation was motivated by two main issues. The first involved the criteria used to evaluate the appropriateness of inferences generated from if-then sentences. What is the best criterion for "correct" reasoning? The second involved age differences in reasoning with if-then sentences. Under what circumstances will there be substantial developmental differences and in what situations will they be minimal?

Most psychological studies of people's understanding of if-then sentences have been based on the assumption that

these sentences are conditionals. It is common to represent a conditional in general terms such as 'If p then q' where 'p' stands for the antecedent proposition and 'q' represents the consequent proposition. With respect to the sentence, "If the animal is a dog, then it has a tail," 'p' refers to "the animal is a dog," and 'q' to "it has a tail". According to propositional logic, a conditional is interpreted as true or false solely on the basis of the truth values of its antecedent and consequent propositions (Quine, 1982). The conditional is true when the antecedent and consequent are both true or both false, or when the antecedent is false and the consequent is true. The conditional is false when the antecedent is true and the consequent is false.

Several psychological studies have been done to determine whether people interpret if-then sentences as conditionals, in accordance with propositional logic (see Evans, 1982, for a review). Two types of tasks have been used in this endeavor: the Wason selection and evaluation tasks (Wason & Johnson-Laird, 1972) and the conditional syllogism task (Taplin, 1971). The Wason tasks have been used exclusively in the studies of reasoning in adults whereas the conditional syllogism task has been applied both to studies of reasoning in adults and to studies of reasoning in children. Because the present investigation has a developmental perspective, only the conditional syllogism task will be discussed. Readers interested in

the Wason tasks can find an excellent discussion of the similarities and differences between these tasks and the conditional syllogism task in Marcus and Rips (1979).

The conditional syllogism task measures peoples' interpretations of if-then sentences by examining their performance on various types of problems. When a sentence is a conditional, a certain pattern of performance is predicted on four classic logic problems: modus ponens, modus tollens, inversion, and conversion. Each of these problems consists of two premises followed by a conclusion that is either valid or invalid on the basis of the premise information.

The first premise for all four logic problems is the conditional itself, 'if p then q'. The second premise varies from one problem type to another but it always consists of a single proposition, either 'p', 'q', or their respective negations. The second premise is then followed by a conclusion that contains the other proposition not mentioned in the second premise. For example, if the second premise was 'p' or 'not-p' (the negation of p), then the conclusion would be 'q' or 'not-q'.

An example of a modus ponens problem is as follows:  
 If the animal is a dog, then it has a tail. (if p then q)  
 The animal is a dog. (p)  
 Therefore, it has a tail. (q)

This conclusion to a modus ponens problem is valid, assuming the propositional logic interpretation of if as the conditional connective.

An example of a modus tollens problem is:

If the animal is a dog, then it has a tail. (if p then q)

The animal does not have a tail. (not-q)

Therefore, it is not a dog. (not-p)

Again, this is a valid conclusion to the problem, based on the conditional norm set by propositional logic.

Inversion and conversion problems, on the other hand, do not yield single valid conclusions. The following is an example of an inversion problem:

If the animal is a dog, then it has a tail. (if p then q)

The animal is not a dog. (not-p)

Therefore, it does not have a tail. (not-q)

This is not a valid conclusion. Both 'not-q' and 'q' are possible conclusions and the solution to the problem is indeterminate.

The following is an example of a conversion problem:

If the animal is a dog, then it has a tail. (if p then q)

The animal has a tail. (q)



Therefore, it is a dog. (p)

Again, this is not a valid conclusion. Both 'p' and 'not-p' are possible and the solution to the problem is indeterminate. In summary, a conditional interpretation of an if-then sentence permits the conclusion 'q' to modus ponens problems, 'not-p' to modus tollens problems but no definite conclusion to either inversion or conversion problems.

Although logicians assumed that if-then sentences are logical conditionals, people often do not interpret them as such, but instead adopt a biconditional interpretation (cf. Wildman & Fletcher, 1979). The truth of a biconditional interpretation is based totally on the truth values of the antecedent and consequent propositions, as is the case with that of the conditional interpretation. The biconditional is true when the antecedent and consequent are both true or both false and is false when the antecedent is true and the consequent is false, or when the antecedent is false and the consequent is true.

With regards to the conditional syllogism task, a biconditional interpretation differs from a conditional interpretation with respect to performance on the conversion and inversion problems but not with respect to performance on modus ponens and modus tollens. On conversion problems, a biconditional interpretation permits

the conclusion 'p' from the two premises 'if p then q' and 'q'. On inversion problems, it permits the conclusion 'not-q' from the premises 'if p then q' and 'not-p'.

Researchers have generally assumed that biconditional interpretations of if-then sentences are examples of fallacious reasoning because propositional logic was used as the criterion for appropriate reasoning. According to propositional logic, if-then sentences are conditionals, not biconditionals. (Ennis, 1969; Inhelder & Piaget, 1958). The biconditional interpretation exists in propositional logic but it is associated with sentences of the form if and only if, not with if-then sentences. Researchers who have found that people interpret if-then sentences predominately as biconditionals rather than as conditionals have appealed to a number of factors to explain this deviation from the criterion set by propositional logic.

When adults have been studied, it was usually assumed that biconditional reasoning is used not because the adults do not correctly understand if-then sentences to be conditionals, but because adults are sensitive to invited inferences arising from the semantics, pragmatics, and context of the particular if-then sentence. For instance, consider the sentence, "If it rains, then the parade will be cancelled." This sentence often invites the inference that if it doesn't rain then the parade will not be cancelled, consistent with a biconditional rather than a conditional interpretation. For other examples, the reader

is referred to Fillenbaum (1975, 1976) for an excellent discussion of how pragmatic inferences lead adults to interpret promises and threats as biconditionals and to Legrenzi (1970) and Rips and Marcus (1977) for discussions of the role of the relationship between 'p' (the antecedent) and 'q' (the consequent) in determining whether a given if-then sentence will be interpreted as a conditional or as a biconditional. In all of these studies, the adults' use of a biconditional interpretation is explained by factors within the reasoning materials and is not thought to reflect a lack of conditional reasoning competence. In fact, the adoption of a biconditional interpretation for some sentences is thought to be adaptive because it shows a sensitivity to the intent of the utterance (see Geis & Zwicky, 1971 or Fillenbaum's, 1975, 1976 research on inducements).

This raises the question of whether if-then sentences should be evaluated exclusively in terms of a conditional interpretation since it is not always appropriate in natural language contexts. The position taken in this paper is that both conditional and biconditional interpretations of if-then sentences are appropriate and the choice of one interpretation over the other for any particular if-then sentence is determined not from the semantics of if-then, but from the context in which it is used (see Footnote 1). The context of if-then is determined by the semantics, pragmatics, and context of the

particular propositions being joined by if and then. This highlights the notion that more factors other than just the truth values of the antecedent and consequent propositions are involved in the interpretation of if-then sentences. The form of an if-then sentence is not sufficient grounds for determining its appropriate interpretation.

In contrast to the studies on adult reasoning, most studies using children as subjects have tended to attribute children's biconditional interpretations to a deviant understanding of if-then, such as a defective lexical entry for the meaning of if (Knifong, 1974; Taplin, Staudenmayer & Taddonio, 1974), or to a lack of competence with conditional logic (O'Brien & Overton, 1980, 1982; Overton, Brynes & O'Brien, 1985), or to a lack of adequate processing strategies for dealing with the conditional (Staudenmayer & Bourne, 1977).


Taplin et al. (1974) argued that the meaning of if changes from a conjunctive (simply joining the antecedent and consequent propositions by and), to a biconditional, to a conditional with increasing age. In the Taplin et al. study, the content of the sentences was not considered and only abstract materials were used. All sentences were of the form, "If there is an R, then there is a D," with various letters substituted for R and D. Thus, no semantic or pragmatic cues from the antecedent and consequent propositions were available to subjects. One may contend that the use of abstract content such as this is the best

way to study reasoning because it eliminates the effects of various semantic and pragmatic factors and lets one examine reasoning in isolation. However, a strong case can be made against generalizing from performance on abstract reasoning tasks to performance on more natural tasks. Roberge and Flexor (1979) found that adolescents' ability to solve abstract reasoning problems was not a good predictor of their reasoning performance in meaningful problem-solving situations. The content of a sentence influences the encoding of the premise information and thus alters the apparent reasoning process by controlling the nature of premises on which the reasoning is based (see Evans, 1982, 1983; Falmagne, 1975). The fact that children and adults interpret abstract if-then sentences differently is not sufficient evidence to conclude that children interpret if-then in general in a different manner than do adults, particularly given children's general disproportionately poorer performance compared to adults on any task of an abstract nature (see for example, Piaget, 1928). Young children may not have had sufficient experience with abstract or context-free situations and this factor, rather than differential meanings of if-then, may be responsible for the age differences in performance on the abstract sentences (see Hutchins, 1980 for the relevance of this line of argument to cross-cultural research on reasoning).

Instead of postulating that children give if a different linguistic interpretation as did Taplin et al.,

other developmental researchers have focussed more on processing deficits in children (O'Brien & Overton, 1980, 1982; Overton, Brynes & O'Brien, 1985; Staudenamy & Bourne, 1977). These researchers acknowledge that both young and older children respond to *if* as though it were the biconditional, but the crucial age difference is that only older children (i.e., children who have reached Piaget's stage of formal operations) can be trained to give the conditional interpretation. The assumption made by O'Brien and colleagues is that the younger child's biconditional interpretation reflects a lack of competence with propositional logic, whereas the older child's biconditional interpretation should be regarded as a false negative. The older children really are competent with propositional logic because they can be trained to use it appropriately.

On the surface of it, the position of O'Brien et al. seems to be a reasonable one. However, it becomes less plausible if one examines the details of the training procedures used. The main problem with the training procedures is that they require participants to maintain and process several bits of information simultaneously in working memory. A problem arises because there are age differences in the amount of available resources in working memory (Case, Kurland & Goldberg, 1982; Manis, Keating & Morrison, 1980). The younger children's lack of sufficient resources in working memory may have prevented them from



responding successfully to the training procedures, regardless of any postulated competence or incompetence with propositional logic. It is not clear from the O'Brien et al. studies whether younger children would be successful with training procedures which place fewer demands on working memory. The young child's apparent incompetence on the procedures of O'Brien et al. may reflect nothing more than the child's limited working memory. It may have nothing to do with incompetence in dealing with propositional logic. Children as young as 10 years have benefited from other training procedures designed to improve conditional reasoning (Lane, 1983; Lee, 1985).

When one compares the explanations given for biconditional reasoning in adults to those given for biconditional reasoning in children, one is struck with an obvious inconsistency. When adults interpret if-then sentences as biconditionals, their reasoning competence is not questioned and explanations for their performance are based on factors inherent in the if-then sentences (such as the nature of the propositions linked by if and then). However, when children interpret if-then sentences as biconditionals, it is assumed that the children's immature reasoning competence is the source of the deviant performance and no appeal to content or pragmatic factors is made (see Kuhn, 1977, and Romain, Connell, & Braine, 1983 for exceptions). This may be due in part to the

pervasive influence of Piagetian assumptions in the area of conditional reasoning (see O'Brien, Brynes & Overton, 1985 and O'Brien and Overton, 1980, 1982 for examples of Piaget's influence). According to Piaget's theory, as stated in Inhelder and Piaget (1958), biconditional interpretations of if-then sentences are instances of pre-formal reasoning and are to be expected in children younger than 11 or 12 years. No appeal to other factors is necessary.

The purpose of the present investigation was to determine whether the nature of the propositions linked by if-then is responsible for children's tendency to interpret if-then sentences as biconditionals, thus casting doubt on the conclusion that children have an invalid understanding of if-then. The studies had two major innovations. The first is the assumption that the content and context of the antecedent and consequent propositions are always crucial in determining which interpretation (i.e., conditional or biconditional) of an if-then sentence is appropriate. Earlier investigators have been reluctant to view biconditional interpretations of if-then sentences as "correct." In propositional logic, a biconditional interpretation is associated with the form, 'p if and only if q'. (Quine, 1982). However, this expression is not commonly used in the English language (McCawley, 1981; Ray & Findley, 1984). It is not surprising, then, that the concept of biconditionality is expressed through the use of if-then (Geis & Zwicky, 1971).



The second major innovation of the present studies is a reconsideration of age differences in reasoning with if-then. When the content of the propositions linked by if-then clearly supports one interpretation over the other, then no age differences in performance are predicted across a broad range of ages (i.e., from 5 years of age to adulthood). This prediction was confirmed in Experiment 1. However, when the context of if-then is ambiguous as to which interpretation is more appropriate, age differences in performance are predicted. With ambiguous contexts, adults should be more likely to give conditional interpretations than should younger subjects (i.e., 9- to 12-year-olds). This prediction was confirmed in Experiment 2. Experiment 3 examined in more detail one variable that influenced whether if-then sentences were treated as conditionals or as biconditionals. In addition, Experiment 3 examined the effectiveness of two new dependent measures at discriminating between if-then sentences interpreted as conditionals, and if-then sentences interpreted as biconditionals.

Before describing the present series of studies, it is important to examine manipulations that have been used by previous researchers to affect the likelihood of if-then sentences being treated as either conditionals or biconditionals. First, some of the literature on adult reasoning will be reviewed, followed by a discussion of studies that have included children as subjects.

Literature on Adults' Reasoning with If-then Sentences. Staudenmayer (1975) found that adults interpreted sentences with specific agents, such as the sentence, "If I turn the switch, then the light comes on," as conditionals, presumably because people infer that the antecedent, "I turn the switch," is sufficient but not necessary for the light to come on. On the other hand, adults interpreted sentences that did not contain specific agents, such as the sentence, "If the switch is turned, then the light comes on," as biconditionals, presumably because people infer that the antecedent, "the switch is turned," is both necessary and sufficient for the light to come on.

Although Staudenmayer found a statistically significant difference in responding to the two types of sentences, caution is needed in generalizing from these results. Even though subjects treated more sentences as conditionals when they had specific agents than when they did not have specific agents, 54.9% of the subjects still did not treat sentences with specific agents as conditionals. Rather, they treated them as biconditionals.

It is relevant to consider why so many subjects were not sensitive to Staudenmayer's manipulation. Consider the sentence, "If I turn the switch, then the light comes on." In order to see that this sentence is a conditional, one must acknowledge the possibility of the light coming on without "I" turning the switch. Depending on the context

or the lack of context, a person infers for the sentence, he or she may or may not be led to accept this possibility. Suppose that people infer that only information actually stated in the sentence is relevant to their reasoning and, accordingly, do not postulate that other people exist that could possibly turn the switch. In this case, "I turning the switch," is necessary for the light to come on and a biconditional interpretation is appropriate. Some of Staudenmayer's subjects may have failed to treat the sentences as conditionals because there were too many degrees of freedom in possible interpretations of the context of the sentences and these different contexts influenced whether a conditional or biconditional interpretation was deemed appropriate. If more control over the contexts of the sentences is achieved, perhaps reasoning will be more consistent across subjects and it will be easier to evaluate whether the necessary/sufficient distinction, as used by Staudenmayer, is truly a useful one. In any case, Staudenmayer's study is important because it provided an objective criterion for distinguishing between conditional and biconditional interpretations that can be applied to natural language situations.

Legrenzi (1970) used a manipulation similar to the necessary/sufficient distinction of Staudenmayer, except that Legrenzi postulated that the important component of the manipulation was whether a situation is strictly

binary. In a binary situation, only two alternatives exist and they are mutually exclusive. One alternative necessarily leads to one outcome and the other alternative leads to a different outcome. Thus, only one alternative leads to any given outcome, which is consistent with a biconditional interpretation.

The binary situation used by Legrenzi involved a concrete apparatus consisting of an inclined plane divided into two channels, one on the left and one on the right. The experimenter rolled a ball down the inclined plane and the ball entered one of the two channels. As it passed through the channel it released a microswitch which, in turn, lit either a red or a green light. The same colored light was always associated with the same channel. In this situation, 73% of adult subjects treated sentences such as, "If the ball rolls to the left, the green light is lit," as biconditionals. That is, subjects said that the only two situations that were possible were that the ball rolls to the left and lights the green light or the ball rolls to the right and lights the red light. The remaining 27% of the subjects responded that the only situation compatible with the if-then sentence was when the ball rolls to the left and lights the green lamp, and that other potential situations were either incompatible or irrelevant to the verification of the if-then sentence. As well, no subjects interpreted the if-then sentences as conditionals (i.e., that it was possible for the ball to roll right and light

the green lamp as well as for it to roll left and light the green lamp or to roll right and light the red lamp).

Rips and Marcus (1977) varied Legrenzi's task and found that the important variable causing sentences to be treated as biconditionals in Legrenzi's study was not the binary nature of the task, inasmuch as the same pattern of results was obtained in nonbinary situations. Rips and Marcus initially proposed that the perceived causal relationship between antecedents and consequents (i.e., between specific channels and certain colored lights being lit) was responsible for the high level of biconditional responding. However, this hypothesis was not confirmed when Rips and Marcus compared performance on the Legrenzi type of task to performance on two other tasks: the well-known selection task of Wason (see Wason & Johnson-Laird, 1972) in which sentences are of the form, "If there is a vowel on the front of the card, then there is an even number on the back," and a task involving descriptions of fish such as, "If a fish is red, then it has stripes." In both of these tasks, there is no causal relationship between the antecedents and the consequents. Rips and Marcus found that performance on these two tasks did not differ from that on the Legrenzi task. However, performance on all three tasks varied with the type of instructions given to subjects. The experimenter described a correlation between antecedents and consequents to some subjects but for other subjects, the two dimensions were

varied independently (i.e., the two dimensions were uncorrelated). Subjects who were led to believe that the antecedents and consequents were correlated gave relatively more biconditional responses whereas the other subjects gave more conditional responses. Rips and Marcus stressed that,

the crucial factor is not the type of relation (causal or noncausal), but rather the form of the relation believed to obtain between Antecedent and Consequent values; in other words, the function mapping the Antecedent onto the Consequent range.  
(p.205)

Rips's and Marcus's position corresponds to Staudenmayer's assertion that the perceived necessity of the antecedent to the consequent is the crucial factor in determining whether an if-then sentence will be treated as a biconditional or as a conditional. If the antecedent is necessary, then a biconditional interpretation is adopted. If, on the other hand, the antecedent is not necessary but only sufficient, then a conditional interpretation is given.

Fillenbaum (1975, 1976, 1978) showed that the intent of an if-then sentence influences whether a conditional or a biconditional interpretation is adopted. In particular, inducements (such as promises or threats) are likely to be given biconditional interpretations because the very force of these utterances is that the consequent is solely dependent upon the antecedent. Consider the following promise, "If you wash the car, then I will give you \$10.00." From this promise, it is not usual for people to

infer that they would receive the \$10.00 from me, regardless of whether they wash my car, although this inference is logically consistent with a conditional interpretation of the statement (either [p and q] or [not-p and q] are possible). On the other hand, consistent with a biconditional interpretation, people would normally infer that they will receive the \$10.00 from me, only if they wash my car.

Markovits (1984) investigated an individual difference variable that predicted the frequency with which adults gave conditional interpretations. The variable was measured by having subjects generate alternatives to a hypothetical situation. For instance, participants were told, "When David has homework to do, he gets into a bad mood. I saw David after school today and he was in a bad mood. Can you imagine what could have put David in a bad mood?" Some adults generated several alternatives whereas others only generated a few. In addition, participants were given modus ponens, modus tollens, inversion, and conversion problems to solve for if-then sentences that had nothing to do with the hypothetical situation. Markovits found that the adults who generated several alternatives to the hypothetical situation also tended to give more conditional interpretations, and fewer biconditional interpretations, than did adults who only gave a few alternatives to the hypothetical situation. Markovits concluded that adults who were good at generating several

alternatives were also good at generating alternate antecedents in the context of if-then sentences. By recognizing alternate antecedents, these subjects are then more likely to adopt a conditional interpretation because the antecedent stated in the if-then sentence is perceived as sufficient but not necessary for the corresponding consequent.

Markovits (1985) provided more evidence that performance on conditional reasoning problems is mediated by an awareness of alternate antecedents. In Markovits' (1985) study, some of the reasoning problems explicitly stated alternate antecedents, such as in the problem, "Since September, John's school has often been closed. The first time, it was because of a teacher's strike. Later on, a fault in the plumbing obliged the director to close the school for several days. Since then, classes have resumed normally. However, winter has begun and John knows that if there is a snow storm in the night then school will be closed the next day." Other problems did not state any alternatives to the antecedent mentioned in the if-then sentence, such as in the problem, "If there is a snow storm in the night then school will be closed the next day." In general, adults gave more conditional interpretations to the problems that explicitly stated alternate antecedents than they did to the other problems. The problems that explicitly stated alternate antecedents emphasized that the antecedent given in the if-then sentence was only



sufficient but not necessary for the consequent, consistent with a conditional interpretation.

In summary, the studies of adults' interpretations of if-then sentences emphasized the importance of the realization that either the antecedent of an if-then sentence is necessary to the consequent (resulting in the adoption of a biconditional interpretation) or is merely sufficient to the consequent (resulting in the choice of a conditional interpretation). This was relevant both to reasoning within specialized, novel tasks (as in Degrenzi, 1970) and to reasoning in natural language contexts (as in Fillenbaum, 1975, 1976, 1978). We will now turn to the studies that have included children as subjects.

Literature on Children's Reasoning with If-then Sentences. A common feature of all the studies reviewed in this section is a manipulation designed to increase the frequency with which children give conditional interpretations, although the specifics of this manipulation vary from study to study.

In Bucci (1978, Expt.2), the manipulation was the breadth of the predicate terms in sentences such as, "All football players are strong," or, "All birds have feathers." In the first sentence, the predicate term, "are strong," was defined as a broad predicate term because being strong is characteristic of other things besides football players. In terms of the necessary/sufficient distinction made

earlier in this paper, the subject of the sentence is sufficient for the predicate to be true but it is not necessary. The second example concerning birds and feathers was defined as a narrow predicate item because having feathers is not usually associated with anything other than birds. In terms of the necessary/sufficient distinction, being a bird is both necessary and sufficient for having feathers.

Before discussing the results of Bucci's study, it is important to note that Bucci did not use a conditional reasoning task, but rather employed a class reasoning task. The particular class reasoning task she used is similar to conditional reasoning tasks because it involved the four standard logic problems commonly used in conditional reasoning studies: modus ponens, modus tollens, inversion and conversion. The difference between a conditional reasoning task and a class reasoning task is in the form of the premises. As previously stated, conditional reasoning problems have a major premise of the form, 'if p then q,' followed by a minor premise that is one of 'p', 'not-p', 'q', or 'not-q' and a conclusion that involves the proposition not mentioned in the minor premise. Class reasoning problems, on the other hand, are of the form, All p are q. X is a p. Is X a q? Other verbs such as have and can, may be substituted for the verb to be in the premises. The class reasoning problems always include a general statement about the relationship between two classes followed by a

statement and conclusion concerning a particular entity's status in the two classes. The class reasoning problem just cited can be translated into the following conditional problem: If X is a p, then X is a q. X is a p. Is X a q? An example of a conversion problem in the form of a class reasoning task is, All p are q. X is a q. Is X a p? An example of an inversion problem is, All p are q. X is not a p. Is X a q?

A response of yes to the sample conversion problem is analogous to a biconditional interpretation; a response of maybe is analogous to a conditional interpretation. Similarly, a response of no to the sample inversion problem is analogous to a biconditional interpretation; a response of maybe is analogous to a conditional interpretation.

In Bucci's study, performance on the inversion and conversion problems differed between sentences with broad predicate terms and sentences with narrow predicate terms. Sentences with broad predicate terms were more often given a conditional interpretation than were sentences with narrow predicate terms. This was true of the performance of adults and six- to eight-year-old children, and 11- to 12-year-old children tended to perform accordingly. These results suggest that even six- to eight-year-old children treat problems of the same form (e.g., two conversion problems) differently, depending on the specific content of the problems (whether they have narrow predicate terms or broad predicate terms).

Some caution is needed, however, in generalizing the results of Bucci's study on class reasoning to conditional reasoning tasks. Although class and conditional reasoning tasks both involve the same standard logic problems, children usually do not perform as well on conditional reasoning tasks as they do on class reasoning tasks (Roberge and Paulus, 1971) and this performance difference interacts both with age and with the type of content used in the problems (Roberge & Paulus, 1971). Class and conditional reasoning are often regarded as separate reasoning dimensions (cf. Ennis, 1978). Thus, if the results of Bucci's class reasoning study are to be generalized to conditional reasoning, then her study should be replicated using problems of the form used in conditional reasoning tasks.

Kuhn (1977) examined conditional and class reasoning in children from grades 1 through 4 and in general found that her subjects gave more conditional responses compared to subjects in other studies, such as in Ennis, Finkelstein, Smith and Wilson (1969), Roberge and Paulus (1971), and Taplin et al. (1974). Kuhn attributed her subjects' higher levels of conditional responding to the materials she used. She used "a simple conversational context in which the conditional interpretation would be the most natural one." (p.345) However, she did not specify why the conditional interpretation would be the most natural one and it is not clear what specific

characteristics of her materials led to relatively more conditional responding. The following is an example of the type of material she used:

I am going to tell you something else about the city of Tundor [experimenter points to Tundor photograph]. All of the people of Tundor are happy. Here is a picture of Jean [photograph of a girl is placed adjacent to Tundor photograph]. Jean lives in Tundor. Now tell me the right answer, yes, no or maybe. Is Jean happy? Let me say it once more. All the people in Tundor [points] are happy. Jean [points] lives in Tundor. Is Jean happy?

It is unclear what factors or combination of factors are responsible for the increased level of conditional responding to Kuhn's materials. Did the concreteness of her task and the use of pictures play a role? Task concreteness has been shown to improve the reasoning of children ranging in age from grade 1 to grade 3 (Kodroff & Roberge, 1975). Did the fact that the problems were presented in the form of class reasoning rather than conditional reasoning cause children to perform better on her task (cf. Roberge & Paulus, 1971)? As well, another possibility is the specific relationship between the antecedents and the consequents used in Kuhn's study. Perhaps the antecedents used in this study were more likely to be perceived as sufficient but not necessary for the consequents because of sentence content. When an antecedent is perceived as only sufficient but not necessary for a consequent, a conditional interpretation is supported. Unfortunately, Kuhn (1977) did not provide enough details to permit an evaluation of these alternative explanations.

Rumain, Connell and Braine (1983) were more explicit about the manipulations they used to affect the level of conditional responding. In the first of two studies, they found that children as young as ten years interpret if-then sentences as conditionals, provided that the sentences are presented with expanded premises rather than with simple premises. An example of an expanded premise is, "If the bottom has 7, then the top has B. But if the bottom doesn't have 7, then the top may have B, or it may have some other letter. And if the top has B, then the bottom may have 7 or it may have some other number." These premises explicitly state that the antecedent (e.g., the top has 7) is merely sufficient but not necessary for the occurrence of the consequent (e.g., the top has B). Thus, the expanded premises emphasize the conditional interpretation. The simple premise counterpart to this example is, "If the bottom has 7, then the top has B." Simple premises do not suggest that the consequent can occur with other antecedents and accordingly, they should not be as likely to be given conditional interpretations as are the expanded premises.

In general, ten-year-old children and adults gave more conditional interpretations to the expanded premises than they did to the simple premises. There was a problem, though, in the use of the expanded premises with the children. Over 50% of the children made errors on a comprehension test of the premises. The wording of the

premises caused the children difficulty and may have depressed their performance (and that of some of the adults as well, inasmuch as 37.5 % of the adults made errors on the comprehension test).

Another problem with the study by Romain et al. is in the interpretation of the results. The expanded premises explicitly stated that the antecedent was merely sufficient but not necessary for the occurrence of the consequent. These premises, then, may produce higher rates of conditional responding simply because subjects remember that the premises stated that the consequent could occur without the antecedent. Thus, the increased rate of conditional responding with expanded premises may be due to a memory phenomenon alone, rather than to an improvement in making the appropriate inferences.

This interpretational problem was overcome in Experiment 2 of Romain et al. (1983). In this second study, as in the first experiment, two types of premises were presented: complex premises and simple premises. An example of a complex premise is, "If there is a pig in the box, then there is an apple in the box. If there is a dog in the box, then there is an orange in the box. If there is a tiger in the box, then there is an orange in the box." The important feature of complex premises is that a consequent, such as having an orange in the box, is associated with more than one antecedent, thus supporting a conditional interpretation. In addition, unlike the

expanded premises used in Experiment 1, the complex premises did not explicitly state that the consequent could occur with alternative antecedents. This information was only implicit in the complex premises and simply remembering these premises would not lead one automatically to a conditional interpretation, as it might with the expanded premises used in Experiment 1. As in Experiment 1, the simple premises did not allude to the possibility that the consequent could occur with other antecedents. An example of a simple premise used in Experiment 2 is, "If there is a duck in the box, then there is a peach in the box."

Compared to performance with simple premises, complex premises led to more conditional responding in adults, ten-year-olds and seven-year-olds. These results can be interpreted in light of the necessary/sufficient distinction mentioned earlier in this paper. In the sample complex premise, having a dog in the box was sufficient for having an orange in the box, but it can be inferred that it is not necessary because, if there is a tiger in the box, then there is an orange in the box. In the single premise example, however, having a duck in the box might be interpreted as necessary to having a peach in the box because no other conditions for having a peach in the box have been mentioned. Thus, although not mentioned by Romain et al., the necessary/sufficient distinction made by Staudenmayer may be relevant to their results.



The results of Romain et al. for both Experiments 1 and 2, clearly support the position that children are capable of giving conditional interpretations to if-then sentences, but it is unclear whether these results can be generalized to children's comprehension of if-then sentences in everyday language or whether the results are peculiar to the particular reasoning tasks used by Romain et al. In both experiments, the reasoning tasks contained clues to the conditional interpretation, and it is not clear whether the children would have given as many conditional responses without this prompting. If-then sentences occurring in everyday language rarely are accompanied by such obvious clues to a conditional interpretation. Rather, if-then sentences are usually simple premises such as, "If your shoelaces are untied, then you will trip." With respect to this example, in order to generate a conditional interpretation, one must consider the possibility that the shoes of a person who trips may be tied or untied. This information can be discovered by searching one's general knowledge about tripping and recalling that people trip for other reasons besides untied shoelaces (i.e., there are alternative antecedents for the consequent). From a cognitive perspective, for example, attaining this information is effortful (Hasher & Zacks, 1979), and as well, it requires the person to attempt deliberately to search for it. Children may have performed better than usual on Romain et al.'s tasks because the demands in these tasks were reduced

by not requiring subjects to generate alternative antecedents. This may have particularly affected the young children's performances because young children often fail to perform a cognitive activity spontaneously, although they can do the activity if given some prompting (cf. Guttman, Levin, & Pressley, 1977). Romain et al.'s results may overestimate children's performance with if-then sentences that occur in everyday language without the presence of prompts.

#### EXPERIMENT 1

Experiment 1 of the current investigation extended the experiments of Romain et al. to more natural language materials and in it children were not directly given clues to the possibility of the consequent occurring with alternative antecedents. Thus, the present study had high ecological validity, as did the study reported by Kuhn (1977), but the present one was more explicit about the manipulation used to affect the likelihood of if-then sentences being treated as conditionals. The present study incorporated Staudenmayer's necessary/sufficient distinction by building it into the stimulus materials through the content of the antecedent and consequent propositions.

Some of the items used in Experiment 1 were similar to those used by Bucci (narrow vs. broad predicate terms) but the items were presented in a conditional reasoning task, rather than in a class reasoning task. In addition,

Bucci's items all entailed noncausal relationships between the subjects and the predicates. The noncausal relationship used in Bucci's items was one in which the predicate described a particular characteristic of the subject, as in the examples, "All birds have feathers," and "All football players are strong." Some of the items in the present study entailed noncausal relationships whereas others involved causal ones, such as in the example, "If it is hot, then popsicles melt." A mix of causal and noncausal items was used because conditionals cited in the literature include both types of relationships (see for example, Rips and Marcus, 1977), as do conditionals that are common in everyday language, and it is not clear whether the same factors (such as the necessary/sufficient distinction) are relevant to performance with all the various types of items (cf. Roberge, 1982).

A mix of causal and noncausal items was included in two sets of items. In one set, the items were assumed to bias a conditional interpretation because it could be inferred that the antecedent was sufficient but not necessary for the consequent to be true, as in the examples, "If an animal is a dog, then it has a tail," or, "If Tom delivers papers, then Tom earns money." In the other set, the items were assumed to bias a biconditional interpretation because it could be inferred that the antecedent was both necessary and sufficient for the consequent to be true, as in the examples, "If an animal is

a bird, then it has feathers," and "If there is air, then people can breathe." The major prediction of Experiment 1 was that subjects should give more conditional interpretations to items designed to bias a conditional interpretation than to items designed to bias a biconditional interpretation, despite the fact that the two types of items have identical logical forms. It was proposed that the content of the propositions connected by if and then and the inferences that these propositions invite influences whether people interpret if-then sentences as either conditionals or biconditionals.

The hypothesis was tested using four groups of subjects ranging from kindergarten to grade 12. The age range was broader than the range sampled by Bucci (1978), Kuhn (1977), or Romain et al. (1983).

#### Method

##### Subjects

The subjects were students enrolled in the London Ontario public school system. Parental permission forms were given to children in two classes at each of the following grade levels: kindergarten, grade 4, grade 8, and grade 12. Of the children whose parents granted them permission to participate in the study, 20 were randomly selected from each grade level to serve as subjects. From each of these groups of 20 subjects, a group of 10 randomly selected participants received if-then sentences assumed to bias a conditional interpretation and the other students

received if-then sentences assumed to bias a biconditional interpretation. The mean age of the children was 5 years 6 months for the kindergarteners, 9 years 5 months for the grade 4 students, 13 years 5 months for the grade 8 students and 17 years 2 months for the grade 12 students. The corresponding standard deviations were 3 months for the kindergarten sample, and 4 months for each of the other grade samples.

Five males and 5 females were in each grade by bias condition except in the group of kindergarten children who received the biconditional materials (6 boys and 4 girls), the grade 4 students who received the conditional materials (4 boys and 6 girls), and the grade 12 students who received the conditional materials (6 boys and 4 girls) or the biconditional materials (4 boys and 6 girls). Sex was not included as a variable in the current study because there were consistent previous findings of no sex differences in conditional reasoning (Meehan, 1984; Roberge & Mason, 1978).

It should be noted that a total of 27 rather than 20 kindergarten children was sampled. The data of seven of these children were eliminated from the study because of an obvious response bias in that they responded "yes" to all questions. Of the seven, two children were dropped from the group that received the if-then sentences assumed to bias a conditional interpretation; the other five were excluded from the group that received the if-then sentences

assumed to bias a biconditional interpretation. The children with response biases tended to be younger (mean age = 5 years 2 months,  $SD=3$  months) than the other kindergarten children sampled, although this trend was not statistically significant,  $X^2(1)=0.25$ ,  $p>.05$ .

### Materials

Two sets of five if-then sentences were used in the study, one set designed to bias a conditional interpretation, and the other set designed to bias a biconditional interpretation. Table 1 lists both sets of items.

It is important to note that the two sets of items were identical in logical form (i.e., all were of the form, If p then q). The items were constructed so that their antecedent and consequent propositions would bias either a conditional or biconditional interpretation. Consider the following two examples: "If an animal is a dog, then it has a tail," and, "If an animal is a bird, then it has feathers." It can be inferred from the first example, that being a dog is sufficient for having a tail but it is not necessary because many other animals have tails. Thus, a conditional interpretation of the sentence is appropriate. However, being a bird is usually both necessary and sufficient for having feathers and in this case, a biconditional interpretation is most appropriate. Stated in general terms, for any statement, 'If p then q', a conditional interpretation will be biased if there are

Table 1

## Items used in Experiment 1

Items Assumed to Bias a Conditional Interpretation

- 1.) If Tom delivers papers, then he earns money.
- 2.) If Mary is a baby, then she drinks milk.
- 3.) If an animal is a dog, then it has a tail.
- 4.) If the boy plays hockey, then the boy is a good skater.
- 5.) If today is Monday, then there is school.

Items Assumed to Bias a Biconditional Interpretation

- 1.) If it is hot, then popsicles melt.
- 2.) If an animal is a bird, then it has feathers.
- 3.) If there is air, then people can breathe.
- 4.) If there is lightning, then there is thunder.
- 5.) If a plant is cared for properly, then it grows.

obvious examples of 'not-p' and 'q' both being true, whereas a biconditional interpretation will be biased if there are no obvious examples of 'not-p' and 'q' both being true.

The effectiveness of the biasing procedure was pretested using a group of ten adults, including graduate students, research assistants, and secretaries. All ten pilot subjects were given both sets of items, and as a group proved to be sensitive to the manipulation. The mean number of conditional responses to the items assumed to bias a conditional interpretation was 4 (SD=1.63) on the inversion problems and 3.3 (SD=1.77) on the conversion problems. The mean number of conditional responses to the items assumed to bias a biconditional interpretation was 2.1 (SD=1.91) on the inversion problems and 1.9 (SD=2.23) on the conversion problems. Subjects gave more conditional responses to the items assumed to bias a conditional interpretation than they did to the items assumed to bias a biconditional interpretation,  $t(9)=2.375$ ,  $p<.05$  and  $t(9)=2.190$ ,  $p<.05$  for inversion and conversion problems respectively.

A major concern in the design of the items was their vocabulary. All words used in the if-then sentences referred to concepts that should be familiar to both the young and older children. Because subjects were not required to read the reasoning problems, it was not necessary to ensure that the words were at the basal



reading level of the youngest subjects, as long as the concepts were familiar to the children.

Each particular if-then sentence occurred four times in the materials, once each in the modus ponens, modus tollens, conversion, and inversion problems. In each type of problem, only affirmative conclusions were included because of the disproportionately greater difficulty encountered by younger subjects when dealing with negatives (Wildman & Fletcher, 1977).

#### Procedure

Each participant was tested individually in a private room in his or her school and received only one of the sets of materials (i.e., either the materials assumed to bias a conditional interpretation or the materials assumed to bias a biconditional interpretation) in order to ensure that the task was short enough to be within the attention spans of the youngest children. Most subjects required approximately 10 minutes to complete the entire task.

All participants, regardless of grade, were first read the instructions that appear in Appendix I. According to the instructions, children were told two things and then were asked a question to which they could respond yes, no, maybe or, don't know. The meaning of these response alternatives was explicitly defined in the instructions. In previous studies, the maybe and don't know response alternatives were often combined under one category, such

as a category of can't tell. These two responses were kept separate in the present study because a response of can't tell can be interpreted as meaning that the conclusion is indeterminate according to the premise information, or that the subject is uncertain about which answer to choose. Answers for the former reason are of interest in the present study and it is important to have a measure of them that is not confounded with answers for the latter reason.

After children were read the instructions, they were presented the problems in a blocked random fashion. The four problems (modus ponens, modus tollens, conversion, and inversion) corresponding to the same if-then sentence were presented in a block. The order of problems within a block was random, as was the ordering of the blocks. All problems were orally presented by the experimenter, who re-read any problems if necessary, although it was required in only a few instances.

After subjects responded to a problem, they were asked why they gave the answer they did. This question was included in an attempt to discover more about how the children were solving the problems. As well, it was a check to see whether the children were making similar inferences to those intended by the design of the materials.

After subjects answered this question, they were then asked to rate how difficult that particular problem was for them to solve. They chose their ratings from the following alternatives: very easy, somewhat easy, somewhat difficult, or very difficult. After rating the problem's difficulty, children went on to the next problem. Approximately half of the kindergarten and grade 4 children did not give ratings for the problems. Often these children would just shrug their shoulders and would not give a verbal response when asked for ratings. If a child did not respond within 10 seconds, the experimenter went on to the next problem. In addition, if a child failed to respond to four consecutive problems, he or she was not asked to rate subsequent problems. Perhaps the younger children's reluctance to rate the difficulty of solving the problems was related to deficits in their metaknowledge about solving problems and/or to their failure to monitor how they thought they were performing while engaged in the task (cf. Schmidt, Schmidt & Tomalis, 1984).

No feedback was given after any of the problems, but after the last problem, children were thanked for their participation and were told that they had done very well.

### Results

The findings generally supported the major prediction of the study in that participants gave significantly more conditional interpretations to the conditionally biased items than to the biconditionally biased items. This was

true of the performance of subjects at all of the grade levels. The details of the results are presented in two sections. The first section, deals with performance on the logic problems and the second focuses on subjects' ratings of problem difficulty. Subjects' explanations for their responses to the logic problems are described in Footnote 2.

#### Performance on the Logic Problems

Responses to the logic problems are shown in Table 2 and were scored according to a conditional interpretation. A score of 1 was given each response of yes on modus ponens problems, each response of no on modus tollens problems, each response of maybe or no on conversion problems, and each response of maybe or yes on inversion problems. All other responses were given a score of 0. Footnote 3 describes the rationale used to determine this scoring scheme.

Total scores were calculated for each participant for each type of problem, with 5 being the maximum possible subject score per problem type. These scores were then used as the dependent variable in a 4x2x4 split-plot analysis of variance, in which the between-subjects factors were grade (k, 4, 8, 12) and bias (conditional, biconditional), and the within-subjects factor was type of problem (modus ponens, modus tollens, conversion, inversion). Table 3 shows the mean and standard deviation for each grade, bias, and type of problem combination.

Table 2  
 Percent Frequency of Response Alternatives  
 to the Logic Problems used in Experiment 1

		Conditionally Biased Materials			
		Yes	No	Maybe	Don't Know
Grade 12	MP	86	2	12	0
	MT	2	68	28	2
	C	40	0	60	0
	I	6	30	64	0
Grade 8	MP	78	2	20	0
	MT	20	40	40	0
	C	34	2	62	2
	I	6	20	72	2
Grade 4	MP	84	0	16	0
	MT	6	72	20	2
	C	48	0	52	0
	I	2	34	64	0

Table 2 (continued)

		Yes	No	Maybe	Don't Know
Grade K	MP	78	6	10	6
	MT	18	74	6	2
	C	42	6	50	2
	I	0	52	48	0

## Biconditionally Biased Materials

		Yes	No	Maybe	Don't Know
Grade 12	MP	86	2	10	2
	MT	2	84	14	0
	C	80	2	18	0
	I	0	78	20	2
Grade 8	MP	84	4	12	0
	MT	2	76	22	0
	C	76	2	22	0
	I	2	72	26	0

Table 2 (continued)

		Yes	No	Maybe	Don't Know
Grade 4	MP	92	2	6	0
	MT	8	90	2	0
	C	94	4	2	0
	I	8	86	6	0
Grade K	MP	96	4	0	0
	MT	24	76	0	0
	C	94	4	0	2
	I	4	90	4	2

Table 3

## Mean Conditional Performance on the Problems in Experiment 1

## Items Assumed to Bias a Conditional Interpretation

Problem	Grade							
	K		4		8		12	
	M	SD	M	SD	M	SD	M	SD
modus ponens	3.90	1.20	4.20	1.03	3.90	0.99	4.30	1.34
modus tollens	3.70	1.01	3.60	1.17	2.00	1.25	3.40	1.58
conversion	2.80	1.81	2.60	2.01	3.20	1.48	3.00	1.83
inversion	2.40	1.78	3.30	1.95	3.90	0.99	3.50	1.96

## Items Assumed to Bias a Biconditional Interpretation

Problem	Grade							
	K		4		8		12	
	M	SD	M	SD	M	SD	M	SD
modus ponens	4.80	0.42	4.60	0.52	4.20	1.32	4.30	0.95
modus tollens	3.80	1.03	4.50	0.71	3.80	1.03	4.20	1.03
conversion	0.20	0.42	0.30	0.48	1.20	1.55	1.00	1.05
inversion	0.40	0.70	0.70	0.67	1.40	1.51	1.00	1.47

Note: Maximum score=5



The F-ratios to test the main effect of type of problem and the interactions involving type of problem were evaluated with a conservative Greenhouse-Geisser correction for their degrees of freedom. This was a control for the positive bias associated with F tests of repeated measures (Kirk, 1968). All other F-ratios were evaluated using conventional degrees of freedom. First the results of this general analysis are summarized and then the details of the specific effects are described.

The analysis of variance indicated a significant main effect for bias,  $F(1,72)=51.31$ ,  $p<.001$  and for type of problem,  $F(1.61,115)=63.40$ ,  $p<.001$ , but not for grade,  $F(3,72)=0.58$ ,  $p=.630$ . As well, there was a significant bias x problem interaction ( $F(1.61,115)=31.51$ ,  $p<.001$ ) but the grade x bias, grade x problem, and grade x bias x problem interactions were not significant,  $F(3,72)=0.40$ ,  $p=.753$ ,  $F(4.82, 115)=2.04$ ,  $p=.081$ ,  $F(4.82, 115)=0.73$ ,  $p=0.600$  respectively.

Effect of bias. Because there was a significant bias x problem interaction, the effect of bias was examined separately for each of the four problem types, collapsing over grade. Each comparison was done using the Scheffé post-hoc procedure (Kirk, 1968) and  $\alpha=.01$ . The compared means and their corresponding standard deviations are presented in Table 4. The most important finding from these comparisons was that on both inversion and conversion problems, children who received the conditionally biased

items gave more conditional responses than children who received the biconditionally biased items, consistent with the major prediction of this study. Footnote 4 shows that this bias effect is not merely due to the low variability in the biconditional conditions.

There was no effect of bias on modus ponens problems. In both of the bias conditions, performance was at a near ceiling level, consistent with the perfect or almost perfect performance on these problems reported in other studies (see for example, Rumain et al., 1983).

On modus tollens problems, children who received the biconditionally biased items scored higher than did children who received the conditionally biased items. This finding was somewhat puzzling because the content manipulation was not designed to affect performance on the modus tollens problems. There should not have been any bias effect on modus tollens problems. The finding of an effect of bias on modus tollens problems can be explained in terms of response biases and an overgeneralization of the maybe response (see Table 2). The tendency to overgeneralize the maybe response to modus tollens problems does not seem to be related to age. It was most prevalent in the grade 8 students of the present study but, among others, Kuhn (1977) found it in children from grades 1 to 4, Wildman and Fletcher (1977) found it in high school students and Taplin (1971) found it in adults. It seems to be more related to performance on inversion and conversion

problems than to age. For example, in the present study, 80% of the time that grade 8 students who received the conditionally biased materials answered maybe to modus tollens, they also responded maybe to the corresponding inversion and conversion problems. It is interesting that maybe responses to modus tollens were not related more to performance on inversion problems than on conversion problems. Modus tollens (i.e., reasoning from not-q to not-p) and inversion problems (i.e., reasoning from not-p to not-q) are the converse of each other. Revlis (1975) and Revlin and Leirer (1978) proposed that many of the errors observed on syllogistic reasoning tasks could be attributed to subjects' illicit conversion of the premises such as converting "All A are B;" to "All B are A." The present results suggest a more general bias to give the maybe response, rather than simply a conversion error because maybe responses on modus tollens were related equally to performance on inversion and conversion problems. The lack of evidence for conversion errors is consistent with Begg's and Harris's (1982) studies of syllogistic reasoning.

Effect of type of problem. The effect of type of problem was determined separately for each level of bias (conditional, biconditional) because of the significant problem x bias interaction. All comparisons were done using dependent t-tests with 39 degrees of freedom and the alpha level was .01 for each comparison. Table 4 shows the

Table 4  
Means Used in Tests of Type of Problem and Bias  
in Experiment 1

Problem	Bias			
	Conditional		Biconditional	
	M	SD	M	SD
modus ponens	4.08	1.14	4.48	0.80
modus tollens	3.18	1.25	4.08	0.95
conversion	2.90	1.78	0.68	0.88
inversion	3.23	1.34	0.89	1.09

individual means used in these comparisons, and Table 5 includes the  $t$ -values and significance levels. The following was the pattern of results for the materials assumed to bias a biconditional interpretation: Subjects scored higher on modus ponens problems than on either inversion or conversion problems, higher on modus tollens problems than on either inversion or conversion problems, and marginally higher ( $p=.019$ ) on modus ponens problems than on modus tollens problems. No other comparisons were significant. This pattern of results is consistent with that found in several other studies in the literature (cf. Shapiro & O'Brien, 1970; Wildman & Fletcher, 1977). In the set of items assumed to bias a conditional interpretation, participants scored higher on modus ponens problems than on either modus tollens or conversion problems. No other comparisons were significant.

Two main points can be made about the effects of type of problem. First, in both sets of items, performance on modus ponens was generally better than performance on the other types of problems. Although this is not of primary concern to the current investigation, it does add some support to Braine's (1978) assertion that the modus ponens inference is inherent in the meaning of if-then.

Second, it was revealing that there was a different pattern of problem type effects in the two sets of materials. It appears, then, that the relative difficulty of problem types is not constant, but varies instead with

Table 5  
Effects of Type of Problem in Experiment 1

Conditionally-Biased Materials

Comparison	T Value	P(2-Tail)
MP vs MT	4.04	.000
MP vs C	2.95	.005
MP vs I	2.31	.027
MT vs C	0.69	.492
MT vs I	-.06	.955
C vs I	-1.65	.108

Biconditionally-Biased Materials

Comparison	T Value	P(2-Tail)
MP vs MT	2.45	.019
MP vs C	13.56	.000
MP vs I	13.50	.000
MT vs C	12.20	.000
MT vs I	10.47	.000
C vs I	-1.36	.183

the content or bias of the reasoning materials.

#### Ratings of Problem Difficulty

Subjects' ratings of problem difficulty were given scores ranging from 1 to 4. Ratings of very easy were scored as 1, somewhat easy as 2, somewhat difficult as 3, and very difficult as 4. The ratings were totalled separately for each type of problem, with 20 being the maximum possible subject score per problem type. These totals were used as the dependent variable in a 2x2x4 split-plot analysis of variance. The between-subjects factors were grade (8, 12), and bias (conditional, biconditional). The within-subjects factor was type of problem (modus ponens, modus tollens, inversion, conversion). Ratings of the kindergarten and grade 4 children were not included in this analysis because 40 to 50% of the data were missing for these two age groups.

The F-ratios to test the main effect of problem type and the interactions involving type of problem were evaluated with a conservative Greenhouse-Geisser correction for degrees of freedom. All other F-ratios were evaluated with conventional degrees of freedom. Table 6 shows the means and standard deviations corresponding to each combination of grade, bias, and type of problem.

\* The results of the analysis of variance are as follows: There was a significant main effect of problem type,  $F(3, 108) = 7.85$ ,  $p < .000$ , but no main effect of either

Table 6  
Ratings of Problem Difficulty in Experiment I

	Grade			
	8		12	
	M	SD	M	SD
<u>Conditionally-Biased Items</u>				
Modus ponens	1.60	0.39	1.20	0.21
Modus tollens	1.78	0.15	1.54	0.51
Conversion	1.66	0.38	1.28	0.24
Inversion	1.68	0.42	1.48	0.44

Biconditionally-Biased Items

Modus ponens	1.28	0.36	1.28	0.50
Modus tollens	1.64	0.55	1.48	0.46
Conversion	1.62	0.55	1.36	0.35
Inversion	1.60	0.39	1.56	0.52

Note: Maximum possible rating is 4 (very difficult).



grade,  $E(1,36)=3.88$ ,  $p=.057$ , or bias,  $E(1,36)=0.22$ ,  $p=.642$ . None of the interactions was significant,  $F<.90$ ,  $p>.438$ .

Effect of type of problem. The details of the effect of problem type were examined using six dependent  $t$ -tests that included all possible pairwise comparisons of the four problem types. The following pattern of results was revealed: Modus ponens problems were rated as significantly easier than were all other problem types. The corresponding  $t$ -values for these comparisons were  $t(39)=-4.40$ ,  $p=.000$  for modus ponens versus modus tollens,  $t(39)=-2.63$ ,  $p=.012$  for modus ponens versus conversion, and  $t(39)=-3.60$ ,  $p=.001$  for modus ponens versus inversion. Conversion problems were rated as marginally easier than modus tollens problems,  $t(39)=-2.06$ ,  $p=.046$ , but were not rated significantly different than inversion problems,  $t(39)=-1.62$ ,  $p=.113$ . The ratings given to modus tollens and inversion problems did not differ significantly,  $t(39)=0.51$ ,  $p=.613$ .

The rating of modus ponens problems as easier than the other problem types is consistent with the subjects' problem solving performance. Not only did subjects rate modus ponens as easier to solve, but they also solved them better than the other problem types.

Effect of response type. An additional set of analyses compared ratings of problem difficulty given to maybe responses and those given to other responses (either

yes or no, depending on the type of problem): When subjects gave maybe responses, they interpreted the solution to problems as indeterminate. When participants gave either yes or no responses, they interpreted the solution to problems as determinate. The rationale for comparing these two groups of responses was to determine whether interpretation of a problem's solution as indeterminate was judged to be more difficult than interpretation of a problem's solution as determinate. Three independent *t*-tests were calculated: One compared maybe responses ( $M=1.6585$ ,  $SD=0.6328$ ) to no responses ( $M=1.4040$ ,  $SD=0.6047$ ) on inversion problems, one compared maybe responses ( $M=1.6296$ ,  $SD=0.6791$ ) to yes responses ( $M=1.3772$ ,  $SD=0.5858$ ) on conversion problems, and one compared maybe responses ( $M=1.7885$ ,  $SD=0.6955$ ) to no responses ( $M=1.5224$ ,  $SD=0.6796$ ) on modus tollens problems. No comparison was done on modus ponens problems because only a few maybe responses were given to these problems.

The result of all three comparisons was that maybe responses were rated by subjects as more difficult than the other response types. The corresponding *t*-values were:  $t(179)=2.76$ ,  $p<.001$  for inversion,  $t(193)=2.77$ ,  $p<.001$  for conversion and  $t(184)=2.38$ ,  $p<.02$  for modus tollens. These results are consistent with previous studies that have found that subjects generally perform better on problems with a determinate solution than on problems with an indeterminate solution (see Evans, 1982, for a review).

### Discussion

The most pertinent finding of Experiment 1 was the highly significant bias effect on performance on inversion and conversion problems, in accord with the major prediction of this study. These results supported the position that reasoning with if-then sentences was influenced not only by the logical form of the sentences or by the semantics of the if-then connective, both of these were constant across the two types of bias, but was also affected systematically by the content of the antecedent and consequent propositions linked by if-then. When the content of the propositions suggested that the antecedent was sufficient but not necessary to the consequent, if-then sentences were interpreted as conditionals. However, when the content of the propositions suggested that the antecedent was necessary to the consequent, if-then sentences were viewed as biconditionals. In addition, all age groups, from kindergarten to grade 12, were sensitive to the content biasing manipulation. Thus, children's as well as adults' biconditional interpretations of if-then sentences can be explained by the content of the propositions linked by if-then and by the inferences that this content suggests. If children's biconditional interpretations are to be regarded as an indication of reasoning incompetence, then adults' biconditional responding should be viewed in a similar fashion.

It is informative to consider the results of Experiment 1 in relation to other studies in the literature, specifically in relation to Bucci (1978, Expt. 2), Kuhn (1977), and Romain et al. (1983). Experiment 1 extended Bucci's finding of content effects on a class reasoning task to a conditional reasoning task. Moreover, the pattern of results across the two studies was very similar, except that the content effects were even stronger in the present case. Participants in Bucci (1978, Expt. 2) gave 17% more conditional interpretations to broad predicate items than to narrow predicate items whereas subjects in the present study gave 46.25% more conditional interpretations to items assumed to bias a conditional interpretation than to items assumed to bias a biconditional interpretation.

Differences in the degree of the effectiveness of the content manipulation were mainly due to a higher rate of conditional responding by Bucci's subjects to narrow predicate items. Some of these items were not especially "narrow," such as the item, "All birds hatch from eggs." Other animals, such as snakes and mosquitos, also hatch from eggs, thus supporting a conditional rather than a biconditional interpretation.

Experiment 1 clarified some aspects of Kuhn (1977). Kuhn suggested that the use of real-world content is needed by children in order for them to adopt a conditional interpretation, the assumption being that the

meaningfulness or the familiarity of the content per se induces more conditional interpretations. An important outcome of the present study was that it was not just the meaningfulness of the content that biased a conditional interpretation, but as well it is the perceived relationship between the antecedent and the consequent clauses that biased one interpretation over another. Concrete, meaningful content can bias a biconditional as well as a conditional interpretation, as was shown by the results of the current experiment.

The present results agree nicely with those of Romain et al. (1983). In Romain et al.'s studies, children as young as seven years gave conditional interpretations when given clues that the consequent could occur with alternate antecedents. However, it was not clear whether the children's performance would generalize to reasoning in the absence of these clues (i.e., to more naturalistic reasoning situations). The results of Experiment 1 suggest that this generalization was valid. Children gave conditional interpretations even in the absence of clues. As well, the present study extended Romain et al.'s findings to a younger sample (i.e., to 5-year-olds).

#### EXPERIMENT 2

Experiment 1 showed that children will give more conditional responses to if-then sentences when the content of the sentences suggests that the antecedent is only sufficient but not necessary for the consequent. A further

question is whether children will treat if-then sentences as conditionals when there is no content in the if-then sentence to facilitate this interpretation. The answer apparently is that they will not (see Bucci, 1978; Paris, 1973; Roberge & Flexor, 1979; Staudenmayer & Bourne, 1977; Taplin et al., 1974; Wildman & Fletcher, 1977, 1979). With abstract materials (i.e., materials in which the content is irrelevant to reasoning, as in, "If there is an A, then there is a D.") children do not usually give conditional interpretations. But then again, neither do adults (Berry, 1983; Taplin & Staudenmayer, 1973; Taplin et al., 1974). For instance, in the studies by Taplin and colleagues, none of the nine-year-old, seven-year-old or thirteen-year-old subjects consistently treated if-then sentences as conditionals and only 10% of the fifteen-year-olds and 13.1% of the seventeen-year-olds did so. As well, a mere 2.8% to 3.6% of the adult subjects consistently gave conditional interpretations.

One problem with abstract materials is that they are ambiguous because of the lack of contextual information to determine whether a conditional or a biconditional interpretation is warranted. The purpose of Experiment 2 was to examine reasoning with abstract materials when an attempt has been made to alleviate the ambiguity in these materials concerning whether a conditional or a biconditional interpretation is more appropriate.

Previous researchers may not have been sufficiently concerned about the ambiguity in abstract reasoning materials, perhaps because of the strong rationalist assumptions underlying much of this research, which has followed the Henle (1962) tradition. It was commonly accepted that logical form, as in 'If p then q', dictated reasoning by specifying valid and invalid inferences. The content and context of the propositions linked by if-then were viewed as relevant in explaining deviations in performance from given logical patterns, rather than in establishing these patterns. Thus, there was no perceived ambiguity in abstract materials because these materials, like any others, were not ambiguous in logical form. However, if one does not accept the notion that logical form is the sole criterion for the evaluation of reasoning, but emphasizes as well the role of pragmatic, semantic, and contextual criteria, then abstract materials are no longer unambiguous as to their "correct" interpretation. In natural language contexts, if-then is used sometimes to express a conditional relationship and at other times a biconditional relationship (Geis & Zwicky, 1971; Ray & Findley, 1984).

The purpose of Experiment 2 was to examine reasoning with abstract materials when the ambiguity in these materials was resolved by specifying examples that supported one interpretation over the other. The previous findings of age differences with abstract materials (eg.,

Taplin et al., 1974) may have arisen because of the ambiguity in the materials, rather than differences in understanding of if-then. In addition, the slight age differences in frequency of conditional responding are offset by the observation that at all ages, subjects give predominately biconditional interpretations and that conditional interpretations are not the norm, even for adults (Legrenzi, 1970; Marcus & Rips, 1979; Markovits, 1984; Taplin, 1971; and Taplin & Staudenmayer, 1973). Perhaps a biconditional interpretation is easier than a conditional one because the former does not involve the consideration of indeterminacy or asymmetrical relationships.

A second motivation for Experiment 2 was to extend the generality of the results of Experiment 1. One criticism of using real-world reasoning material, as was done in Experiment 1, is that people may not really "reason" from them, but instead simply rate the perceived truth of conclusions, independent of the premises. This can be checked in Experiment 1 by comparing responses given to modus ponens and inversion problems, and those given to modus tollens problems with responses given to conversion problems. Modus ponens and inversion problems contain the same conclusions, as do modus tollens and conversion problems. If people are just rating the perceived truth of conclusions then they should give the same response to modus ponens and inversion problems and they should give



the same response to modus tollens as to conversion problems. This was clearly not the case in Experiment 1, in which only 17% of responses given to modus ponens were consistent with the corresponding responses given to inversion problems and only 23.8% of responses given to modus tollens were consistent with the corresponding responses given to conversion problems. To have some basis for comparison, the average agreement within problem types was calculated to be 87% for modus ponens, 78% for modus tollens, 83% for conversion, and 82% for inversion. These measures were not directly comparable to the measures of between-problem consistency because the former involved measures of consistency across different items and the latter only included measures across identical items. In any case, the content manipulation of Experiment 1 may have been as effective as it was partly because it relied on inferences from general knowledge and these inferences are extremely compelling (see Geis & Zwicky, 1971).

In Experiment 2, no inferences from general knowledge were applicable; instead, the subjects were required to generate inferences from the abstract examples given. If the manipulation in Experiment 2 produces results consistent with those found in Experiment 1, then it provides more support for the position that the perceived relationship (i.e., necessary vs. sufficient) between the antecedent and the consequent was responsible for the striking content effect found in Experiment 1.

The biasing manipulation in Experiment 2 entailed the use of examples following each abstract problem. Each problem was of the form, "If there is an A, then there is a B," with various letters substituted for A and B. The examples consisted of pairs of letters. The crucial characteristic of these pairs was whether they included the consequent of the if-then sentence paired with letters other than the one mentioned in the antecedent of the if-then sentence. In the examples corresponding to the conditional items, consequents of the if-then sentences were paired with letters other than the antecedent of the corresponding if-then sentence. These examples were designed to bias a conditional interpretation because they highlighted the fact that the antecedent stated in the if-then sentence was merely sufficient but not necessary for the occurrence of the consequent. The consequent occurred in the examples with alternative antecedents. On the other hand, in the examples used with the biconditional items, consequents were only paired with the letter used in the antecedent of the corresponding if-then sentence, suggesting that the antecedent was necessary to the occurrence of the consequent. In the set of neutral items, no examples followed the problems and the choice of a conditional versus a biconditional interpretation was ambiguous.

#### Method

##### Subjects

The subjects included 39 grade 4 students, 39 grade 7 students, and 48 introductory psychology students. The grade 4 and grade 7 students were enrolled in the same London Ontario public school and were randomly selected from a larger pool of students whose parents had granted them permission to participate in the study. The introductory psychology students were registered at the University of Western Ontario and participated in the study for partial course credit. The mean age of participants was 9 years 8 months for the grade 4 sample, 12 years 7 months for the grade 7 sample and 19 years 6 months for the introductory psychology students. The corresponding standard deviations were 3 months, 8 months, and 15 months for the grade 4, grade 7, and introductory psychology students, respectively.

One third of the subjects in each age group were randomly chosen to receive conditionally biased items, another third to receive the biconditionally biased items, and the final third to receive the neutral items.

### Materials

Three sets of items were designed so that one set presumably biased a conditional interpretation, another biased a biconditional interpretation, and the third set contained neutral items. All items in each set were of the form, "If there is an A, then there is a B," with various letters substituted for A and B. Each set of items included five problems of each of the four types, modus

ponens, modus tollens, conversion, and inversion. All problems were typed on 4"x6" (10.16 cm x 15.24 cm) index cards, with one problem per card (see Appendix II). Each problem in one set of items was identical to the corresponding problem in the other two sets of items, but the sets of items differed in the examples that followed the problems. The examples were created to bias one interpretation over the other and these examples were the major manipulation used in Experiment 2. The examples, where relevant, always occurred at the bottom of the corresponding index card, below the problem.

In the set of neutral items, no examples followed the problems. An example of a neutral item can be found in Appendix II. In the other two sets of items, five examples followed each problem. The first example always consisted of the letters used in the if-then sentence (for the sentence, If A then B, the first example would be AB), and the last two examples were always pairs of letters other than those used in the if-then sentence (e.g., LP and XR). The crucial examples were the second and third ones. For the set of conditionally biased items, the second and third examples consisted of the consequent of the if-then sentence paired with a different antecedent (e.g., WB and FB), thus emphasizing that the antecedent stated in the if-then sentence was only sufficient but not necessary for the consequent. The second and third examples for the set of biconditionally biased items consisted of letters other

than those used in the if-then sentence (e.g., WI and EV). The consequent of the if-then sentence was not paired with alternate antecedents, suggesting that the antecedent was necessary for the consequent. Sample conditional and biconditional items can be found in Appendix II.

Several constraints were placed on the particular letters used in the examples. First the same letter was never used in consecutive items, nor were any letters repeated within an item except for the repetition of the consequent in the items assumed to bias a conditional interpretation. As well, the letters used in examples for the conditional items matched those used in the corresponding biconditional items except for the letters used in the consequents of the second and third examples.

#### Procedure

All participants were tested individually and were read the instructions presented in Appendix II. Each subject was shown the index cards corresponding to one of the sets of 20 items (i.e., either the conditionally biased, biconditionally biased, or neutral items) and the experimenter read each item aloud as the corresponding card was placed in front of the subject. One third of the subjects in each age group received the items assumed to bias a conditional interpretation, another third were presented the items assumed to bias a biconditional interpretation, and the last third received the neutral items.

All subjects were presented the items in the same blocked random order. There were five blocks of four problems, that is, one each of modus ponens, modus tollens, inversion, and conversion. The order of problems within the blocks and the order of the blocks were both random.

Participants were not given any feedback on their performance until after they had answered the last problem. At this time, they were thanked for their participation and the details of the study were explained to them.

### Results and Discussion

Subjects' responses to the logic problems are presented in Table 7 and were scored using the same criteria as in Experiment 1. All responses consistent with a conditional interpretation were scored 1. All responses inconsistent with a conditional interpretation were scored 0.

Total scores were computed for each participant for each type of problem, with 5 being the maximum possible subject score per type of problem. Subjects' total scores were used as the dependent variable in a 3x3x4 split-plot analysis of variance. The between-subjects factors in this analysis were grade (4, 7, university) and bias (conditional, neutral, biconditional). The within-subjects factor was problem type (modus ponens, modus tollens, inversion, conversion). Table 8 includes the means and standard deviations for each grade by bias by problem type.

Table 7

Percent Frequency of Response Alternatives  
to the Logic Problems used in Experiment 2

University Students and Neutral Materials

	Yes	No	Maybe	Don't Know
modus ponens	95	0	5	0
modus tollens	0	36.25	62.5	1.25
conversion	23.75	6.25	70	0
inversion	0	48.75	51.25	0

University Students and Conditional Materials

modus ponens	93.75	1.25	5	0
modus tollens	6.25	67.5	26.25	0
conversion	20	1.25	78.75	0
inversion	10	17.5	72.5	0

University Students and Biconditional Materials

modus ponens	98.75	1.25	0	0
modus tollens	3.75	75	21.25	0
conversion	76.25	2.5	21.25	0
inversion	3.75	91.25	5	0

Table 7 (continued)

Grade 7 and Neutral Materials

	Yes	No	Maybe	Don't Know
modus ponens	92.31	4.62	3.07	0
modus tollens	15.39	75.38	9.23	0
conversion	90.77	0	9.23	0
inversion	0	98.46	1.54	0

Grade 7 and Conditional Materials

modus ponens	76.92	3.08	20	0
modus tollens	9.23	76.92	13.85	0
conversion	75.38	0	23.08	1.54
inversion	.0	46.15	53.85	0

Grade 7 and Biconditional Materials

modus ponens	100	0	0	0
modus tollens	9.23	80	10.77	0
conversion	92.31	0	7.69	0
inversion	0	92.31	7.69	0



Table 7 (continued)

Grade 4 and Neutral Materials

	Yes	No	Maybe	Don't Know
modus ponens	73.85	10.77	10.77	4.61
modus tollens	35.38	46.15	13.86	4.61
conversion	84.61	0	12.31	3.08
inversion	0	87.69	10.77	1.54

Grade 4 and Conditional Materials

modus ponens	84.62	3.08	12.30	0
modus tollens	21.54	58.46	20	0
conversion	56.92	0	41.54	1.54
inversion	0	30.77	69.23	0

Grade 4 and Biconditional Materials

modus ponens	96.92	3.08	0	0
modus tollens	21.53	72.31	3.08	3.08
conversion	98.46	0	1.54	0
inversion	0	98.46	1.54	0

Table 8

Mean Conditional Performance on Problems in Experiment 2

	Grade					
	4		7		University	
	M	SD	M	SD	M	SD
<u>Neutral Materials</u>						
modus ponens	3.69	1.49	4.62	1.12	4.75	1.00
modus tollens	2.31	1.97	3.78	1.64	1.81	2.01
conversion	0.62	1.19	0.46	1.39	3.81	1.60
inversion	0.54	0.78	0.08	0.28	2.56	2.22
<u>Conditionally-Biased Materials</u>						
modus ponens	4.23	0.93	3.85	1.58	4.69	0.70
modus tollens	2.92	1.85	3.85	1.86	3.38	1.75
conversion	2.07	1.89	1.15	1.46	4.00	1.32
inversion	3.46	1.61	2.69	2.39	4.13	1.36
<u>Conditionally-Biased Materials</u>						
modus ponens	4.85	0.38	5.00	0.00	4.94	0.25
modus tollens	3.62	1.80	4.00	1.63	3.75	1.65
conversion	0.08	0.28	0.38	1.12	1.19	1.87
inversion	0.08	0.28	0.38	0.96	0.44	0.89

combination.

The  $F$ -ratios to test the main effect of type of problem and the interactions involving problem type were evaluated with a conservative Greenhouse-Geisser correction for their degrees of freedom. All other  $F$ -ratios were evaluated with conventional degrees of freedom.

The results of the analysis of variance are simple to summarize. All main effects and two-way interactions were clearly significant ( $p < .002$ ) and the three-way interaction was marginally significant ( $p = .0377$ ). Table 9 shows the  $F$ -values associated with each effect.

Effect of bias. Separate tests of bias were done for each grade by problem combination because of the significant grade x bias and problem x bias interactions. Because performing these separate tests necessitated a considerable number of comparisons, each comparison was done using the Scheffé post-hoc procedure with  $\alpha = .01$ . The Scheffé test was chosen because it is conservative and helps control for Type 1 errors (Kirk, 1968). Footnote 5 shows that the nonparametric Mann-Whitney U procedure produced the same pattern of results as did the Scheffé tests.

The Scheffé tests showed that, for all three grade levels, there was no significant effect of bias on either *modus ponens* or *modus tollens*. This finding was expected because the bias manipulation was not intended to affect

Table 9  
Analysis of Variance for Performance  
on Problems in Experiment 2

Source	df	F	p
Grade (G)	2.117	18.16	.0000
Bias (B)	2.117	29.74	.0000
Problem (P)	3.351	125.48	.0000
GxB	4.117	4.46	.0022
PxC	6.351	8.06	.0000
PxB	6.351	16.68	.0000
PxCxB	12.351	2.02	.0377

performance on either of these two types of problems. The reader will recall that biconditional and conditional interpretations predict the same response on these problems. It also provided some evidence that subjects assigned to one type of biased condition (conditional, biconditional, or neutral) did not differ on the average from participants assigned to other conditions on such factors as general reasoning ability. This finding is important because it rules out the effects of such confounding factors as general reasoning ability, which could have accounted for bias effects on performance on other problem types.

There was an interesting pattern of bias effects on both inversion and conversion problems. First, participants who received the conditionally biased items scored higher than subjects who received the biconditionally biased items. This difference was significant for comparisons at all three grade levels, except for the comparison that involved grade 7 students on conversion problems. This latter comparison failed to reach statistical significance ( $.01 < p < .05$ ), although the trend was in the predicted direction. In general, the finding of more conditional responses to conditionally biased items than to biconditionally biased items replicated the similar finding of Experiment 1, and extended them to more abstract materials and older subjects (university students).

For the set of neutral items, there was a different pattern of performance on inversion and conversion problems for the various age groups. The university students gave as many conditional interpretations to the neutral versions of these problems as they did to the conditionally biased versions, which was significantly more than they gave to the biconditionally biased versions. On the other hand, the grade 4 students gave significantly fewer conditional interpretations to the neutral items than they did to the conditionally biased items. The grade 4 students' scores on the neutral items did not differ significantly from their scores on the biconditionally biased items. The pattern of results for the grade 7 students on inversion problems was the same as that of the grade 4 students. On conversion problems, however, grade 7 students in all three of the bias conditions scored at a similar level and there were no significant bias effects.

In general, the pattern of bias effects on inversion and conversion problems using neutral items replicated earlier studies that used abstract, arbitrary, or non-thematic materials (Bucci, 1978; Taplin et al., 1974; Wildman & Fletcher, 1977, 1979). That is, with these types of materials, older subjects were more likely to give conditional interpretations than were younger subjects. However, when the reasoning materials included context that supported one interpretation over the other (i.e., as in the conditionally biased and biconditionally biased

materials), then age differences in conditional reasoning were minimal. When the context pointed to a conditional interpretation, people of all ages sampled were most likely to give a conditional interpretation (with the exception of the grade 4 and 7 students on the conversion problems). When the context pointed to a biconditional interpretation, all age groups were most likely to give a biconditional interpretation.

Effect of grade. The effect of grade was determined separately for each bias by problem combination because of the significant grade x bias and grade x problem interactions. Each comparison was done using the Scheffé post-hoc procedure and alpha was set at .01 per comparison. A conservative alpha level was chosen to control for Type I errors. The comparisons for the biconditional materials are described first, followed by those for the conditional materials, and then those for the neutral materials.

For the set of biconditionally biased materials, there was no effect of grade on any of the problem types. Participants in all grades scored relatively high on modus ponens and modus tollens and relatively low on inversion and conversion. These results were consistent with the comparable ones of Experiment 1. In Experiment 1, there were no effects of grade from kindergarten to grade 12. In Experiment 2, there were no effects of grade from grade 4 to university. These consistent findings of no grade differences in reasoning are noteworthy given the wide

range of ages compared .

For the conditionally biased materials, there was no effect of grade for any problem type except conversion, on which university students scored higher than did either the grade 4 or grade 7 students. The cause of this grade difference in performance on conversion problems is unclear. One possible explanation is that it reflects a general grade difference in the frequency of adopting a conditional interpretation with these materials. This explanation is unlikely, however, because there were no grade differences in performance on inversion problems. Differences on both inversion and conversion problems would be expected if there were differences in the frequency of conditional interpretations. As well, there were no grade differences in performance on inversion or conversion problems with the conditionally biased materials used in Experiment 1.

A more plausible explanation is that the grade differences in performance on conversion problems in Experiment 2 were due to factors within the reasoning materials and not due to differences in reasoning ability per se. The reader will recall that conditionally biased items used in Experiment 2 had five examples of pairs of letters following each item. The first example contained the antecedent and the consequent from the if-then sentence. The second and third examples consisted of the consequent from the if-then sentence paired with an



alternative antecedent and the last two examples were made up of letters other than the antecedent and the consequent of the if-then sentence. Consider the following conversion problem. If there is an A, then there is a B. There is a B. Is there an A? Examples corresponding to this problem are: AB, EB, KB, LZ, and MC. When solving this problem, subjects looked at the examples, beginning with the first one (AB). Based on the first example alone, the answer to the conversion problem is yes, which is not consistent with a conditional interpretation. Younger subjects may not have realized that, even though they had found an example consistent with one of the available answers (yes, no, maybe, or don't know), it was possible for a later example to contradict this answer. This would explain why the younger students were more likely to respond yes rather than maybe to the conversion problems, thereby accounting for the grade differences in performance.

For the set of neutral materials, there were no grade effects on performance on either modus ponens or modus tollens problems. However, on both inversion and conversion problems, university students scored significantly higher than did either the grade 7 or grade 4 students, and the differences between the groups were substantial ( $M=3.81$  and  $2.56$  for university students,  $M=0.46$  and  $0.08$  for grade 7 students, and  $M=0.62$  and  $0.54$  for grade 4 students on conversion and inversion problems, respectively). These dramatic grade differences in

performance with the neutral items contrast sharply with the minimal grade differences in performance with conditionally biased and biconditionally biased materials. Older subjects were relatively more likely than younger subjects to give a conditional interpretation to if-then sentences that lacked context. The results of Experiment 2, indicate that older subjects' higher frequencies of conditional responding, compared to that of younger subjects may be due to characteristics of the reasoning materials (i.e., to a lack of context), rather than to a difference in general reasoning ability. Whether or not grade differences occur in the frequency of conditional responding depends on the nature of reasoning materials used to measure conditional reasoning. When the context of the reasoning materials suggests a conditional interpretation over a biconditional interpretation, or vice versa, then there are few, if any, grade differences in performance.

Effect of type of problem. The effect of type of problem was determined separately for each grade by bias combination because of the significant problem x grade and problem x bias interactions. Each comparison was done using dependent t-tests with 15 degrees of freedom for the comparisons that involved the university students and 12 degrees of freedom for the comparisons that involved the other students. Alpha was set at .01 for each comparison. This section describes in turn the comparisons based on the

conditionally biased, biconditionally biased, and neutral material.

For the set of conditionally biased materials, type of problem had no effect on the performance of the university students. There were, however, effects of problem type on the performance of the grade 7 and grade 4 students. Grade 7 students scored higher on modus ponens and modus tollens than on conversion problems. Grade 4 students scored higher on modus ponens than on conversion problems and marginally higher on modus ponens than on modus tollens problems ( $p=.017$ ). These findings replicated Experiment 1.

For the set of biconditionally biased materials, university, grade 7, and grade 4 students, all scored higher on modus ponens and modus tollens than on inversion and conversion problems. In addition, university students scored higher on modus ponens than on modus tollens problems. This pattern of results replicated the comparable ones of Experiment 1 and was consistent with the results of earlier studies (c.f., Shapiro & O'Brien, 1970; Wildman & Fletcher, 1977).

For the set of neutral materials, university, grade 7, and grade 4 students, all scored higher on modus ponens than on inversion problems. Grade 7 students and grade 4 students scored higher on modus ponens than on conversion problems. University students scored marginally higher on modus ponens than on conversion problems ( $p=.014$ ). Grade 7

students had higher scores on modus tollens than on either inversion or conversion problems. University students had higher scores on modus ponens than on modus tollens problems.

In summary, the major point about the effects of type of problem was that performance was consistently better on modus ponens than on the other problem types, regardless of the type of reasoning material or the age of the reasoner. This finding replicated Experiment 1 and provided support for Braine's (1978) assertion that modus ponens is of central importance to the interpretation of if-then.

### EXPERIMENT 3

The third experiment had two main purposes. The first was to provide more evidence for the hypothesis that the bias effects found in Experiments 1 and 2 were attributable to subjects' realization that either the antecedent of an if-then sentence is necessary to the consequent, thereby resulting in the adoption of a biconditional interpretation, or is merely sufficient to the consequent, resulting in the choice of a conditional interpretation. Experiment 3 tested this hypothesis directly by having participants rate the necessity of the antecedent to the consequent and then correlating these ratings with subjects' performances on the logic problems. Strong correlations between the ratings and performance on inversion and conversion problems were predicted because these problems discriminate between conditional and

biconditional interpretations. On the other hand, the ratings were not predicted to correlate with performance on modus ponens or modus tollens problems because these problems do not discriminate between conditional and biconditional interpretations.

The second purpose of Experiment 3 was to examine how conditionally-interpreted and biconditionally-interpreted if-then sentences differ on other dependent measures besides performance on the logic problems. Two other dependent measures were considered, namely, subjects' paraphrases of the if-then sentences and subjects' judgments of the equivalence of if-then sentences to various alternative expressions. These measures were similar to ones used by Fillenbaum (1975, 1976, 1978) to discriminate between different types of inducements.

#### Method

##### Subjects

Forty-nine introductory psychology students participated in Experiment 3. They were all registered at the University of Western Ontario and participated in the study for partial course credit. The mean age of participants was 19 years, 7 months and the standard deviation was 3 years 3 months.

All participants were given the logic problems to solve and were asked to rate the necessity of the antecedent of an if-then sentence to the corresponding

consequent. As well, 15 of the subjects wrote the paraphrases and the other 34 participants made the equivalence judgments.

Only adult subjects were sampled because the tasks necessitated the use of fairly sophisticated verbal skills, in the case of the paraphrases, and required subjects to give subjective ratings. Recall from Experiment 1 that the younger subjects encountered substantial difficulty when asked to make ratings of problem difficulty. It was felt that ratings of perceived necessity and equivalence judgments in the present experiment would also be too difficult for children.

### Materials

The 20 if-then sentences presented in Table 10 were used in this study. These sentences included a variety of relationships between their antecedents and consequents including: (a) causal relationships, as in the sentence, "If the car is out of gas, then the car will stall"; (b) whole-part relationships (e.g., "If the tree is an oak, then the tree has acorns"); (c) relationships of inducement (e.g., "If you mow the lawn, then I will give you \$5.00"); and (d) relationships describing a universal truth (e.g., "If you are my daughter, then I am your father"). In addition, 10 of the sentences were assumed to bias a conditional interpretation because the consequent was possible with alternative antecedents, as in the sentence, "If the building is a school, then the building

TABLE 10

## Items used in Experiment 3

- 1) If you do the dishes, then I will take the garbage out.
- 2) If you mow the lawn, then I will give you \$5.00.
- 3) If you throw a temper tantrum, then you will be punished.
- 4) If the houseplant is cared for properly, then it grows.
- 5) If the forests are sprayed with DDT, then the birds will die.
- 6) If iron is heated in fire, then it turns red.
- 7) If the car is out of gas, then it will stall.
- 8) If Fred sleeps in, then he will be late for school.
- 9) If Mary delivers papers, then she earns money.
- 10) If the animal is a bird, then it has feathers.
- 11) If the tree is an oak, then it has acorns.
- 12) If the animal is a fish, then it has gills.
- 13) If the animal is a cat, then it has a tail.
- 14) If the tree is a spruce, then it has needles.

## . Table 10 (continued)

- 15) If the building is a school, then it has windows.
- 16) If today is Sunday, then tomorrow is Monday.
- 17) If the basketball team gets a basket, then they score two points.
- 18) If Tom is Mary's husband, then Mary is Tom's wife.
- 19) If today is Christmas, then it is December.
- 20) If you are my daughter, then I am your father.



has windows." The other 10 sentences were assumed to bias a biconditional interpretation because the stated antecedent is usually perceived as necessary to the consequent, such as "If the animal is a fish, then the animal has gills."

The same 20 sentences were used in all parts of Experiment 3, including the logic problems, the ratings of the necessity of antecedents to consequents, the paraphrase task, and the equivalence judgment task.

For the logic problems, each if-then sentence occurred once in a modus ponens, once in a modus tollens, once in a conversion, and once in an inversion problem.

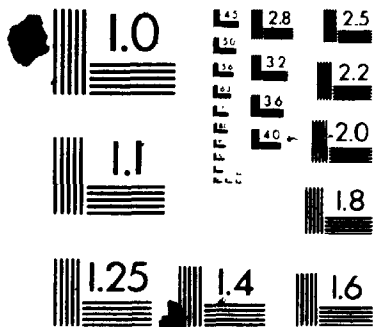
For the equivalence judgment task, each if-then sentence was followed by five statements, each of which described a relationship between the antecedent and the consequent of that particular if-then sentence. Consider the general if-then sentence, if p then q. The five statements that followed this sentence were: 'if q then p' (labelled the biconditional-reverse), '(p and q) or (not-p and not-q)' (labelled the biconditional-or), 'not-p or q' (labelled the conditional-or), 'not-p and q' (labelled the conditional-and), and 'only if p then q' (labelled the biconditional-only-if). The labels in parentheses were chosen by the author and are not the standard logical labels for the expressions. The first word in each label, i.e., the word conditional or biconditional, refers to the

2

OF/DE

2

MICROCOPY RESOLUTION TEST CHART  
NBS 1010a  
ANSI and ISO TEST CHART No. 2



interpretation that the expression is consistent with. The next word or words in the label, i.e., reverse, and, or, or only if, are just descriptions that discriminate the three biconditional expressions from each other, and distinguish between the two conditional expressions. The five expressions will be referred to by these labels in subsequent sections of this paper.

### Procedure

Each participant was tested individually and completed the logic problems task first. At the beginning of this task, subjects were read the instructions that were used in Experiment 1 (see Appendix I). The logic problems were then presented in a blocked random fashion. The four problems (modus ponens, modus tollens, inversion, and conversion) corresponding to the same if-then sentence were presented in a block. Each problem was presented on a separate page. The order of problems within a block and the order of the blocks were random. Participants completed the logic problems at their own pace.

After completing the logic problems task, subjects rated each of the if-then sentences. All participants were asked to rate on a scale from 1 to 10 the necessity of the antecedent of an if-then sentence to the consequent (i.e., the likelihood that the consequent could occur with alternative antecedents). A rating of 1 indicated that the antecedent was necessary to the consequent (i.e., it was very unlikely that the consequent could occur with

alternative antecedents). A rating of 10 corresponded to the antecedent not being necessary for the consequent (i.e., it was very likely that the consequent could occur with alternative antecedents).

After subjects gave a rating for an if-then sentence, they were asked either to paraphrase the if-then sentence or to make the equivalence judgments between the if-then sentence and five other expressions.

The instructions for the paraphrase task were identical to those used with the paraphrase task in Fillenbaum (1975). Subjects were told to "rephrase each sentence as accurately as possible without changing the meaning of the sentence." The only restriction on participants' paraphrases was that they were not permitted to include the words if and then. Subjects were told to "work on the assumption that you are trying to communicate each sentence to someone else so that the other person might get the sense of it as fully and exactly as possible," and that their job was, "not to improve the sentences or make them more sensible, but to paraphrase them, rewording each in a way that captures its meaning as accurately as possible."

For the equivalence judgment task, subjects were told that each if-then sentence would be followed by five other statements about the antecedent and the consequent of the if-then sentence. Participants were asked to rate on a

scale of 1 to 10 how similar in meaning each of the five statements was to the if-then sentence. A rating of 1 meant that the statement did not mean the same thing at all as the if-then sentence. A rating of 10 meant that the statement was totally synonymous with the if-then sentence. The order in which the five statements followed an if-then sentence was random.

Each if-then sentence was presented on a separate page. The ordering of the pages was random. Subjects completed the tasks at their own pace and were not given any feedback on their performance until they had completed all three tasks. At this time, participants were thanked for their participation and the details of the study were explained to them.

### Results and Discussion

The results are presented in three sections, beginning with the correlations between performance on the logic problems and the ratings of the necessity of antecedents to consequents, then the results of the equivalence judgment task and finally the results of the paraphrase task.

#### Performance on the Logic Problems and Ratings of Necessity of Antecedents to Consequents

Subjects' responses to the logic problems were scored using the same criteria that were used in Experiments 1 and 2. All responses consistent with a conditional interpretation were scored 1, and all responses

inconsistent with it were scored 0. For each of the 20 if-then sentences, four averages were computed over subjects, one for each type of problem. In addition, the mean rating of necessity was computed for each sentence. For each of the sentences, the mean performance of participants on each type of problem and the mean rating of the necessity of the antecedent to the consequent are shown in Table 11. Pearsonian correlations were then computed between these averages. Alpha was set at .01 per correlation.

The results of these analyses were as predicted. The ratings of the necessity of antecedents to consequents were substantially correlated with performance on conversion and inversion problems,  $r=.88$ ,  $p<.001$ , and  $r=.89$ ,  $p<.001$ , respectively. Ratings of the necessity of antecedents to consequents were not correlated with performance on either modus ponens or modus tollens problems,  $r=.09$ ,  $p=.356$ , and  $r=-.12$ ,  $p=.312$ , respectively, even after a correction was done for the restricted range associated with performance on these two types of problems (see Lindeman, Merenda & Gold, 1980, p.59-60). The corrected correlation coefficients were .10 and -.24 for modus ponens and modus tollens, respectively.

These results confirmed the hypothesis that it was the perceived necessity, or lack of necessity, of a given antecedent to the consequent that was the crucial determinant of whether a biconditional or conditional

Table 11

Averages used in the correlations between problem type  
and the rating of the necessity of antecedents to consequents  
in Experiment 3

modus ponens	modus tollens	conversion	inversion	rating
.980	.592	.327	.347	4.00
1.000	.776	.388	.429	3.39
.980	.918	.122	.061	2.47
.980	.837	.408	.469	2.35
1.000	.939	.429	.327	2.84
1.000	.918	.061	.020	0.84
.980	.878	.367	.204	3.33
1.000	.918	.020	.061	1.06
.878	.531	.449	.429	3.32
.939	.898	.429	.286	2.53
.959	.918	.551	.551	3.17
1.000	.939	.531	.571	3.87
1.000	.796	.592	.571	4.21
.959	.837	.673	.653	5.02
1.000	.939	.653	.673	5.00
1.000	.816	.612	.735	5.16
1.000	.918	.694	.653	5.31
.959	.796	.714	.776	6.71
1.000	.939	.714	.714	5.69
1.000	.837	.735	.776	7.57

Table 11 (continued)

Note: The first ten rows correspond to the biconditionally biased items and the last ten rows correspond to the conditionally biased items.



interpretation was given on the logic problems.

### Equivalence Judgments

Subjects' equivalence judgments were used as the dependent measure in a 2x5 repeated measures analysis of variance. The two independent factors in this analysis were bias (conditional, biconditional) and type of statement compared to if p then q (biconditional-reverse, biconditional-or, conditional-or, conditional-and, biconditional-only-if). Table 12 shows the mean and standard deviation for each bias by type of statement combination. Both bias and type of statement were within-subjects factors, so all F-ratios were evaluated with a conservative Greenhouse-Geisser correction for their degrees of freedom.

The results of this analysis were as follows: There were significant main effects for both bias and type of statement,  $F(1,33)=48.39$ ,  $p<.001$ , and  $F(4,132)=36.42$ ,  $p<.001$ , respectively. As well, there was a significant bias x type of statement interaction,  $F(4,132)=34.68$ ,  $p<.001$ .

Effect of bias. Because of the significant bias x type of statement interaction, the effect of bias was determined separately for each type of statement and likewise, the effect of type of statement was determined separately for each level of bias. All comparisons were done using dependent t-tests and alpha was set at .01 per

Table 12  
Experiment 3 Equivalence Judgments

Type of Statement	Bias			
	Conditional		Biconditional	
	M	SD	M	SD
Biconditional-reverse	5.61	1.93	7.23	1.66
Biconditional-or	6.20	1.90	7.51	1.57
Conditional-or	4.08	2.77	3.94	2.81
Conditional-and	2.70	1.59	2.07	1.13
Biconditional-only-if	5.19	1.87	7.12	1.77

comparison. The tests of the effect of bias are described first, followed by a description of the tests of the effect of type of statement.

Subjects rated the three biconditional statements as more similar in meaning to the if-then sentences assumed to bias a biconditional interpretation than to the if-then sentences assumed to bias a conditional interpretation,  $t(33)=10.47$ ,  $p<.001$ ,  $t(33)=5.86$ ,  $p<.001$ , and  $t(33)=8.71$ ,  $p<.001$ , respectively. Participants rated the conditional-and statement as more similar in meaning to the conditionally biased if-then sentences than to the biconditionally biased if-then sentences,  $t(33)=3.13$ ,  $p<.005$ . For the conditional-or statement, there was no difference in ratings given to conditionally biased if-then sentences compared to those given to biconditionally biased if-then sentences,  $t(33)=0.66$ ,  $p=.511$ .

This pattern of bias effects is important for two reasons. First, it is consistent with the pattern of bias effects found on the logic problems. The equivalence judgments thus provide an alternative dependent measure for discriminating between if-then sentences assumed to bias a conditional interpretation and ones assumed to bias a biconditional interpretation. Second, the presence of bias effects on equivalence judgments emphasizes the point that the subjects did not interpret if-then sentences only on the basis of a uniform semantics for the the connective if-then. If they did, there would have been no bias

effects because all of the if-then sentences had the same connective. Instead, they apparently incorporated the content of the propositions linked by the connective if-then when interpreting the sentences and when rating the similarity of the meaning of one sentence to that of another statement.

Effect of type of statement. The effects of type of statement are described first for the conditionally biased items and then for the biconditionally biased items. For the set of conditionally biased items, the three biconditional statements were rated as more similar in meaning to the if-then sentences than was the conditional-and statement,  $t(33) > 6.00$ ,  $p < .001$ . As well, the biconditional-or statement was rated more similar in meaning to the if-then sentences than were either the biconditional-only-if statement or the conditional-or statement,  $t(33) > 2.96$ ,  $p < .01$ . No other comparisons were significant at or beyond the  $\alpha = .01$  level.

For the set of biconditionally biased items, the three biconditional statements were judged more similar in meaning to the if-then sentences than were either of the conditional statements. The smallest  $t$ -value for these comparisons was  $t(33) = 5.07$ ,  $p < .001$ . In addition, the conditional-or statement was judged more similar in meaning to the if-then sentences than was the conditional-and statement,  $t(33) = 3.52$ ,  $p = .001$ . No other comparisons were significant at or beyond the  $\alpha = .01$  level.

The most striking aspect of these results was that the statements rated most similar in meaning to the if-then sentences were the biconditional statements rather than the conditional statements. For example, with the set of conditionally biased items, the biconditional-reverse and the biconditional-or statements were judged most similar in meaning to the if-then sentences. With the set of biconditionally biased items, all three biconditional statements were judged most similar in meaning to the if-then sentences. Under no circumstances were conditional statements judged more similar in meaning to the if-then sentences than were the biconditional statements. This finding highlights the point that if-then as used in natural language is not necessarily equivalent in meaning to if-then used in standard propositional logic. In standard propositional logic, if-then is customarily interpreted as a conditional (i.e., as material implication) and is commonly represented by the expression not-p or q. These findings challenge the appropriateness of using standard propositional logic as the norm for human reasoning.

#### Paraphrase Task

The number of biconditional paraphrases was totalled separately for the set of biconditional if-then sentences (total=12) and for the set of conditional if-then sentences (total=4). Table 13 presents examples of biconditional paraphrases. The crucial feature of these paraphrases was

Table 13

## Responses Scored as Biconditional Paraphrases in Experiment 3

1) If the houseplant is cared for properly, then it grows.

Houseplants must be cared for properly to grow.

To be sure that your houseplant grows, you must care for it properly.

2) If you are my daughter, then I am your father.

I am your father so you are my daughter.

You are my daughter because I am your father.

3) If Tom is Mary's husband, then Mary is Tom's wife.

Tom is Mary's husband because Mary is Tom's wife.

4) If the animal is a bird, then it has feathers.

An animal with feathers is a bird.

5) If the tree is an oak, then it has acorns.

The tree has acorns, therefore it is an oak.

6) If the animal is a fish, then it has gills.

An animal with gills is a fish.

7) If the animal is a cat, then it has a tail.

An animal with a tail is a cat.

8) If the tree is a spruce, then it has needles.

A tree with needles is a spruce.

## Table 13 (continued)

9) If the building is a school, then it has windows.  
A building with windows is a school.

Note: The first six if-then sentences are from the set of biconditionally biased items. The last three are from the set of conditionally biased items.

that they explicitly stated that the consequent was only possible with the antecedent from the if-then sentence. There were significantly more biconditional paraphrases to the biconditional if-then sentences than to the conditional if-then sentences,  $\chi^2(1)=8.00, p<.01$ .

Although the paraphrase measure discriminated statistically between biconditional and conditional if-then sentences, a major shortcoming associated with this measure was that the frequency of biconditional paraphrases was very low relative to the total number of paraphrases. The total number of paraphrases for the biconditional if-then sentences was 150 (10 sentences x 15 subjects) and yet only 12 of these, or 8%, were biconditional paraphrases. Similarly, the percentage of biconditional paraphrases given to conditional if-then sentences was only 2.7%.

Over 90% of the paraphrases given to both biconditional and conditional if-then sentences were of one of three general types. The first type involved the substitution of when or because for if, as in the paraphrase, "When the basketball team gets a basket, they score two points," for the sentence, "If the basketball team gets a basket, then they score two points." The second type entailed the conjunction of the antecedent and the consequent, such as in the paraphrase, "Tom is Mary's husband and Mary is Tom's wife," for the sentence, "If Tom is Mary's husband, then Mary is Tom's wife." The third type consisted of condensing the antecedent and the



consequent into a single proposition, such as in the paraphrase, "Cats have tails," for the sentence, "If an animal is a cat, then it has a tail." All three types of paraphrases are irrelevant to the conditional/biconditional distinction. The major problem with the paraphrase method is that it does not place enough constraints on subjects' performance and they are most likely to give responses irrelevant to the conditional/biconditional distinction.

#### GENERAL DISCUSSION

The major finding from this investigation was that the interpretation of if-then sentences was based not only on logical form or on the semantics of if, but also on the content of the antecedent and consequent propositions as well as the context of the sentence. In Experiment 1, participants were from four grade levels, ranging from kindergarten to grade 12. All grade levels gave biconditional interpretations to sentences in which the antecedent was necessary to the consequent but gave conditional interpretations to sentences in which the antecedent was only sufficient to the consequent. In Experiment 2, subjects were from three grade levels, ranging from grade 4 to university. All grade levels interpreted the same abstract if-then sentences as conditionals in one context but as biconditionals in another context. However, university students gave more conditional interpretations than did the other participants when abstract sentences were presented without context. In

Experiment 3, university students' interpretations of if-then sentences were correlated with their ratings of the necessity of the antecedent to the consequent. In addition, conditional and biconditional if-then sentences led to different equivalence judgments and paraphrases. Considered together, the results of the three experiments provided consistent evidence that content and context are crucial factors in the interpretation of if-then sentences and that any theoretical account of reasoning with these sentences must consider both the content and the context of the sentences.

A major concern, then, is whether current theoretical accounts of if-then can accommodate these content and context effects. Several theories of if-then reasoning within both philosophy and psychology range from truth table accounts, to inference rule schemata, to possible worlds semantics. This discussion focuses on the psychological theories, but some of the philosophical positions will be mentioned briefly, just to point out their relevance to psychology. An example of each of the major types of psychological theory will be described and the theory's ability to handle content and context effects will be discussed.

#### Theories of If-Then Reasoning

Truth table accounts. According to most truth table accounts, if-then should be regarded as the truth-functional conditional of propositional logic. This

position arose from the classical view within philosophy and is best exemplified in psychology by Piaget's theory of logical reasoning (Inhelder & Piaget, 1958). Although there are subtle differences between Piaget's theory and its philosophical counterparts (cf. Ennis, 1978), all the theories maintain that reasoning is based solely on the truth values of the antecedent and consequent propositions. That is, an if-then sentence is judged to be true when its antecedent and consequent are both true or both false, or when its antecedent is false and its consequent is true, regardless of the content of the propositions or of the context of the sentence. According to Piaget's theory, sensitivity to content and context, as was found in the present experiments, is indicative of immature or pre-formal reasoning and is to be expected in children younger than 11 or 12 years, but not in older children or adults. The performance of the kindergarten, grade 4, and grade 7 students in the present experiments, then, is consistent with Piaget's theory, but that of the grade 8, grade 12 and university students is not. In fact, if we were to interpret the findings of the present studies in terms of Piaget's theory, we would have to conclude that participants from all age groups were incompetent or pre-formal reasoners.

Because of other empirical demonstrations of adults and adolescents performing below the level predicted by his theory (see Linn, 1983, for a review), Piaget (1972),

revised his theory somewhat, stating that competence at this level is only expressed on certain tasks, such as those that involve content from a person's occupational or professional domain, but not on other tasks. However, this reformulation of his theory still does not account for the present data.

A major issue is whether the use of the truth-functional conditional is an appropriate criterion of "correct" reasoning with if-then, given that adults often do not interpret if-then in this manner (cf. Scholnick & Wing, 1983; Taplin, 1971) and their alternate interpretations are usually more appropriate to the intent of the if-then sentence, given its context (cf. Geis & Zwicky, 1971). In light of these findings, several psychologists have replaced the conditional as the criterion of appropriate reasoning with other truth-functional interpretations, notably the biconditional interpretation (Wildman & Fletcher, 1979) and the defective truth table interpretation (Wason & Johnson-Laird, 1972). According to the defective truth table interpretation, an if-then sentence is true when its antecedent and consequent are both true and is false when its antecedent is true and its consequent is false. No judgment is made about the truth of the sentence when its antecedent is false.

Neither the biconditional nor the defective truth table interpretations is sufficient to account for the content and context effects on if-then reasoning because

these interpretations, like the conditional one, imply that if-then sentences have only one appropriate interpretation. The experiments reported in this dissertation have clearly shown that if-then sentences have at least two interpretations, and the choice of one or the other is dependent on the content of the propositions linked by if-then and on the context of the sentence. It is impossible for any truth-functional interpretation alone to account for the content and context effects.

Although a truth-functional account by itself cannot account for reasoning with if-then, one possibility is that it can provide an adequate account in combination with other factors (cf. Ennis, 1978). The other factors, then, would be responsible for explaining the content and context effects. The truth-functional meaning, however, would have to be derived by some method other than the consultation of truth tables. There is now empirical evidence that truth tables are not used in natural reasoning, regardless of the presence or absence of other factors (Johnson-Laird, 1983, chp. 3). Accordingly, most current students of reasoning have abandoned truth-functional approaches (cf. Braine, 1978; Cheng & Holyoak, 1985; Evans, 1982, 1983; Johnson-Laird, 1983; Rips, 1983).

Inference rule schemata accounts. Inference rule schemata approaches were motivated by a desire to set up formal systems of reasoning that more closely resembled actual or natural reasoning than did the traditional

truth-functional approaches. An inference rule schema is simply "a rule to the effect that a certain proposition can immediately be concluded when certain other propositions have been established." (Braine, 1978, p. 3). It is customary to use a notation for inference rules in which the proposition that is the conclusion is written below a horizontal line (called the inference line), and the premises on which the conclusion is based are written above the line. The following is an example of an inference rule schema for reasoning with if-then:

$$\frac{\text{if } p \text{ then } q; p}{q}$$

There have been several inference rule schemata approaches within psychology (cf. Braine, 1978; Braine, Reiser & Romain, 1984; Johnson-Laird, 1975; Osherson, 1975; Rips, 1983). For the purpose of this discussion, the focus will be on Braine's approach because it makes explicit claims about the meaning of if-then.

According to Braine (1978), if-then means the same as the inference line in an inference rule. For the sentence if p then q, the expression if-then is simply a way of stating that one is entitled to conclude q, given p (i.e., to make a modus ponens inference). This is all that if-then means. It provides no information about the justification for concluding q, given p (i.e., whether q is caused by p, entailed by p, or simply co-occurs with p), nor does it govern the generation of modus tollens.

inversion, or conversion inferences. These inferences are not inherent in the meaning of if-then and they are explained by another component in an overall model of reasoning.

Braine (1978) postulated that inference rule schemata are only a part of an overall model of reasoning. Other components include a comprehension mechanism, a mechanism for selecting the appropriate inference rule schema, heuristics for planning the sequence of reasoning, and some definition of working memory. He did not elaborate on these components because he assumed that they were not unique to reasoning, and that the inference rule schemata were what set deductive reasoning apart from other cognitive processes. However, he acknowledged that predictions from inference rule schemata to reasoning performance always necessitate a consideration of the other reasoning components.

The results of the present studies are consistent with Braine's theory. In all three experiments, subjects from all age groups consistently made more modus ponens inferences than other inference types. Braine's limitation of the meaning of if-then to modus ponens emphasizes an important, yet previously ignored possibility. Other theoretical accounts, particularly those with a truth-functional perspective, have tended to attribute all inferences generated from if-then sentences to the semantics of if-then. Braine's approach opens up the

possibility that only a subset of the inferences generated from an if-then sentence arise from if-then. The others are generated in response to the propositions linked by if-then and to the context of the sentence. One of the problems with the other approaches is that they attempted to ascribe too much meaning to if-then, at the expense of developing theories that either cannot capture the multitude of interpretations that if-then sentences are given in natural language contexts, or postulate that if-then has no constant meaning (cf. Wason & Johnson-Laird, 1972). According to Braine, if-then does have a constant meaning. It is just more restrictive than people had previously assumed. ✓

If one accepts Braine's position that if-then is simply justification to make modus ponens inferences, then one is left with the task of explaining how content and context account for the other inferences generated from if-then sentences. Braine et al. (1984) briefly mentioned that other inferences, such as inversion and conversion, may arise from "conversational implicatures" (Grice, 1975), but the details of how natural language situations influence the generation of inferences are left unspecified. What is needed, then, is a theory to explain exactly how content and context regulate inference generation. The present investigation provided a first step in this direction. In Experiment 3, the generation of inversion and conversion inferences was shown to be highly



correlated with participants' ratings of the necessity of antecedent to consequent propositions ( $r=.89$  and  $.88$ ). This is an example of a specific aspect of content that explains over 77% of the variance in performance on the two types of inference problems.

Possible worlds semantic approach. This approach was developed within philosophy as an attempt to provide a uniform semantics of if-then. The approach can be traced back to Ramsey's (1931) idea that the way to evaluate an if-then sentence is first to add its antecedent to your stock of beliefs, and then to determine whether or not its consequent is true. If you believe that the antecedent necessarily entails the consequent, then you will infer that the consequent is true, and hence that the if-then sentence as a whole is true. As well, if you already believe that the consequent is true, regardless of the antecedent, then you will consider the if-then sentence true. Ramsey's test, however, works only for sentences in which you have no prior opinion about the truth of the antecedent (according to his procedure you automatically add the antecedent to your stock of beliefs).

Stalnaker (1968) modified Ramsey's test so that it could accommodate prior opinions about the truth of the antecedent. According to Stalnaker's version, when you believe the antecedent to be true, your evaluation of the sentence is based on your belief about the consequent. On the other hand, when you believe the antecedent to be

false, its addition to your beliefs will require some of them to be changed in order to avoid inconsistencies, and which particular changes are made depend on pragmatics.

Stalnaker used the above test procedure to formulate a set of truth conditions for if-then, based on the notion of possible worlds. The sentence if p then q is true "if and only if q is true in that possible world in which p is true but which otherwise differs minimally from the world in question (usually, of course, the actual world)" (Johnson-Laird, 1983, p. 15). Possible worlds are philosophical inventions for hypothetical states of affairs and the set of possible worlds is infinite (Stalnaker, 1981). Thus it is impossible for ordinary reasoners to consider "all possible worlds". The concept of possible worlds is important to psychology nonetheless because it has led to the psychological concept of a mental model (see Johnson-Laird, 1983). A mental model is a hypothetical situation, but unlike possible worlds, the set of mental models is finite and thus it is conceivable that they are within the working memory limits of ordinary reasoners.

Johnson-Laird (1983) developed a theory of if-then reasoning, based on mental models. According to his theory, people construct a set of mental models for an if-then sentence from both the content of the sentence and inferences from general knowledge. A mental model is an internal representation that has "a similar relation-structure to the process it models."

(Johnson-Laird, 1983, p.4). but, "need neither be wholly accurate nor correspond completely with what it models." (Johnson-Laird, 1983, p.3). Models are constructed by the recursive use of procedural semantics that aim at representing both the antecedent and the consequent and the relation between them. Models are formulated first in response to the antecedent, and then these models are revised given the consequent and the relation of the antecedent to the consequent. After the models are completed, reasoners formulate informative conclusions that are consistent with all models corresponding to the if-then sentence.

The content and context effects reported in the present dissertation can be explained by the mental models approach in that conditional if-then sentences would be modelled differently than biconditional if-then sentences. In the conditional models, the relationship between the antecedent and the consequent would be one of sufficiency but not necessity because one could infer that the consequent is highly probable with other antecedents. On the other hand, in the biconditional models, the antecedent would be necessary to the consequent because one could infer that it is unlikely for the consequent to occur with other antecedents. When people are required to reason about an if-then sentence, they do so by referring to the set of mental models they have constructed for that particular sentence. Thus, they will reason differently

with conditional and biconditional if-then sentences because the mental models constructed for these sentences include different relations between the antecedents and consequents.

An important aspect of Johnson-Laird's approach is that it emphasizes that if-then can be interpreted in different ways, depending on the context in which it is used. As well, his approach stresses that the interpretation of many if-then sentences is grossly under-determined, even when context is taken into account. Thus, there is no guarantee of the validity of many everyday inferences because it is impossible to know for sure whether the reasoner considered models in which the inferences are invalid.

A novel aspect of Johnson-Laird's theory is that reasoning is conceptualized without recourse to certain internalized rules of logic, notably truth tables or inference rule schematas. All reasoning is seen as the product of consulting mental models for truth conditions and counterexamples to conclusions. There is no mention of "rules in the head" as is prevalent in the other theories.

One problem with Johnson-Laird's approach is that it is not explicit about whether particular inferences come from if-then or from the context in which it is used. Johnson-Laird (1983) stressed that the meaning of an if-then sentence is built up compositionally from the

meanings of its components but he did not elaborate on what the meanings of the components are. In particular, his theory is somewhat vague about the meaning of the if-then component. He claimed that, "if has a single unequivocal semantics that leaves a role to be played by the interpretation of the clauses that it connects." (Johnson-Laird, 1983, p. 62). If-then is a clue to the construction of mental models, but it is not clear whether the expression if-then by itself is responsible for any inferences, such as the modus ponens inference. Johnson-Laird (1983) admitted that his theory still needed some elaboration. The meaning of if-then is one area that needs more attention.

Pragmatic reasoning schemas. Cheng and Holyoak (1985) postulated the presence of pragmatic reasoning schemas to account for the context sensitivity of human reasoning. Pragmatic reasoning schemas are "abstract knowledge structures induced from ordinary life experiences, such as 'permissions', 'obligations', and 'causations'." (Cheng & Holyoak, 1985, p. 395). A pragmatic reasoning schema "consists of a set of generalized context-sensitive rules which, unlike purely syntactic rules, are defined in terms of classes of goals (such as taking desirable actions or making predictions about possible future events) and relationships to these goals (such as cause and effect or precondition and allowable action)." (Cheng & Holyoak, 1985, p. 395)

According to Cheng and Holyoak (1985), content and context effects on the interpretation of if-then sentences arise because different if-then sentences evoke different underlying schemas and the schemas differ in the inferences they entail. For example, Cheng and Holyoak (1985) found that sentences that evoked a permission schema were more likely to lead to a conditional interpretation than were sentences with an arbitrary relationship between the antecedent and consequent.

Pragmatic reasoning schemas have trouble accounting for the content and context effects found in the present dissertation. The main problem is that sentences from the same schema, for example causations, occurred in both the conditional and biconditional materials and led to different inferences depending on the set of materials the sentence was in. In Experiment 3, the conditional causal statement "If the car is out of gas, then it will stall," led to 61.2% conditional responses on conversion problems and 73.5% conditional responses on inversion problems, whereas the biconditional causal statement, "If the basketball team gets a basket, then they score two points," led to 36.7% conditional responses on conversion problems and 20.4% conditional responses on inversion problems. Thus, the schema distinction is not the primary one that causes an if-then sentence to be interpreted as a conditional rather than a biconditional. Although the concept of a pragmatic reasoning schema may be useful for

some types of sentences, such as inducements, it is not universally applicable. It may account for reasoning with inducements because most inducements share the property that the antecedent is necessary to the consequent. In other schemas, such as causations, some representative sentences suggest that the antecedent is necessary to the consequent, whereas others suggest that it is only sufficient. The results reported in the present dissertation suggest that the crucial variable affecting the interpretation of if-then sentences is not pragmatic schema type but the perceived necessity of the antecedent to the consequent.

Another problem with Cheng and Holyoak's theory is that it places considerable emphasis on the notion of a schema but it has not provided independent evidence for the existence of schemas. In fact, the utility of schematic representations in memory is currently somewhat questionable (cf. Alba & Hasher, 1983). At this stage of theory development, Cheng and Holyoak may be better off not relying so heavily on the concept of schemas but instead use a more theoretically neutral concept that does not entail as many representational commitments.

Conclusions about the theories. Of the four types of theories considered, Braine's inference rule schemata approach and Johnson-Laird's mental model approach were found to best account for the content and context effects of the present dissertation, however, neither provided a

complete account. On the one hand, Braine's theory was useful because it limited the meaning of if-then to modus ponens inferences but the theory did not provide enough detail about how context (conversational implicatures) accounted for the other inferences commonly generated from if-then sentences, such as those made on inversion and conversion problems. On the other hand, Johnson-Laird's theory of mental models can, in principle, account for any inferences generated from an if-then sentence but this theory is not specific about whether all inferences arise from the context in which if-then is embedded or whether some are central to the expression itself.

One possible conclusion may be that a synthesis of Braine's and Johnson-Laird's theories is what is required for a complete account of reasoning with if-then sentences. Braine's theory would provide the meaning of if-then, whereas Johnson-Laird's would explain the content and context effects commonly observed when people reason with if-then sentences.

Another possibility is to adopt an approach based on reasoning by analogy from exemplars, which entails no rules, no schemata, and no abstract entities. According to this perspective, people reason by thinking of specific examples that correspond to the if-then sentence. Exemplar views have been formulated to explain the representation of concepts (see Smith & Medin, 1981) and the learning of grammar (e.g. Vokey & Brooks, in press). Studies of



reasoning with Wason's selection task suggest that an exemplar-based approach may also be relevant to theories of if-then reasoning (see Griggs, 1983 for a review).

Wason's selection task consists of an if-then sentence such as, 'If there is an E on one side of the card, then there is a 4 on the other side,' followed by an array of four cards, one each with E, K, 4, and 7. Subjects are told that each card has a letter on one side and a number on the other side, and that their job is to select just those cards that are necessary to turn over in order to decide whether the if-then sentence is true or false.

The correct solution, based on a conditional interpretation, is to turn over E and 7. However, fewer than 10% of adults give this solution (Manktelow & Evans, 1979) and instead choose just E, or E and 4.

In light of the poor performance on Wason's task with arbitrary content (i.e., letters and numbers), several studies were conducted using thematic, realistic, and concrete materials to see whether performance could be improved (see Evans, 1982, for a review). The results of these studies were mixed. In some (e.g. Johnson-Laird, Legrenzi & Sonino Legrenzi, 1972), the use of thematic materials resulted in improved performance, but in other studies (e.g. Griggs & Cox, 1982), there was no improvement in performance. In Johnson-Laird et al. (1972), the if-then sentence was "If a letter is sealed,

then it has a 50 lire stamp on it." This rule was followed by 5 envelopes: one sealed, one not sealed, one with a 50 lire stamp, one with another type of stamp, and one blank. The study was conducted in England, and each subject was asked to "imagine that he was a post-office worker sorting letters" and to "select those envelopes that you definitely need to turn over to find out whether or not they violate the rule." (Johnson-Laird et al., 1972, p. 397). This version of the task led to 81% correct selections, which was an improvement of about 71% over the performance typically found with the arbitrary version! However, Griggs and Cox (1982) failed to replicate Johnson-Laird et al.'s findings using American students as subjects and American and Mexican stamps and addresses in their materials. In the Griggs and Cox study, the thematic materials did not improve performance over that typically found with the arbitrary version of the task.

One explanation for the lack of facilitation with the thematic materials in Griggs and Cox's study concerns the relevance of the materials to the subjects. Postal rules of this sort were common in England but not in the United States. Therefore, the rule would not have been as relevant to the experiences of Griggs and Cox's American subjects as it was to the British participants in the Johnson-Laird et al. study.

Golding (1981) re-ran an updated version of the Johnson-Laird et al. postal problem using two age groups

of British subjects, one group over 45 years and one group under 45 years. The rationale for comparing the age groups was that the older group would have had experience with postal rules similar to the one in the problem whereas it is unlikely that the younger group would have had this experience. Postal rules similar to the one in the problem had not been used in England for several years. Golding predicted that only the older group should show enhanced performance with the postal rule problem and this is exactly what she found. The older group was correct 59% of the time on the postal problem, whereas the younger group was correct only 9% of the time. These results strongly suggest that thematic materials enhance performance because they cue subjects to recall their experience with the specific situation or analogous situations in which the correct solution is salient. Subjects seemed to be reasoning by exemplars or analogies. Consistent with this interpretation is the finding that the facilitation observed with thematic materials does not transfer to the arbitrary version of the task (Golding, 1981). Reasoning seems to be dependent on the specific exemplars and analogies.

An exemplar-based, reasoning by analogy approach is also supported by two major findings of the present investigation. The most impressive support was the finding that, in Experiment 3, subjects' ratings of the likelihood that the consequent could occur with alternative

antecedents was highly correlated with performance on inversion and conversion problems. Subjects interpreted the problems as conditionals when they were able to generate several situations in which the consequent occurred with alternative antecedents. On the other hand, participants interpreted the problems as biconditionals when they were could think of only a few, if any, instances in which the consequent occurred with alternative antecedents. Additional support for an exemplar-based analogical reasoning approach was also provided by the observation in Experiment 2, that the same abstract if-then sentences were treated as conditionals or biconditionals, depending on the type of examples that followed the problems. The performance of both children and adults was highly influenced ( $p < .001$ ) by the types of examples following the if-then sentences.

One issue about exemplar-based approaches that still needs to be addressed is whether the use of examples explains all inferences generated from an if-then sentence. The results of the present investigation provided evidence for exemplar-based reasoning on inversion and conversion problems but the role of exemplar-based reasoning on modus ponens problems was not established. The frequency of modus ponens inferences did not vary much across sentences. They were generated in the present study anywhere from 75% of the time to 100% of the time, depending on the sentence. Because modus ponens is so common, it must arise either (a)

from characteristics of exemplars that are more or less constant across sentences (an example of 'p and q' for any sentence, 'if p then q'), or (b) from the semantics of if-then, independent of the use of exemplars. Experiment 1 provided indirect evidence that the former may be the more accurate interpretation. In this experiment, participants were asked why they responded to the problems the way they did. Of interest, are the explanations participants gave when they failed to make modus ponens inferences. Every time subjects failed to make a modus ponens inference, they justified their answer by citing examples of 'p and not-q', or by stating a general claim about the existence of 'p' without 'q'. Consider the problem, "If an animal is a dog, then it has a tail. The animal is a dog. Does it have a tail?" Sample subject explanations for not generating the modus ponens inference to this problem were "sheep dogs don't have tails," "dogs don't have to have a tail to be a dog," and "some dogs' tails get cut off." Thus, subjects' explanations indicated that they were rejecting modus ponens inferences because of information derived from specific exemplars.

Another issue concerning exemplar-based reasoning is its relationship to Johnson-Laird's (1983) mental models approach. A mental models approach is very similar to an exemplar-based approach because both entail the consideration of specific examples and inferences from general knowledge. This is all that is necessary for an

exemplar-based theory. Johnson-Laird's theory of mental models, however makes several additional claims such as the presence of propositional representations in addition to exemplars, as well as the requirement that use of exemplars be computable (Johnson-Laird, 1983). At present, it is unclear whether there is sufficient empirical justification for these additional claims.

In conclusion, an exemplar-based, reasoning by analogy approach seems to provide the best account of the content and context effects on if-then reasoning. This approach is more parsimonious than an inference rule schemata approach, assuming that the use of exemplars can account for the generation of all inference types. An inference rule schemata approach only accounts for modus ponens inferences and the other inferences must be explained by conversational implicatures. As well, although both an exemplar-based theory and a mental models approach can in principle account for all inferences generated from if-then sentences, it is not clear whether the additional assumptions entailed by a mental models approach are necessary.

#### Are There Age Differences in Reasoning?

The results of the present experiments indicated that there were minimal age differences in if-then reasoning, as long as the content or the context of the reasoning materials made it clear which interpretation of a sentence was relevant. However, there were substantial age

differences (in the order of  $p < .001$ ) in the neutral condition of Experiment 2 when the reasoning materials were of the form, "If there is an A, then there is a B," and no examples were included to suggest whether a conditional or a biconditional interpretation was appropriate. In this condition, university students gave more conditional interpretations than did either the grade 7 or the grade 4 participants.

One explanation for the age differences in reasoning with some materials but not with others is that a different type of reasoning is required in the two instances. Markman (1978) distinguished between empirical and logical reasoning. Empirical reasoning is content- and context-dependent and entails the consideration of general knowledge but not the use of abstract reasoning principles. For example, suppose a person were given the problem, "If an animal is a dog, then it has a tail. The animal is not a dog. Does it have a tail?" Based on empirical reasoning, the person would answer maybe, assuming he or she could think of non-dogs that have tails (e.g., cats) and do not have tails (e.g., snakes). However, it would not be assumed that the person had the abstract principle that the answer to any problem of the form, "If p then q. not-p. q?" is indeterminate. If the person did possess this abstract rule and used it to regulate his or her reasoning, then this would be an example of logical rather than empirical reasoning. Logical reasoning is based

totally on abstract principles and because of this, it is constant across content and contextual variations. In addition, based on logical reasoning, people will choose conclusions that are factually incorrect, as long as the conclusions are consistent with the logical principle.

The distinction between empirical and logical reasoning has been used to explain age differences in reasoning (e.g. Sophian & Huber, 1984), as has a similar distinction proposed by Bucci (1978) between pragmatic (i.e., empirical) and structural (i.e., logical) reasoning. Young children are assumed to be capable only of empirical reasoning, whereas adults are thought to rely on logical reasoning. The results of the present dissertation are consistent with the position that children's reasoning is based on empirical grounds because of the substantial content and context effects found in Experiments 1 and 2. However, the relevance of logical reasoning to the performance of the adults is questionable. First, there is unambiguous evidence that at least some adult reasoning is empirical, not logical, because of the content and context effects on their performance in the present studies and in other areas, such as hypothesis testing (cf. Tschirgi, 1980). In the present experiments, the only instance in which one could possibly conclude that adults were using logical reasoning was in the neutral condition of Experiment 2. However, we still need to explain why adults who are capable of logical reasoning apparently abandon



that strategy whenever the task permits empirical reasoning. What then is the utility of logical reasoning? Perhaps logical reasoning is specific to reasoning tasks that are totally de-contextualized and thus force people to reason in a context-free manner. Adults are able to meet this challenge but children are not. However, this may not be relevant to inferential thinking in everyday situations, whether it be in science, advertising, legal proceedings or other areas, since all of these domains require reasoning in context. The present investigation has shown that reasoning in context is influenced by the context and thus is more empirical than logical in nature. The concept of logical reasoning is not appropriate for understanding the everyday inferences generated by people of any age because logical reasoning cannot, in principle, account for the content and context effects.

The notion of logical reasoning may have arisen because the primary concern of researchers in the area of human reasoning was to find evidence that "people possess rational rule systems to allow them to solve problems in a context-free manner." (Evans, 1983, p.8). There was an overwhelming concern with proving human reasoning to be rational or irrational (cf. Cohen, 1981) but now there is some question as to whether this is a useful enterprise. For example, Wason (1983) contended that the issue of rationality is, in principle, unresolvable because the notion of rationality is unfalsifiable. Whenever

performance does not conform to the logical standards, one can postulate the presence of other factors that led performance to go astray, even though competence was intact. Thus, how can one tell when reasoning is irrational (Cohen, 1981)?

In conclusion, it may be that adults are capable of context-free (i.e., logical) reasoning and that children are incapable of it but this does not appear to be relevant to understanding everyday reasoning. Whenever reasoning occurs in a context, both adults' and children's reasoning will be strongly influenced by the specific context. In effect, there will be few, if any, age differences in reasoning, provided that all age groups are familiar with the content of the reasoning material.

#### Implications for Instruction

The present results are relevant to the design of instructional programs aimed at teaching children to reason with if-then sentences. Most instructional programs for logical reasoning (e.g. Lee, 1985; Lipman, 1985), are based on the assumption that if-then is equivalent to the truth-functional conditional. The findings reported in the current dissertation strongly suggest that the truth-functional conditional is an inappropriate criterion of "correct" reasoning unless the content and context effects on its interpretation are acknowledged. Children need to be alerted to the fact that sometimes if-then is used in the context of a conditional, and other times in

the context of a biconditional. Children could be explicitly taught to ask themselves about the likelihood of the consequent of an if-then sentence occurring with alternative antecedents. When alternative antecedents are likely, then a conditional interpretation should be adopted. However, when they are unlikely, then a biconditional interpretation is appropriate.

Emphasizing the importance of both conditional and biconditional interpretations of if-then is crucial if children's inferences are to be adaptive to natural language contexts. The traditional downplaying of biconditional interpretations does an injustice to children's reasoning because biconditional interpretations are common in natural language (e.g., Geis & Zwicky, 1971).  
Generality of Content and Context Effects

The strong effect of content and context on reasoning with if-then is consistent with similar findings in other domains. For example, content and/or context have recently been shown to influence performance on (a.) Piaget's conservation task (Donaldson, 1978; Moore, Nelson-Piercy, Abel & Frye, 1984), (b.) responses to the Wason Selection Task and THOC problem (Griggs, 1983; Griggs & Cox, 1982; Newstead, Griggs & Warner, 1982; Wason, 1983), (c.) statistical inference (Kahneman & Tversky, 1982), (d.) construction of mental maps (Siegel, 1985), (e.) causal attributions (Dix & Herzberger, 1983), (f.) social inferences (White, 1984), (g.) interpretation of the

concepts before and after (Carni & French, 1984). (h.) determination of the temporal ordering of sentences with because and if (Emerson & Cerkoski, 1980). (i.) discrimination between inclusive and exclusive meanings of or (Braine & Romain, 1981; Newstead & Griggs, 1983; Newstead, Griggs & Chrostowski, 1984). and (j.) performance on class inclusion tasks (Miller & Barg, 1982; Smith, 1982).

These pervasive content and context effects suggest that developmental theories in general should focus on context-specific abilities (cf. Donaldson, 1978) rather than general abilities that are invariant across different contexts (e.g. Inhelder & Piaget, 1958). It is only by understanding the ways in which context influences cognition, that we will be able to understand everyday cognitive activities.

### References

- Alba, J. W. and Hasher, L. (1983). Is memory schematic? Psychological Bulletin, 93, 203-231.
- Begg, I and Harris, G. (1982). On the interpretation of syllogisms. Journal of Verbal Learning and Verbal Behavior, 21, 595-620.
- Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. Quarterly Journal of Experimental Psychology, 35, 39-49.
- Braine, M. D. S. (1978). On the relation between the natural logic of reasoning and standard logic. Psychological Review, 85, 1-21.
- Braine, M. D. S., Reiser, B. J. and Romain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G. H. Bower (Ed.). The psychology of learning and motivation (vol. 18). New York: Academic Press.
- Braine, M. D. S. and Romain, B. (1981). Development of comprehension of "or": Evidence for a sequence of competencies. Journal of Experimental Child Psychology, 31, 46-70.
- Bucci, W. (1978). The interpretation of universal affirmative propositions. Cognition, 6, 55-77.
- Carni, E. and French, L. A. (1984). The acquisition of "before" and "after" reconsidered: What develops? Journal of Experimental Child Psychology, 37, 394-403.

- Case, R., Kurland, D. M. and Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386-404.
- Cheng P. W. & Holyoak, K. J. (1985). Pragmatic reasoning schemas. Cognitive Psychology, 17, 391-416.
- Dix, T. and Herzberger, S. (1983). The role of logic and salience in the development of causal attribution. Child Development, 54, 960-967.
- Donaldson, M. (1978). Children's minds. Glasgow: Fontana.
- Emerson, H. F. and Gerkoski, W. L. (1980). Development of comprehension of sentences with "because" or "if". Journal of Experimental Child Psychology, 29, 202-224.
- Ennis, R. H. (1969). Ordinary logic. Englewood Cliffs, N.J.: Prentice-Hall.
- Ennis, R. H. (1978). Conceptualization of children's logical competence: Piaget's propositional logic and an alternative proposal. In L. S. Siegel and C. J. Brainerd (Eds.): Alternatives to Piaget: Critical essays on theory. New York: Academic Press, 201-261.
- Ennis, R., Finkelstein, M., Smith, E. and Wilson, N. (1969). Conditional logic and children. Ithaca, N.Y.: Cornell Critical Thinking Readiness Project (Eric Document Reproduction Service No. ED 040 437).
- Evans, J. S. B. T. (1982). The psychology of reasoning. London: Routledge and Kegan Paul.

- Evans, J. St. B. T. (1983). Linguistic determinants of bias in conditional reasoning. Quarterly Journal of Experimental Psychology, 35, 635-644.
- Falmagne, R. J. (1975). Reasoning: Representation and process in children and adults. Hillsdale, N.J.: Lawrence Erlbaum.
- Fillenbaum, S. (1975). 'If': Some uses. Psychological Research, 37, 245-260.
- Fillenbaum, S. (1976). Inducements: On the phrasing and logic of conditional promises, threats and warnings. Psychological Research, 38, 231-250.
- Fillenbaum, S. (1978). How to do some things with if. In J. W. Cotton and R. L. Klatzky (Eds.). Semantic factors in cognition. Hillsdale, N.J.: Lawrence Erlbaum, 169-214.
- Geis, M. L. and Zwicky, A. M. (1971). On invited inferences. Linguistic Inquiry, 2, 561-566.
- Golding, E. (1981). The effect of past experience on problem solving. Paper presented at the Annual Conference of the British Psychological Society, Surrey University, April. Described in Griggs (1983) on page 24.
- Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.). Syntax and semantics (vol. 3). New York: Academic Press.
- Griggs, R. A. (1983). The role of problem content in the selection task and in the THOG problem. In J. St. B. T. Evans (Ed.). Thinking and reasoning. London:

Routledge & Kegan Paul, 16-43.

Griggs, R. A. and Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 73, 407-420.

Guttmann, J., Levin, J. R. and Pressley, M. (1977). Pictures, partial pictures and young children's oral prose learning. Journal of Educational Psychology, 69, 473-480.

Hasher, L. and Zacks, R. T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 108, 356-388.

Henle, M. (1962). On the relation between logic and thinking. Psychological Review, 68, 366-378.

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

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

Johnson-Laird, P. N., Legrenzi, P. and Sonino Legrenzi, M. (1972). Reasoning and a sense of reality. British Journal of Psychology, 63, 395-400.

Kahneman, D. and Tversky, A. (1982). On the study of statistical intuitions. Cognition, 11, 123-142.

Kirk, R. E. (1968). Experimental design: Procedures for the behavioral sciences. Belmont: Brooks/Cole Publishing Co.



- Knifong, J.D. (1974). Logical abilities of young children - Two styles of approach. Child Development, 45, 78-83.
- Kodroff, J. K. and Roberge, J. J. (1975). Developmental analysis of the conditional reasoning abilities of primary-grade children. Developmental Psychology, 11, 21-28.
- Kuhn, D. (1977). Conditional reasoning in children. Developmental Psychology, 13, 342-353.
- Lane, D. S. (1983). Using observation and action instruction to facilitate conditional reasoning performance in early adolescents. Journal of Early Adolescence, 3, 335-347.
- Lee, S. (1985). Children's acquisition of conditional logic structure: Teachable? Contemporary Educational Psychology, 10, 14-27.
- Legrenzi, P. (1970). Relations between language and reasoning about deductive rules. In G. B. Flores D'Arcais and W. J. M. Levelt (Eds.). Advances in psycholinguistics. Amsterdam-London: North-Holland Publishing Co., 322-333.
- Levin, J. R. (1985). Methodological and statistical "bugs" in research. In C. Brainerd and M. Pressley (Eds.). Basic processes in memory and development: progress in cognitive development. New York: Springer-Verlag.

- Lindeman, R. H., Merenda, P. F. and Gold, R. Z. (1980). Introduction to bivariate and multivariate analysis. Glenview, Illinois: Scott, Foresman & Co.
- Linn, M. C. (1983). Content, context and process in reasoning during adolescence: Selecting a model. Journal of Early Adolescence, 3, 63-82.
- Lipman, M. (1985). Thinking skills fostered by philosophy for children. In J. W. Segal, S. F. Chipman and R. Glaser (Eds.). Thinking and learning skills (vol. 1): Relating instruction to research. Hillsdale, N. J.: Lawrence Erlbaum, 83-108.
- Manktelow, K. I. and Evans, J. St. B. T. (1979). Facilitation of reasoning by realism: Effect or non-effect? British Journal of Psychology, 70, 477-488.
- Manis, F. R., Keating, D. P. & Morrison, F. J. (1980). Developmental differences in the allocation of processing capacity. Journal of Experimental Child Psychology, 29, 156-169.
- Markovits, H. (1984). Awareness of the 'possible' as a mediator of formal thinking in conditional reasoning problems. British Journal of Psychology, 75, 367-376.
- Markovits, H. (1985). Incorrect conditional reasoning among adults: Competence or performance? British Journal of Psychology, 76, 241-247.
- McCawley, J. D. (1981). Everything that linguists have always wanted to know about logic. Chicago: University of Chicago Press.

- Meehan, A. M. (1984). A meta-analysis of sex differences in formal operational thought. Child Development, 55, 1110-1124.
- Miller, L. K. and Barg, M. D. (1982). Comparisons of exclusive versus inclusive classes by young children. Child Development, 53, 560-567.
- Moore, C., Nelson-Piercy, C., Abel, M. and Frye, D. (1984). Precocious conservation in context: The solution of quantity tasks by nonquantitative strategies. Journal of Experimental Child Psychology, 38, 1-6.
- Newstead, S. E. and Griggs, R. A. (1983). The language and thought of disjunction. In J. St. B. T. Evans (Ed.). Thinking and reasoning. London: Routledge & Kegan Paul, 76-106.
- Newstead, S. E., Griggs, R. A. and Chrostowski, J. J. (1984). Reasoning with realistic disjunctives. Quarterly Journal of Experimental Psychology, 36, 611-627.
- Newstead, S. E., Griggs, R. A. and Warner, S. A. (1982). The effects of realism on Wason's THOG problem. Psychological Research, 44, 85-96.
- O'Brien, D. P. & Overton, W. F. (1980). Conditional reasoning following contradictory evidence: A developmental analysis. Journal of Experimental Child Psychology, 30, 44-60.

- O'Brien, D. P. & Overton, W. F. (1982). Conditional reasoning and the competence-performance issue: A developmental analysis of a training task. Journal of Experimental Child Psychology, 34, 274-290.
- Osherson, D. (1975). Logic and models of logical thinking. In R. J. Falmagne. Reasoning: Representation and process in children and adults. Hillsdale, N.J.: Lawrence Erlbaum, 81-91
- Overton, W., Brynes, J. P. & O'Brien, D. P. (1985). Developmental and individual differences in conditional reasoning: The role of contradiction training and cognitive style. Developmental Psychology, 21, 692-701.
- Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. Journal of Experimental Child Psychology, 16, 278-291.
- Peel, E. A. (1967). A method for investigating children's understanding of certain logical connectives used in binary propositional thinking. British Journal of Mathematical and Statistical Psychology, 20, 81-92.
- Piaget, J. (1928/1959). Judgment and reasoning in the child. Patterson, N.J.: Vittlefield & Adams.
- Piaget, J. (1972). Intellectual development from adolescence to adulthood. Human Development, 15, 1-12.

- Ramsey, F. P. (1931). General propositions and causality. In F. P. Ramsey. The foundations of mathematics and other logical essays. London: Routledge and Kegan Paul.
- Quine, W. V. O. (1982). Methods of logic. Cambridge: Harvard University Press.
- Ray, J. L. & Findley, P. (1984). Open and restricted context sentences and reasoning with conditional propositions. Communication Monographs, 51, 243-252.
- Revlín, R. and Leirer, V. D. (1978). The effect of personal biases on syllogistic reasoning: Rational decisions from personalized representations. In R. Revlín and R. E. Mayer (Eds.). Human Reasoning. Washington, D. C.: V. H. Winston and Sons, 51-82.
- Revlis, R. (1975). Syllogistic Reasoning: Logical decisions from a complex data base. In R. Falamagne (Ed.). Reasoning: Representation and process in children and adults. Hillsdale, N.J.: Lawrence Erlbaum, 93-133.
- Rips, L. J. (1983). Cognitive processes in propositional reasoning. Psychological Review, 90, 38-71.
- Rips, L. J. and Marcus, S. L. (1977). Suppositions and the analysis of conditional sentences. In M. A. Just and P. A. Carpenter (Eds.). Cognitive processes in comprehension. Hillsdale, N.J.: Lawrence Erlbaum.

- Roberge, J. J. (1982). Linguistic factors in conditional reasoning. Quarterly Journal of Experimental Psychology, 34, 275-284
- Roberge, J. J. & Flexor, B. K. (1979). Propositional reasoning in adolescence. Journal of General Psychology, 100, 85-91.
- Roberge, J. J. & Mason, E. J. (1978). Sex differences in conditional reasoning. Journal of General Psychology, 99, 41-48.
- Roberge, J. J. and Paulus, D. H. (1971). Developmental patterns for children's class and conditional reasoning abilities. Developmental Psychology, 4, 191-200.
- Rumain, B., Connell, J. and Braine, M. D. S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. Developmental Psychology, 19, 471-481.
- Schmidt, C. R., Schmidt, S. R. & Tomalis, S. M. (1984). Children's constructive processing and monitoring of stories containing anomalous information. Child Development, 55, 2056-2071.
- Scholnick, E. K. and Wing, C. S. (1983). Evaluating presuppositions and propositions. Journal of Child Language, 10, 639-660.
- Seigel, A. W. (1985). Martha Muchow: Anticipations of current issues in developmental psychology. Human Development, 28, 217-224.

- Shapiro, B. J. and O'Brien, T. C. (1970). Logical thinking in children ages six through thirteen. Child Development, 41, 823-829.
- Smith, E. E. and Medin, D. L. (1981). Categories and concepts. Cambridge, Mass.: Harvard University Press.
- Smith, L. (1982). Class inclusion and conclusions about Piaget's theory. British Journal of Psychology, 73, 267-276.
- Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.). Studies in logical theory: Oxford: Blackwell.
- Stalnaker, R. C. (1981). Indicative conditionals. In W. L. Harper, R. Stalnaker and G. Pearce (Eds.). Ifs. Boston: D. Reidel, p. 193-209.
- Staudenmayer, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.). Reasoning: Representation and process in children and adults. Hillsdale, N.J.: Lawrence Erlbaum, 55-80.
- Staudenmayer, H. & Bourne, L. E. Jr. (1977). Learning to interpret conditional sentences: A developmental study. Developmental Psychology, 13, 616-623.
- Taplin, J. E. (1971). Reasoning with conditional sentences. Journal of Verbal Learning and Verbal Behavior, 10, 219-225.

- Taplin, J. E. and Staudenmayer, H. (1973). Interpretation of abstract conditional sentences in deductive reasoning. Journal of Verbal Learning and Verbal Behavior, 12, 530-542.
- Taplin, J. E., Staudenmayer, H. and Taddonio, J. L. Jr. (1974). Developmental changes in conditional reasoning: Linguistic or logical? Journal of Experimental Child Psychology, 17, 360-373.
- Vokey, J. R. and Brooks, L. R. (in press). Taming the clever unconscious: Analogic and abstraction strategies in artificial grammar learning. Cognition.
- Wason, P. C. (1983). Realism and rationality in the selection task. In J. St. B. T. Evans (Ed.). Thinking and reasoning. London: Routledge & Kegan Paul, 44-75.
- Wason, P. C. and Johnson-Laird, P. N. (1972). Psychology of reasoning: Structure and content. Cambridge, Mass.: Harvard University Press.
- White, P. (1984). A model of the layperson as pragmatist. Personality and Social Psychology Bulletin, 10, 333-348.
- Wildman, T. M. and Fletcher, H. J. (1977). Developmental increases and decreases in solutions of conditional syllogism problems. Developmental Psychology, 13, 630-636.
- Wildman, T. M. & Fletcher, H. J. (1979). Processing errors in conditional and biconditional reasoning. Contemporary Educational Psychology, 4, 366-380.



## Footnotes

1. The labels, conditional and biconditional, are being used only in the sense that they entail different inferences on inversion and conversion problems. No inferences about their truth-functional nature are intended. For instance, it is possible that in response to the biconditional if p then q sentences, participants are retrieving the corresponding if q then p sentences from their general knowledge. Having if q then p available enables subjects to give what appear to be biconditional interpretations to if p then q by simply generating modus ponens and modus tollens inferences from if q then p. Consider the conversion problem, "If an animal is a bird, then it has feathers. The animal has feathers. Is it a bird?" This problem can be translated into a modus ponens problem for the sentence, "If an animal has feathers, then it is a bird." The modus ponens problem would read, "If an animal has feathers, then it is a bird. The animal has feathers. Is it a bird?" Similarly, the inversion problem corresponding to, "If an animal is a bird, then it has feathers," can be rewritten as a modus tollens problem with "If an animal has feathers, then it is a bird," as the first premise. Thus, responses consistent with a biconditional interpretation may arise either from reasoning from if p then q or from reasoning from if q then p. The present data do not discriminate between these two possibilities. One way to tease this apart, would be to present hypothetical reasoning situations where you could

explicitly manipulate the availability of if q then p but this was beyond the scope of the present studies.

2. The explanations participants gave for their responses to the logic problems suggested that their answers were influenced by their familiarity with the sentence content rather than by general rules. This was true of the explanations given by participants in all of the age groups. Consider the problem, "If Tom delivers papers, then he earns money. Tom does not deliver papers. Does he earn money?" A grade 12 student who gave a conditional response to this problem answered "maybe because he could have another job," whereas one of his peers who gave a biconditional response answered "no, if he doesn't deliver something he doesn't get money for it." An example of a kindergarten student's conditional response to this item was "maybe, on allowance day," and one of her classmate's biconditional response was "no, he didn't do his job." None of the subjects justified their answers by stating a general rule except on modus ponens problems when participants from all age groups would often respond "yes, because it says so" or "it has to be because of the first sentence." However, participants who responded no or maybe to modus ponens problems, explained their choice of these responses by referring to specific examples of the antecedent being true and the consequent being false or by stating that it was possible for the antecedent to be true without the consequent being true.

3. Responses of no to conversion problems and of yes to

inversion problems were regarded as consistent with a conditional interpretation because the few times they occurred (less than 15% of the total responses), subjects often qualified their answer by saying, "no, it doesn't have to be true," or "yes, it can be true." In effect, the subjects' meant to say maybe. In any case, the analyses for Experiment 1 were repeated when these responses were not included as part of the conditional interpretation and the same pattern of effects was found.

4. The reader will notice that in Table 4, the standard deviations associated with problems in the biconditional materials are reduced relative to those in the conditional materials. This may have been caused by ceiling effects on modus ponens and modus tollens and by floor effects on conversion and inversion problems. If this were the case, then it would have led to an artifactually small mean square error term and thus a positive bias in the Scheffé post-hoc tests (Levin, 1985). In order to ensure that the significant effects of bias found using the Scheffé procedure were not simply due to a reduced error term, the comparisons were re-examined using the nonparametric Mann-Whitney U procedure (Kirk, 1968). The Mann-Whitney U procedure produced the same pattern of results as did the Scheffé procedure. That is, on inversion and conversion problems, there were significantly more conditional interpretations to the conditionally biased items than to the biconditionally biased items,  $Z=6.06$ ,  $p<.001$ , and  $Z=5.23$ ,  $p<.001$ , respectively. As well, on modus tollens,

participants scored higher with the biconditionally biased materials than with the conditionally biased materials,  $Z=5.31$ ,  $p<.001$ .

5. In order to ensure that the significant bias effects found using the Scheffé procedure were not simply due to reduced variability within some of the cells, the comparisons were re-examined using the nonparametric Mann-Whitney U procedure. Comparisons that were significant with the Scheffé procedure were also found to be significant ( $p<.01$ ) with the Mann-Whitney U procedure. The Z scores for these comparisons ranged from 2.66 to 4.76.

Appendix I

Instructions for Experiment 1

I'm going to give you some problems to think about. I will tell you two things, and then ask you a question. You will give me an answer to this question. You can answer Yes, No, Maybe or Don't Know. Pretend that I tell you that, "If this bicycle is red, then it is mine." I also tell you that, "The bicycle is red." Is the bicycle mine? If you think that the bicycle is mine, then say Yes. If you don't think that it is mine, then say No. If you don't think I have told you enough to decide whether the bicycle is mine or not, then say Maybe. That is, it might be the case that the bicycle is mine, or it might be the case that the bicycle is not mine. If you don't know the answer to a question, tell me that you Don't Know.

After you have answered each question, I want you to tell me why you gave the answer you did. I will also ask you how hard you found each question. Do you have any questions? Do you understand what you are supposed to do?

Appendix II

Samples of Materials used in Experiment 2

Neutral item

If there is an A, then there is a T.

There is a T.

Is there an A?

Biconditionally-biased item

If there is an A, then there is a T.

There is a T.

Is there an A?

AT QK FX JO BL

Conditionally-biased item

If there is an A, then there is a T.

There is a T.

Is there an A?

AT QT ET JO BL



### Appendix III

#### Instructions for Experiment 2

You will be given a number of cards one at a time. At the top of each card are two sentences followed by a question (and at the bottom of each card are 5 pairs of letters). [Show sample card]

I want you to pretend that the two sentences are true [point to sentences] Then I want you to answer the question. You can answer Yes, No, Maybe or Don't Know. (In order to help you answer the question, look at the examples of the pairs of letters that can occur together).

Let's try an example. If there is a P, then there is an X. There is a P. Is there an X? If you think there is, then answer Yes, if you think there isn't then answer No, if you think there may be an X, but there doesn't have to be one, then answer Maybe and if you don't know which answer to give, then answer Don't Know. (Remember to look at the pairs of letters at the bottom of the card, before you decide which answer to give.) Do you have any questions?

Let's begin. I will read each card while you look at it. (Again, don't forget to look at the pairs of letters at the bottom of each card).

**END**

**2 1 0 4 8 7**

**FIN**