INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI
CONDITIONAL INFERENCES DURING COMPREHENSION: 
IMPLICATIONS FOR THE NATURE AND DEVELOPMENT 
OF HUMAN CONDITIONAL PROCESSING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

The Degree Doctor of Philosophy in the

Graduate School of the Ohio State University

By

Aaron W. Rader, M.A.

*****

The Ohio State University
2001

Dissertation Committee:

Professor Gerald Winer, Adviser

Professor John Gibbs

Professor Vladimir Sloutsky

Professor Neal Johnson

Approved by

Adviser

Psychology Graduate Program
ABSTRACT

The processing of conditional \textit{(if-then)} statements in discourse has been a source of continual debate in philosophy, linguistics, and psychology. A central topic dividing theoretical camps is the nature of the representations that drive conditional inferences and the development of these representations. The reported experiments examine conditional inferences drawn during discourse comprehension to test several current theoretical viewpoints. In particular, theories differ as to whether invalid conditional inferences should occur during comprehension, and at what age range(s).

Three extant psychological theories were tested in four experiments. All experiments included adults, and two also included preadolescents to test developmental predictions. Syntactic theories posit that, during comprehension, people draw inferences that conform to the logically valid, procedural inference schema of modus ponens (MP). MP has the logical syntax of \textit{If }P \textit{ then }Q; P, which licenses the conclusion \textit{Q}. Given MP premises of that form, one should infer the conclusion automatically. The invalid inference of affirming the consequent (AC) should never occur during comprehension because procedural inference schemas for invalid inferences do not exist (an AC inference has premises of the form \textit{If }P \textit{ then }Q; Q, which suggest the conclusion \textit{P}). Mental models theories of conditional processing instead
predict that preadolescents and adults should draw MP and AC inferences whenever a conditional's antecedent and consequent are strongly associated in memory because participants' mental representations of premises will suggest both inferences. When antecedent and consequent are weakly associated, MP inferences should still occur but AC inferences should not. Finally, cognitive-developmental theory proposes that preadolescents below about age 14 tend to misrepresent conditionals as biconditionals (if and only if statements). During comprehension, preadolescents should draw MP and AC inferences because both are warranted under a biconditional interpretation. Older participants form veridical representations and should thus draw only MP inferences.

Four experiments examined these predictions. Experiments 1 and 3 contrasted mental models and syntactic theories by testing adults in an on-line priming task. Experiments 2 and 4 tested those theories and cognitive-developmental theory in a recognition task. In all four experiments, participants read short stories that conformed to either the MP or AC premise forms, but without conclusions. In Experiments 1 and 3, participants' response times to single-word probes from argument conclusions were significantly faster when stories fit either the MP or AC form, as compared to control stories. In Experiments 2 and 4, participants of both age groups reliably (and erroneously) judged that conclusions to MP and AC arguments had been presented, when in fact they had not. Results suggest that both preadolescents and adults draw MP and AC inferences during comprehension, a finding with strong implications for all three theories.
Dedicated to R.K.M. and to my parents
ACKNOWLEDGMENTS

I wish to thank Vladimir Sloutsky for many stimulating discussions, easy access to Upper Arlington schools, and continued support during the course of this project.

I would also like to thank the other members of my dissertation committee (Gerald Winer, John Gibbs, and Neal Johnson) for intellectual guidance throughout the course of this project.

I thank Rosalee Meyer for continued encouragement during the many trying phases of this arduous dissertation.

This research was supported by a grant from the James S. McDonnell Foundation to one of my committee members (V.S.), without which access to the computers and software needed to conduct these experiments would have been very difficult, if not impossible.
VITA

October, 5, 1972 ...................... Born — La Marque, Texas

1994 ............................. B.A., Psychology, The University of Texas at Austin

1996 ............................. M.A., Psychology, The Ohio State University

1995-1998 ........................ Graduate Teaching Associate, Department of Psychology, The Ohio State University

1997,1998 ........................ Graduate Teaching Associate, School of Teaching and Learning, The Ohio State University

1998-2001 ........................ Research Associate, College of Education, The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Psychology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xii</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Current theories of conditional inference</td>
<td>8</td>
</tr>
<tr>
<td>1.1.1 Syntactic (mental logic) theories</td>
<td>10</td>
</tr>
<tr>
<td>1.1.2 Semantic (mental models) theories</td>
<td>14</td>
</tr>
<tr>
<td>1.1.2.1 Standard mental models theory</td>
<td>14</td>
</tr>
<tr>
<td>1.1.2.2 Modified mental models theory</td>
<td>22</td>
</tr>
<tr>
<td>1.1.2.3 The issue of search for counterexamples</td>
<td>26</td>
</tr>
<tr>
<td>1.1.3 Cognitive developmental theory</td>
<td>29</td>
</tr>
<tr>
<td>1.1.4 Summary of theoretical predictions</td>
<td>33</td>
</tr>
<tr>
<td>1.2 Overview of experiments</td>
<td>35</td>
</tr>
<tr>
<td>2. Preliminary experiment: identifying strong and weak associates</td>
<td>44</td>
</tr>
<tr>
<td>2.1 Method</td>
<td>45</td>
</tr>
<tr>
<td>2.1.1 Participants</td>
<td>45</td>
</tr>
<tr>
<td>2.1.2 Materials</td>
<td>45</td>
</tr>
<tr>
<td>2.1.3 Procedure</td>
<td>45</td>
</tr>
<tr>
<td>2.2 Results</td>
<td>46</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>47</td>
</tr>
<tr>
<td>3.1</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>71</td>
</tr>
<tr>
<td>6.1</td>
<td>76</td>
</tr>
<tr>
<td>6.2</td>
<td>77</td>
</tr>
<tr>
<td>A.1</td>
<td>106</td>
</tr>
<tr>
<td>A.2</td>
<td>Stimulus story for the noun <em>lamp</em>. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics .....................................................................................................107</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>A.3</td>
<td>Stimulus story for the noun <em>snow</em>. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics .....................................................................................................108</td>
</tr>
<tr>
<td>A.4</td>
<td>Stimulus story for the noun <em>candy</em>. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics .....................................................................................................109</td>
</tr>
<tr>
<td>A.5</td>
<td>Stimulus story for the noun <em>book</em>. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics .....................................................................................................110</td>
</tr>
<tr>
<td>A.6</td>
<td>Stimulus story for the noun <em>thief</em>. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics .....................................................................................................111</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Mean acceptance rates (+SEs) for Inference test sentences by story version (collapsed over argument form) and age group, Experiment 2</td>
<td>65</td>
</tr>
<tr>
<td>6.1</td>
<td>Mean acceptance rates (+SEs) for Inference test sentences by story version (collapsed over argument form) and age group, Experiment 4</td>
<td>79</td>
</tr>
</tbody>
</table>
The logical connective if...then, and statements based on this connective, are omnipresent in human thinking, reasoning, and discourse. These statements, known as conditionals, are also important from the normative perspectives of logic and philosophy. As a terminological aid, it is noted that an if-then conditional consists of an antecedent (the proposition following if) and a consequent (the proposition following then). I will discuss the importance of conditionals to philosophy, logic, comprehension, and reasoning because these perspectives have helped frame many psychological theories of conditional processing.

Philosophical and linguistic analyses of conditionals abound that seek to classify and describe appropriate inferences based on them (e.g., McCawley, 1981; Sanford, 1989; Scholnick & Wing, 1991; Traugott, ter Meulen, Reilly, & Ferguson, 1986). For example, in some conditionals, the antecedent event may produce the consequent event, and these are considered causal conditionals (e.g., If I depress the brake, then the car slows down). Other conditionals are definitional, in that a property of the item described in the antecedent places it, analytically, in the class of things mentioned in the consequent (e.g., If he is an unmarried adult, then he is a bachelor). Other conditionals are
It is circumstantial, in that the antecedent event can allow the consequent event to occur without technically causing it (e.g., *If it snows today, I’ll read a book*). As another example, conditionals may be classified according to the factual status of the antecedent event. To illustrate, suppose that Ohio State occasionally loses football games, which brings about complaints from disgruntled fans. We may produce conditionals with antecedents that are factual (*If Ohio State loses, many fans complain*). Furthermore, on the day of a specific game, one may utter a conditional with an uncertain but possible antecedent (*If Ohio State loses today, the coach’s job is in jeopardy*). Assuming that Ohio State does lose and the coach is unceremoniously fired, one may utter a conditional whose antecedent is clearly counterfactual given the actual state of affairs (*If Ohio State had won, the coach would still be employed*). All of these distinctions are not always hard and fast, and other distinctions may be proposed. However, they illustrate that conditionals are often regarded as serving many different functions.

In terms of language, the connective *if-then* plays an important role in discourse comprehension and its development. *If-then* may enhance the integration of separate propositions into a coherent mental representation (Franks, 1996, 1997; Lea, 1995; Lea, O’Brien, Fisch, Noveck, & Braine, 1990). It may be used to express a variety of relationships between propositions, such as causal, hypothetical, and pragmatic relationships (e.g., Fillenbaum, 1986). Conditional statements and inferences based upon them appear in the discourse of even young children (Braine & Rumain, 1983; Reilly, 1986). Analyses of linguistic corpora produced in everyday interactions reveal that young children and adults produce causal, definitional, and circumstantial conditionals, typically of the uncertain-but-possible variety, with circumstantial uses appearing to be
most common. Additionally, both logically valid and logically invalid inferences occasionally appear in children’s and adults’ discourse, again illustrating the importance of the conditional to discourse processing (Scholnick & Wing, 1991, 1995). Finally, discourse has been argued to be the vehicle by which children acquire logical understanding of the connective if-then (Falmagne, 1990), suggesting a link between discourse comprehension and reasoning (a topic of great concern in the present work). Understanding how people process conditionals is thus a matter of fundamental importance for theories of discourse comprehension and its development.

Finally, the connective if...then also plays a critical role in the process of reasoning; it thus is fundamental in both formal logic and philosophy of science. In standard propositional logic, a conditional statement is identified with the material conditional. A material conditional’s truth conditions are such that it is true whenever the consequent is true or the antecedent is false. The conditional is thus false only when its antecedent is true and its consequent is false. To make these truth conditions clearer, the truth table for the material conditional is given in Table 1.1. As the table displays, even conditionals with false antecedents are technically true; however, it is important to note that people often consider such possibilities to be irrelevant to the truth of the conditional (Evans, Newstead, & Byrne, 1993).

A conditional may be combined with a categorical premise (sometimes called a minor premise) to yield one of four argument forms. Examples of these forms are displayed in Table 1.2. As the table illustrates, the argument forms of modus ponens (MP) and modus tollens (MT) are logically valid. That is, whenever the premises for an
Table 1.1: Truth table for the material conditional.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$Q$</th>
<th>If $P$ then $Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

MP inference are true, the conclusion's truth is guaranteed—it is not possible for both premises to be true and the conclusion false. On the other hand, the argument forms of affirming the consequent (AC) and denying the antecedent are invalid; it IS possible for the premises to be true but the conclusion false. Thus, if a person draws inferences according to standard propositional logic, she should draw the relevant conclusions to MP and MT arguments, but she should refrain from drawing the conclusion to any AC or DA argument.

In philosophy of science, the *if-then* conditional construction has been proposed as the manner in which to express the relationship between a hypothesis and evidence pertaining to the hypothesis. That is, *if* a given hypothesis or theory is correct, *then* certain observable events or conditions should occur. The hypothesis will be falsified if
<table>
<thead>
<tr>
<th>Validity</th>
<th>Argument Type</th>
<th>Modus Ponens</th>
<th>Modus Tollens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>Modus Ponens</td>
<td>IF ( p ) THEN ( q ): If the weather is nice, Ed takes a walk.</td>
<td>IF ( p ) THEN ( q ): If the weather is nice, Ed takes a walk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( p ): The weather is nice.</td>
<td>NOT-( q ): Ed does not take a walk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( q ): Ed takes a walk.</td>
<td>NOT-( p ): The weather is not nice.</td>
</tr>
<tr>
<td></td>
<td>Affirming the Consequent</td>
<td>IF ( p ) THEN ( q ): If the weather is nice, Ed takes a walk.</td>
<td>IF ( p ) THEN ( q ): If the weather is nice, Ed takes a walk.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( q ): Ed takes a walk.</td>
<td>NOT-( p ): The weather is not nice.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( p ): The weather is nice.</td>
<td>NOT-( q ): Ed takes a walk.</td>
</tr>
</tbody>
</table>

Table 1.2. Examples of the four conditional inferences.
the observables specified in the consequent do not come about (note that this state of affairs corresponds to a modus tollens argument: \textit{IF hypothesis THEN observables; NOT observables; therefore, NOT hypothesis}). This view is of course problematic and has been criticized on many grounds (Howson & Urbach, 1993; Putnam, 1991). However, it still seems to express an intuitively correct way in which hypotheses should be evaluated: make predictions of certain observable phenomena assuming the hypothesis to be true, then look for the occurrence of those phenomena.

A significant step in understanding how people process conditionals, a step that blended logic, philosophy, and development, was taken by the developmentalist Piaget. Piaget proposed that, during adolescence, people develop the competence needed to evaluate conditional hypotheses in accordance with formal logic. Specifically, adolescents and adults are able to conduct hypothesis tests that are informative, in that all irrelevant variables are properly controlled (Inhelder & Piaget, 1958). Importantly, the relationships specified by such hypotheses should be represented in a format structurally analogous to the logical conditional, and the adolescent should also understand all the possibilities compatible with a conditional (i.e., all the “true” entries in the truth table for the conditional). These competencies appear as adolescents develop the ability to represent possibilities compatible with a given state of affairs in a propositional logical format that involves conditionals as well as other propositional logical connectives (Inhelder & Piaget, 1958).

Two points are critical about Piaget’s proposal. First, his argument proposes that normative analyses from logic and philosophy could be relevant to the study of human thinking and its development. The second point is his proposal that adolescents (by age
14-15 years) and adults should possess the capacity mentally to represent conditional relationships in a format corresponding to the conditional of propositional logic. This proposal marks Piaget as perhaps the first “mental logician;” i.e., the first psychologist to argue that people possess an internal mental logic, structurally analogous to formal logic, that is applied to various kinds of reasoning problems. Piaget's proposal has inspired much of the research on human reasoning, including conditional reasoning. Adherents of this tradition typically argue that some cognitive implementation of classical logic, perhaps augmented by a few additional principles, drives reasoning (Braine, 1978; Macnamara, 1986; Overton, 1990; Rips, 1994).

An alternative theoretical tradition argues that people's rationality is bounded by their limited processing capacity (Simon, 1956). As the cognitive complexity of the task increases, it becomes increasingly difficult to pursue normatively warranted solutions. As a result, people use a variety of heuristics and economical cognitive representations to bypass the complexity of the task and reach non-optimal, but generally effective, solutions (e.g., Tversky & Kahneman, 1974; Gigerenzer & Goldstein, 1996; Johnson-Laird, 1983). The bounded rationality approach has also inspired much research on reasoning, including conditional reasoning. Of course, theoretical explanations and predictions generated within this approach differ dramatically from those generated by Piaget.

One area where the two theoretical traditions diverge is in comprehension (in which people have to process large amounts of information quickly) and the formation of a problem representation. The processing of conditionals is one domain for which the two traditions provide quite different explanations. This work is concerned with how
people process conditional statements in discourse, and how several theories of conditional processing and reasoning address this issue. This issue is important because, to reason with conditional premises, one must first form a representation of them. Alternative theories of conditional processing make strikingly different assumptions about the representation of conditionals. These initial premise representations may in turn determine the conditional inferences that are drawn during comprehension. As shall be described, some theories suggest that only valid inferences, in particular MP, should occur during comprehension; whereas other theories indicate that MP and invalid inferences such as AC may be drawn during comprehension. Finally, examination of the conditional inferences drawn during comprehension helps to determine whether deliberate reasoning essentially parallels comprehension, as some theories suggest. A main theme developed in this work will be that conditional inferences in comprehension represent only an initial, “quick and dirty” type of inference, and that deliberate reasoning involves an additional stage or stages beyond comprehension. Major theories of conditional processing will now be reviewed. The focus will be on the theories’ applications to comprehension and deliberate reasoning tasks, but I discuss the latter only when appropriate.

Current Theories of Conditional Inference

Conditionals have received extensive attention in psychology, and a variety of theories have been proposed to explain how people represent and draw inferences from them. However, a number of these theories only seem to pertain to conditionals with specific contents (e.g., Cheng & Holyoak, 1985; Cosmides, 1989; Light, Blaye, Gilly, & Girotto, 1989; Manktelow & Over, 1991; see also review by Evans, Newstead, & Byrne,
1993). Critically, many of these theories are proposed and tested within the context of the Wason card selection task. The task has received enough attention that it should be described in some detail.

In this task, participants are given a conditional rule along with four cards. The rule describes what must be present on each side of each card if the rule holds (e.g., *If a vowel is on one side then an even number is on the other side.*). One side of each card presents the antecedent (either affirmed or negated) and the other side presents the consequent (affirmed or negated). For a conditional *If p then q*, the participant sees cards corresponding to *p, q, not-p* and *not-q*. Typically, the participant is asked to select cards that could potentially falsify the rule. If the rule is treated deductively, one should select the cards corresponding to the antecedent (*p*) and the negated consequent (*not-q*).

This task is very deliberate and complex (as opposed to the automatic, rapid process of comprehension), and it is not well understood despite considerable study. The wide variety of theories built around this task, including a Bayesian hypothesis-testing account, (Oaksford & Chater, 1996), an evolutionary cheater-detection account (Cosmides, 1989), and a non-logical linguistic heuristics account (Evans, 1984, Evans & Over, 1996), attest to this fact. Also, task content can lead to responses quite at variance with logical standards (Evans & Over, 1996). These considerations suggest that content-specific theories based on the selection task do not provide good explanations for the online processing of conditionals in discourse. Three candidates with more promise are the syntactic and semantic theories of reasoning, as well as a cognitive-developmental theory of conditional reasoning. These theories will now be examined, along with how they seem to relate to human comprehension.
Syntactic (mental logic) theories

Syntactic theories are part of a tradition called mental logic, which holds that there is a largely innate logical competence driving the average person’s ability to reason logically and to have intuitions of logical necessity (e.g., Braine, 1978, 1990; Fodor, 1975; Macnamara, 1986; Rips, 1994). Several syntactic theories propose that logically untrained individuals possess procedural inference schemata that apply to the propositional logical form (as defined by propositional logical connectives) of incoming discourse (Braine & O’Brien, 1991; Lea, 1995; Lea, et al., 1990; Sperber & Wilson, 1995; see also Rips, 1995). These schemata are stored in the lexical entries for logical connectives such as if, and, and or. These schemata are procedural because they should apply “spontaneously, without goals, goal setting, or reasoning strategies” (Lea, et al., 1990, p. 363). The schemata thus play an important role in normal language comprehension. Not all schemata proposed by syntactic theorists are argued to apply automatically to discourse, so the following discussion pertains to the subset of schemas that are.

The basic inference process is as follows. A person first abstracts the propositional logical form of the discourse; this abstraction is driven by the activation of the inference schemas stored in the lexical entries for logical connectives. Secondly, whenever the logical form of the input matches one or more of the automatic syntactic inference schemas, the relevant conclusion licensed by the schema(s) is drawn automatically and without conscious deliberation. Note that this account ties reasoning directly to the process of language comprehension; conscious considerations of validity, necessity, and the like are not utilized in this process.
One of these inference schemas is that for modus ponens (MP). Whenever the input matches the propositional form of this schema \( (\text{If } P \text{ then } Q; \ P) \) the conclusion \( (Q) \) should be drawn and enter working memory, even if the person is not trying to reason and has not been instructed to do so. As an example, consider this sample story:

(1) Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was.
(2) He thought that \textit{if it was dark outside, then it was night.}
(3) Still feeling sleepy, Frank arose to look out a window.
(4) He saw that \textit{it was dark outside.}

The italicized statements in sentences (2) and (4) conform to the MP inference schema of several syntactic theories (Braine & O’Brien, 1991; Lea, 1995; Rips, 1994). Thus, someone presented with this story for comprehension should infer that \textit{it was night} without instruction to draw inferences. Indeed, participants given stories like the one above will then erroneously but reliably recognize such inferences as actually having been presented in the stories (Franks, 1996, 1997; Lea, et al., 1990). Additionally, results of priming experiments using a lexical decision task indicate that MP inferences occur online (Lea, 1995, Experiments 3 and 4). The reality of MP inferences in text comprehension thus seems established, though whether these inferences reflect an underlying mental logic is debatable.

However, in addition to valid MP inferences, people often make invalid conditional inferences such as affirming the consequent (AC) (Evans, 1993; Evans, et al., 1993, Johnson-Laird, 1999). Syntactic theories, therefore, make additional assumptions to handle the occurrence of invalid conditional inferences, because these theories do not include logically invalid inference schemas. For example, one suggestion is that failures of memory or attention may cause one to misrepresent the logical form, leading to errors
in that inferences are based on these faulty representations (Cohen, 1981; Rips, 1994, 1995). However, it is hard to see how random errors could lead to systematic response tendencies.

Another proposal, derived from linguistics (Geis & Zwicky, 1971; Grice, 1975), is that invalid inferences arise through other, pragmatic "invited inferences" that a person draws (Braine & O’Brien, 1991; and Rumain, Connell, & Braine, 1983, invoke invited inferences). Specifically, a conditional If $P$ then $Q$ may invite one to infer that If not-$P$ then not-$Q$. Thus, hearing If you mow the lawn, I’ll give you five dollars leads one to believe that If you don’t mow the lawn, I won’t give you five dollars is also true (Geis & Zwicky, 1971). Under such an interpretation, both valid and invalid conditional inferences seem legitimate. However, syntactic theorists have not hypothesized that invited inferences should occur during comprehension, and no direct evidence exists for the occurrence of such. Braine (1978, p. 9) mentions briefly a study claiming to find evidence that invited inferences are computed during deliberate reasoning, but he presents no data or details of the experiment, so his claim is impossible to evaluate.

A final proposal is that a conditional may be misrepresented as a biconditional (If and only if $P$ then $Q$), in which case all conditional inferences are valid according to standard propositional logic. However, this proposal on the face cannot explain AC inferences during comprehension, if in fact they occur. The reason is that no current syntactic theory includes a procedural inference schema for the biconditional! The biconditional idea thus seems useless at first glance. However, a more intriguing version of this proposal is that people may maintain two variants of the conditional premise: (1) If $P$ then $Q$ and (2) If $Q$ then $P$. In this case, either categorical premise (i.e., $P$ or $Q$) would
supply premises conforming to the MP form, thus licensing a valid inference. Thus, an
AC inference may instead be an MP inference based on representation (2) of the
conditional premise. This hypothesis will be addressed in the General Discussion.

Developmentally, mental logic theories make few predictions. Because they
explicitly tie logical competence to language acquisition (Braine, 1990; Braine &
O'Brien, 1991, Rumain, et al., 1983), they posit that procedural inference schemas are
largely acquired at an early age. Supporting this idea, many studies suggest that MP
inferences are routinely drawn by participants of all ages (reviewed in Braine & Rumain,
1983; Evans, et al., 1993). Young children 4.5 years and older seem to produce these
inferences in conversation at rates comparable to adults, although they and adults often
seem to give biconditional interpretations to conditionals in conversation (Reilly, 1986;
Scholnick & Wing, 1995). The one developmental prediction is that young children
below about age 10 should be more likely than older participants to draw invited
inferences, and hence more likely to commit invalid inferences than older participants.
This proposal does have some support (Braine & Rumain, 1983; Rumain, et al., 1983). However, these theories imply that preadolescents and adults, the subjects of this work,
should process conditionals in essentially the same fashion.

Critically, the online occurrence of any invalid inference, such as AC, during
comprehension would be difficult to capture within a syntactic approach. In turn, the
syntactic assumptions about the representation and processing of conditional arguments
during comprehension would be called into question. Syntactic theories thus predict MP
inferences during comprehension but apparently not AC inferences, for both
preadolescent and adult participants.
Semantic (mental models) theories

In this section two semantic theories will be reviewed: the standard theory of mental models and a modification of that theory focused on conditionals.

Standard mental models theory. The second approach to conditional inference is the semantic, which is best exemplified by the standard mental models theory of inference (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991; Johnson-Laird, Byrne, & Schaeken, 1992). According to this theory, people do not parse incoming discourse into propositional logical syntax; hence inferences drawn are not based on inference schemas that correspond to formal logic.

Instead, this theory proposes a three-step inference process:

1. First, a person forms a model of incoming discourse. Typically, she will represent only those possibilities compatible with the truth of the incoming discourse and only those that are explicitly mentioned. Working memory limitations and an implicit assumption that the only relevant possibilities are those that are stated (Evans & Over, 1996; Sperber & Wilson, 1995) may cause these limitations. Participants' representations may thus deviate from standard logical syntax (e.g., a conditional is true whenever its antecedent is false, but the utterance of a conditional does not express these possibilities explicitly and they will probably not be represented).

2. The person draws an informative (non-repetitive) inference from this representation. The inference is generated by combining representations of the individual premises, a process that should eliminate some possibilities and leave one or more possibilities as putative conclusions.
3. Finally, the person searches for any counterexample; that is, a model of the situation in which the discourse premises are true but the inference is false. If no such model is found, the inference is accepted. Thus, the third step predicts that people are sensitive to the logical validity of conclusions, though logically incomplete representations tend to prevent ideal performance. Also, the third step does not tie logical reasoning performance completely to comprehension processes, since sensitivity to validity should be guided by a deliberate process of search (Evans, 2000).

The stage of search for counterexamples seems to be a deliberate process that is not a natural part of comprehension (an idea developed below). The discussion of this theory’s application to comprehension will thus address only the first two stages.

Let us consider the sample story given earlier to examine the process of inference according to the mental models view:

(1) Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was.
(2) He thought that if it was dark outside, then it was night.
(3) Still feeling sleepy, Frank arose to look out a window.
(4) He saw that it was dark outside.

The model theory predicts that a conditional such as that in sentence (2) will be represented in memory with tokens that are analogue representations of the antecedent and consequent, respectively:

dark outside night

... 

These tokens are written on one line because they correspond to a single possibility, or model; i.e., a model in which both are true. The ellipsis corresponds to other “implicit”
possibilities compatible with the conditional’s truth (e.g., any representation of a false antecedent) but that are not made explicit. It is likely that implicit possibilities will not be represented because they do not seem relevant to the discourse and are not explicitly stated (Sperber & Wilson, 1995; see also Evans & Over, 1996; Sloutsky & Goldvarg, 1999). Indeed, a tenet of the Johnson-Laird (1999) theory is that people prefer to work with as few possibilities as possible, so a single possibility would be most preferred.

How are MP inferences accomplished? When the participant learns that it was dark outside (sentence 4), this model is added to her representation of the input:

dark outside

This model matches up with the situation described in the antecedent of the conditional premise. Now step 2 of the process, inference generation, occurs. The models for the conditional premise (sentence 2) and minor premise (sentence 4) are combined in the participant’s representation of the discourse. This combination cancels the implicit models and picks out the proposition that it was night as a conclusion. This entire process accomplishes an MP inference without recourse to logical inference schemas.

AC inferences may occur by a similar process. Suppose that the participant instead encounters the story presented below. Now, the italicized parts of sentences (2) and (4) correspond to the form of an AC inference:

(1) Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was.
(2) He thought that if it was night, then it was dark outside.
(3) Still feeling sleepy, Frank arose to look out a window.
(4) He saw that it was dark outside.
The conditional premise would be represented exactly as in the previous example. When
the participant processes the fourth sentence of this story, this model will be added to the
discourse representation:

night

This model cancels the implicit model and matches the model of the conditional premise,
picking out the proposition that *it was dark outside* (this corresponds to step 2, inference
generation, of the model theory). In this case, the entire process accomplishes an AC
inference. No additional assumptions about invited inferences need be invoked to explain
this process. It is predicted that these AC inferences should occur during comprehension,
a possibility that has not been examined by other researchers.

Importantly, research on discourse processing concurs with the prediction that MP
and AC inferences could occur during comprehension. From this perspective, MP
inferences are predictives inference about “what will happen next” in a text. Although
the issue is debated, many studies of discourse processing indicate that comprehenders
draw predictive inferences when the predicted event is highly constrained by the text,
when it is fairly predictable based on world knowledge, and when no other events are
easily predictable based on the text (Fincher-Kiefer, 1995, 1996; Keefe & McDaniel,
such as that based on the example story meet these criteria; the event in the consequent is
plausible in context and easily predictable, given that it is explicitly stated. Also, no
other outcome is as readily inferred from story events. Some non-logical sentential
connectives serve as signals to make inferences that integrate separate propositions into
coherent structures (Halliday & Hasan, 1976; Millis, Golding, & Barker, 1995; Millis & Just, 1994; Murray, 1997). The occurrence of if-then may be another such signal, and MP is one type of inference that could be drawn based on the conditional.

Other considerations based on studies of comprehension suggest that AC inferences during comprehension could occur as widely as MP inferences. AC inferences may be a type of bridging inference. Bridging inferences help establish local coherence among parts of a text—that is, they facilitate the integration of propositions from adjacent clauses or short sequences of clauses that are in working memory simultaneously (Haviland & Clark, 1974; Gerbsbacher, 1991; Singer, 1994). Bridging inferences are considered obligatory in comprehension, and substantial evidence exists that readers make inferences of this sort (e.g., Fincher-Kiefer, 1995; Magliano, Baggett, Johnson, & Graesser, 1993; McKoon & Ratcliff, 1992; Potts, Keenan, & Golding, 1988).

Maintaining local coherence is also considered a primary goal of comprehenders, one that is sought for automatically during reading (Gerbsbacher, 1991; Kintsch, 1998; McKoon & Ratcliff, 1992). Thus, suppose that one reads both If it was night, then it was dark outside and It was dark outside. Logically, one should refrain from making an AC inference (that it was night), but the logical response requires one to remain uncertain about the conclusion. However, if readers strive to maintain coherence, they should avoid this indeterminacy, and an easy way to do so is to draw an AC inference. As with MP, the connective if-then may be a signal to make such inferences when possible.

Note that this analysis is consistent with the “bounded rationality” approach: conditional inferences in comprehension of the MP and AC variety allow one to maintain an economical and integrated representation of discourse. Although AC conclusions are
logically inappropriate, it is nevertheless possible that some discourse situations, such as the examples presented earlier, actually encourage such inferences. The finding that AC inferences do occur in actual linguistic corpora lends credence to this proposal (Scholnick & Wing, 1991, 1995).

Aside from the discourse processing analysis, phenomena in several domains are consistent with another assumption of the Johnson-Laird approach and with the notion of bounded rationality: people prefer to work with a single possibility that is compatible with some state of affairs. In the case of conditional inference, people might represent only one possibility (a conjunction of the conditional’s antecedent and consequent) and draw inferences that allow the maintenance of a single-possibility representation. Both MP and AC inferences could be drawn during comprehension, if more advanced reasoning strategies are not brought into play. However, other phenomena from a variety of domains also seem consistent with the one-possibility hypothesis. Arguably, representing a single possibility is an overarching goal of many cognitive processes aside from reasoning and comprehension.

Studies of decision-making support the notion of a one-possibility preference. For example, people who had to decide whether to carry out an action such as going to a movie never sought out information about alternative actions. Participants in these studies only seemed to consider the possibility explicitly given to them, unless context suggested clear alternatives about which they could collect information (Legrenzi, Girotto, & Johnson-Laird, 1993). Furthermore, people are more willing to choose an option when there are no competing options available, even if the competing options are less attractive than the target one (Shafir, Simonson, & Tversky, 1993). Apparently, the
competing options make people more reluctant to make a decision. Additionally, when choosing whether to purchase an option, people seem to find a disjunction of possible reasons for the purchase less compelling than either reason presented alone, even when people would tend to make the purchase regardless of which reason actually turned out to be true (Shafir, Simonson, & Tversky, 1993). People prefer a single definite possibility to a disjunction of possibilities when making decisions; the disjunction call for costly computations to reach an optimal decision, whereas the single possibility allows one to reach a decision without as much cognitive effort.

In the domain of hypothesis testing, when participants must formulate a hypothesis to explain some data pattern, there is overwhelming evidence that people only consider one possibility - often one suggested by the task. For example, when given the number sequence 2-4-6 and asked to discover, by suggesting other sequences, the rule used to generate the sequence, people generally generate sequences conforming to the rule increasing even numbers (Wason, 1960). The experimenter’s rule was actually any increasing number sequence. Enough studies replicated this phenomenon, in a variety of domains, that it was given the name “confirmation bias” (e.g., Mynatt, Doherty, & Tweney, 1977). Similar findings exist in studies of children’s scientific reasoning. Children and some adolescents tend to anchor on one explanation of a set of observations, to the extent that they will often misinterpret the observations themselves, and they seek evidence that will support their ideas (Kuhn, Amsel, & O’Laughlin, 1988; Kuhn, Garcia-Mila, Zohar, & Andersen 1995). It is important to note that one rational
analysis demonstrates that, in terms of reducing indeterminacy, confirmation bias may be a more efficient way to discover rules and hypotheses (Klayman & Ha, 1987). The important point is that the phenomenon demonstrates a preference for one possibility.

A similar phenomenon exists in cases of children’s hypothetical and modal reasoning. Children, and even adolescents to an extent, fail to consider multiple possibilities compatible with a given set of data. In particular, when an observation is compatible with more than one possible explanation, children below about age 10 treat the observation as suggesting only one explanation (Byrnes & Overton, 1986; Fay & Klahr, 1996; Pieraut-Le Bonniec, 1980; Scholnick & Wing, 1988). The opposite type of error – treating a determinate data set, compatible with just one possibility, as being indeterminate – virtually never occurs. Instead, children seem to fixate on the first obvious possibility and discount any other alternative explanations. It should be noted that, although older participants are better able to avoid this “premature closure,” they still make many errors. This is not surprising, given the existence of confirmation bias and a tendency to focus on a single possibility in adulthood. Indeed, Mynatt, Doherty, & Dragan (1993) argue that even adults can only focus on and gather information pertaining to one explanation for a data set at a time.

Thus, the standard mental models account suggests that both MP and AC inferences could occur during comprehension in contrast to the syntactic account that predicts only MP inferences. In terms of cognitive economy, good reasons exist to support these predictions. Because this account suggests that participants tend to represent only a single possibility for any conditional premise, it is strengthened by the finding that participants in a variety of tasks seem to prefer to represent one possibility.
However, a sizeable body of research on deliberate conditional reasoning suggests that AC inferences can be systematically enhanced or suppressed, and to this work we now turn, along with a theory of these effects. An examination of this effect in the context of comprehension is important to determine whether conditional inferences in comprehension are the same as those drawn in deliberate reasoning tasks, and hence whether reasoning is essentially a part of the comprehension process.

**Modified mental models theory.** The AC enhancement and suppression effects are compatible with another version of mental models theory developed by Markovits (1984, 1985, 1993). In this theory, conditional processing depends critically on the strength of the memory association between a conditional’s antecedent and consequent. When the conditional’s antecedent and consequent are strongly linked in semantic memory (for example, in a relation of cause and effect, or when one is a category and the other is a highly typical exemplar of that category), participants should draw both MP and AC inferences. For example, participants given the conditional premise below, together with the appropriate minor premise, tend to draw both AC and MP inferences:

*If a dog has fleas, then it will scratch constantly* (Quinn & Markovits, 1998, p. B99).

(Note that the antecedent and consequent in this example were strongly associated as determined by a preliminary study, not by assumption.)

This effect should occur because the conditional’s representation in such cases tends to be a conjunction of the given antecedent and consequent:

\[
\text{dog has fleas} \quad \text{it scratches constantly}
\]
Because the two propositions are strongly linked in memory, no alternative possibilities are likely to be available to participants. Thus, the person should draw MP and AC inferences in the manner described above for standard mental models. However, when an antecedent and consequent are weakly linked in memory, the comprehender’s representation should be different. Consider this premise:

*If the dog has a skin disease then the dog scratches constantly* (Quinn & Markovits, 1998, p. B99).

Participants given such weakly associated antecedents and consequents should retrieve *alternative antecedents* from memory and form a representation that includes a conjunction of the alternative antecedent and consequent. The representation of the premise above would be as follows:

\[
\begin{align*}
\text{dog has skin disease} & \quad \text{it scratches constantly} \\
A & \quad \text{it scratches constantly}
\end{align*}
\]

The representation of the conditional premise contains two possibilities (hence the use of two lines). One possibility is the conjunction of the given antecedent and consequent. A second possibility is the conjunction the consequent and an alternative antecedent, \(A\), that is retrieved from memory when the premise is presented (this alternative is likely to be the most strongly associated antecedent [Markovits, et al., 1998; Janveau-Brennan & Markovits, 1999]). Because this second possibility is present in memory, someone given the consequent should not infer the given antecedent as a conclusion; thus the AC inference should be blocked. However, reasoners who are instead given the antecedent as a minor premise should draw the MP inference, because the state of affairs described in either antecedent brings about the consequent’s state of affairs.
Many studies provide results supporting these predictions with respect to deliberate conditional reasoning. Studies with adults demonstrate that AC inferences are reliably made when no alternative antecedents should be available in memory, but are much less likely when alternatives are explicitly supplied (Byrne, 1989; Byrne, Espino, & Santamaria, 1999) or readily available (Cummins, 1995; Hilton, 1990; Quinn & Markovits, 1998; Thompson, 1994). Some evidence suggests that the effects occur, though to a lesser extent, when the reasoning materials are fairly abstract, so that participants must represent abstract alternatives (Oaksford, Chater, & Larkin, 2000; Schroyens, Schaeken, Fias, & Y'Dewalle, 2000). The finding with explicitly supplied counterexamples has occurred in one study of children (Rumain, et al., 1983), and the effect of antecedent availability in memory has been demonstrated with children and adolescents (Klaczyński & Narasimham, 1998; Markovits, et al., 1998; Janveau-Brennan & Markovits, 1999). One researcher replicated this data pattern in a recognition measure of inferences drawn during comprehension (Franks, 1997), although this study's design (asking explicit questions and allowing participants to re-read texts) was more like a deliberate reasoning study. In all of these studies, the tendency to draw MP inferences was unaffected, and they were usually drawn at ceiling rates. It should be noted that the effects with AC seem to be due to the availability of counterexamples; other possible explanations have not been supported (Byrne, et al., 1999; Schroyens, et al., 2000).

Ironically, Braine (1978) notes, in a few sentences, that the same type of finding occurred in a few very early studies. However, he makes no attempt to incorporate this content effect into his proposal for a mental logic!
Modified mental models theory also has a developmental component. Markovits suggests that the retrieval of alternative antecedents becomes more flexible with age, an idea derived from Piaget's writings on possibility (1987a). Piaget proposed that children become more adept, with age, at retrieving and representing possibilities compatible with a given situation. Markovits predicts that the retrieval of alternative antecedents to conditional premises becomes a more flexible and abstract process with age. Young children, below age 7 or 8, are limited to alternatives that they have directly experienced, and in many cases will be unable to retrieve anything. Preadolescents should be able to retrieve some types of category-based alternatives based on world knowledge, along with memories for directly experienced events. Some adolescent and adult reasoners should be able to represent alternatives in the abstract; they should understand that although a conditional's antecedent is sufficient for its consequent, other alternatives exist that could also bring about the consequent (even if one doesn't know what they are). However, although children and preadolescents are both "slaves" to memory retrieval processes more than adults, even adults can be affected by these processes (Quinn & Markovits, 1998).

The results of the adult studies and developmental studies together indicate that both preadolescents and adults should be susceptible to the enhancement and suppression effects. Indeed, despite the age predictions by Markovits (1993), a main difference between preadolescents and older participants that he and colleagues have found, in the domain of conditional reasoning, is that preadolescents experience more difficulty in reasoning with contrary-to-fact and very abstract premises (Markovits & Vachon, 1989, 1990). The processing of such premises is a separate question not addressed in current
work, as reported studies use premises involving concrete content that should be plausible and sensible based on general knowledge. Also, although children younger than age 8 seem to have difficulty retrieving alternative antecedents, participants by the age of 11-12 are much more adept at doing so (Janveau-Brennan & Markovits, 1998; Markovits, et al., 1998).

The enhancement and suppression effects associated with conditionals have, with one possible exception previously noted (Franks, 1997), been demonstrated only in the context of deliberate reasoning tasks. It is unclear whether they would appear in comprehension tasks that did not instruct participants to reason. Thus, the issue of search for counterexamples (which is directly related to the availability of alternatives) must be addressed.

The issue of search for counterexamples in semantic theories. Both mental models theories are unclear about whether the third step in their hypothesized reasoning process, search for counterexamples, should occur during comprehension. According to the theories, this process is necessary to avoid AC inferences. The standard theory argues that a person must “flesh out” the model of the conditional to include possibilities in which the antecedent is false; these possibilities would serve as counterexamples and prevent the AC inference (Johnson-Laird & Byrne, 1991). The modified theory proposes that this process is critically affected by the strength of association in memory between a conditional’s antecedent and consequent (Markovits, 1993). As has been described, much research in fact supports the modified theory with respect to deliberate reasoning.
Thus, if inferences during comprehension do in fact parallel the inferences made during deliberate reasoning, in that people utilize counterexamples to inferences, the evidence obtained from studies of inference during comprehension should be like that obtained in studies of deliberate inference. Both MP and AC inferences should occur when conditionals’ antecedents and consequents are strongly associated, but only MP inferences should occur when the association is weak. Ironically, such findings would support the assertion that reasoning is a component of comprehension, per syntactic theories (e.g., Braine & O’Brien, 1991); of course, AC inferences are not easily handled by those theories.

If search for counterexamples is a separate process occurring only during deliberate reasoning, however, the parallel should not be expected. Searching for counterexamples might be too computationally demanding to occur automatically during comprehension. The search process requires a person to consider the semantic principle of validity (an inference is true, given the truth of its premises, only if no counterexamples exist to falsify it), but it is doubtful that people try to falsify inferences in comprehension (Kintsch, 1998). Importantly, studies of syllogistic reasoning suggest that search occurs only when participants are explicitly asked to reason, and even then only when participants’ initial mental models of the premises do not support putative conclusions (Byrne, et al., 1999; Evans, 2000; Evans, Handley, Harper, & Johnson-Laird, 1999; Newstead, Handley, & Buck, 1999; see also Polk & Newell, 1995). Only a small minority of participants seem to search in all cases (see Stanovich, 1999). The process
may be easier if counterexamples are actually provided (Byrne, 1989; Byrne, et al., 1999; Rumain, et al., 1983), but it seems unlikely to occur at all unless the participant is explicitly asked about the necessity or validity of a conclusion.

Thus, it is predicted that the systematic suppression and enhancement of AC inferences found in deliberate reasoning studies will not be found in studies of conditional inferences drawn during comprehension. Instead, both MP and AC inferences could be expected because both can occur quickly and do not involve any search for counterexamples. That is, these inferences do not even require step three of the standard mental models theory (search for counterexamples) for their occurrence. Additionally, this prediction is consistent with findings on human comprehension, hypothesis formation, induction, and decision-making. This pattern of findings would in turn point to important differences between comprehension and deductive reasoning.

To summarize predictions of semantic theories, the standard mental models theory (Johnson-Laird & Byrne, 1991) proposes that both MP and AC inferences should occur in comprehension, especially if it is assumed that people do not search for counterexamples to inferences that they draw during comprehension. The modified mental models theory (Markovits, 1993) instead suggests that AC inferences should occur during comprehension when a conditional’s antecedent and consequent are strongly associated in memory—this is because no counterexamples are easily available to these inferences. When a conditional’s antecedent and consequent are not strongly associated, AC inferences should not occur because counterexamples are easily available. MP
inferences should occur regardless of the strength of association. These modified mental models predictions imply that search for counterexamples is not a separate stage, in that counterexamples are retrieved automatically during comprehension of the conditional.

**Cognitive-developmental theory**

In the developmental literature, a prominent theoretical approach derives from Piaget's last writings concerning logical inference (1987a, 1987b). Piaget argued that a fundamental change in reasoning competence occurs during the transitions from childhood to adolescence. Full reasoning competence should not develop until adolescence, when conditional statements can be represented veridically by the standards of classical logic (see also Inhelder & Piaget, 1958). However, Piaget (1987a, b) made the representation of conditional relationships the most fundamental component of reasoning competence; this aspect differs from Inhelder and Piaget (1958). Essentially, Piaget argued that adolescents develop the ability to represent all the possibilities compatible with a conditional premise, whereas preadolescents are unable to do so. This ability appears as part of a more general ability to generate all the possibilities compatible with a statement's logical semantics and to understand the necessary relationships that exist among these possibilities, as well as to distinguish necessary inferences from those that are only possible (Pieraut-Le Bonniec, 1980).

Piaget's ideas inspired the conditional reasoning theory of Overton (1990) and colleagues, which broadly agrees with other cognitive-developmental theories of reasoning (e.g., Franks, 1996; Moshman, 1999; Moshman & Franks, 1986). A key prediction is that children and preadolescents below about age 14 will either misrepresent conditional statements as biconditionals, or will represent only the temporal or causal
relationship that a conditional describes. Misrepresentation occurs because these participants do not understand the asymmetric relationship between a conditional’s antecedent and consequent; they do not understand that the antecedent is merely sufficient for the occurrence of the consequent, but the consequent is necessary for the occurrence of the antecedent. As a result, these participants should often fall prey to committing conditional fallacies such as AC. Older participants understand the asymmetry, represent conditionals veridically, and avoid the fallacies; they can transcend causal or temporal relationships and represent the abstract logical relationship between a conditional’s antecedent and consequent.

The theory also posits that conditionals can be interpreted veridically whenever a relation of *relevance* holds between antecedent and consequent. Roughly speaking, a relevance relation connects antecedent and consequent in some type of meaningful fashion (a formal definition is given in Ricco, 1990). For example, the conditional *If I have enough money, I’ll buy a CD player* contains a causal relation of relevance between antecedent and consequent (in this case, the antecedent event is an enabling condition that could help cause the consequent event). Relevance interacts with developing logical competence in that it can bootstrap logical performance for adolescents (at least 14 years old) but will not significantly improve performance for younger participants (Lourenco & Machado, 1996; Overton, 1990; Ward & Overton, 1990). In other words, relevance is necessary but not sufficient for logical reasoning performance; the requisite competence must be in place.
Where does the cognitive developmental theory fall in the syntactic/semantic debate? This theory agrees with syntactic views in predicting that adults should tend not to fall prey to committing AC inferences. One adherent of the general cognitive developmental approach also proposes that logical inference schemas may be bootstrapped from relevant linguistic experiences, given the appropriate underlying competence; this author clearly aligns the cognitive developmental and syntactic theories (Falmagne, 1990). It disagrees in proposing that preadolescents will in general commit AC inferences more often than older reasoners. With respect to mental models, cognitive developmental theory is similar in tying reasoning competence to participants' representations of possibilities and notions of logical necessity. However, cognitive-developmental theory differs by arguing that mature reasoners represent the logical semantics of conditionals accurately, whereas mental models theories argue that even adults are limited by working memory, considerations of relevance, and availability of counterexamples in memory.

Importantly, cognitive-developmental theory is primarily a theory of how logical competence improves with age, rather than a theory of real-time cognitive processing. Possibly, this may reflect the original intentions of Piaget, who wanted first to identify and describe structures of thought without specifying how they would be functionally implemented (Lourenco & Machado, 1996). As such, it leaves the nature of the procedures that implement logical competence largely unspecified (Overton, 1990). Describing a set of processes or steps that occur during conditional processing is thus difficult within the confines of this theory. A related problem is that the theory is unclear on how closely the representations and processes involved during comprehension match
those formed during reasoning. However, statements of this theoretical viewpoint argue strongly that a person’s conditional representations, apparently in a variety of contexts, are largely determined by the person’s understanding, or lack thereof, of the logical relationship between a conditional’s antecedent and consequent (Overton, 1990; Overton, et al., 1999; Ricco, 1990).

Finally, many of the cognitive-developmental studies are based upon the Wason selection task (O’Brien & Overton, 1980; Overton, Ward, Black, Noveck, & O’Brien, 1987; Ward, et al., 1990; Ward & Overton, 1990). As noted earlier, some researchers argue that this task involves processes other than those of deductive reasoning. Some findings with children support these views (Klaczynski & Narasimham, 1998; Light, Blaye, Gilly, & Girotto, 1989; Light & Legrenzi, 1990), suggesting that the data of Overton and colleagues do not pertain to deductive inference. Children and adolescents seem to be affected by linguistic pragmatics and “cheater-detection” considerations much like adults, so that different problem contexts can lead to logically perfect and logically erroneous performance. Thus, Overton and colleagues’ conclusions must be taken with caution. In particular, general claims about conditional processing and representations across tasks should not be taken at face value, given the overreliance on the Wason selection task.

To summarize, cognitive-developmental theory predicts that preadolescents are incapable of accurately representing conditional relationships. They should thus tend to commit AC inferences as well as MP inferences, and even “relevant” content should not
lead to logically accurate performance in this age group. However, adults should be more likely to avoid AC inferences and commit only MP inferences, because adults are capable of representing conditionals in a logically veridical manner.

Summary of theoretical predictions

Syntactic theories (Braine & O'Brien, 1991; Braine & Rumain, 1983; Lea, 1995; Rips, 1994) argue that people automatically abstract the propositional form of incoming discourse and apply, whenever possible, logically valid inference schemas to that discourse to produce conclusions. The inference process is part of comprehension and does not involve further, more deliberate processes. With conditionals, participants should apply a modus ponens (MP) schema to the discourse whenever possible to produce MP inferences. Invalid inferences such as affirming the consequent (AC) should not happen during comprehension, as there would be no ready explanation for them. These predictions should hold for preadolescents as well as adults since both groups possess the inference schemas in their logico-linguistic competence.

Standard mental models theory (Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) proposes that people represent true possibilities compatible with the logical structure of incoming discourse, but these possibilities will often deviate from normative standards. Conditionals are represented as essentially a conjunction of antecedent and consequent. During comprehension, both MP and AC inferences can be drawn because both can be drawn based on the initial representation of the conditional; this prediction is firmer if it is assumed that participants will not search for counterexamples to these inferences (an issue about which the standard theory is unfortunately vague). The
standard mental models theory is consistent with the notion that inferences during comprehension are only an initial stage, and that true reasoning involves the application of more deliberate processes beyond those involved in comprehension.

The modified mental models theory (Markovits, 1993) proposes that the strength of memory association between a conditional's antecedent and consequent determines how it is represented and thus the inferences that are drawn. Strongly associated antecedents and consequents will lead to both MP and AC inferences, but weakly associated propositions will lead to only MP inferences. However, these effects are predicted on the assumption that alternative antecedents are retrieved automatically in the case of weak memory associations. This approach implies that counterexamples, in the form of alternative antecedents, would be available immediately, rather than resulting from a separate search of memory. Neither version of mental models would predict any significant differences between preadolescents and adults.

Finally, cognitive-developmental theory suggests that conditional inferences are determined by how a person represents the conditional's antecedent and consequent and how the person represents the relationship between them (Franks, 1996; Overton, 1990). Preadolescents cannot conceptualize the abstract and asymmetric logical relationships of necessity and sufficiency that hold between an antecedent and consequent. Instead, they tend to misrepresent conditionals as biconditionals, and this representation will lead preadolescents to draw both MP and AC inferences. Adult participants understand the logic of conditionals and are capable of representing them according to logical standards.
As a result, they should still draw MP inferences but should be much less likely to draw AC inferences. Cognitive-developmental theory is unclear about the processes that implement logical competence, so few specific cognitive operations are posited.

It should be noted that no theory implies that propositional inferences are the only inferences drawn in comprehension. Everyday comprehension involves inferences that connect pronouns to antecedents, inferences about the probable causes and effects of events, the probable instruments used in various actions, and so on. The theoretical discussion just presented only pertains to inferences at the level of entire propositions, and which involve the logical connective if...then.

Overview of Experiments

The four experiments reported here examine the representations formed during the comprehension of conditional arguments through an on-line priming technique (Experiments 1 and 3) and a recognition memory technique (Experiments 2 and 4). These techniques allow insights into the representations and inferences generated during comprehension and are standard in discourse processing research (see reviews in Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; McKoon & Ratcliff, 1992; Zwaan & Radvansky, 1998), although they are less common in the deductive reasoning area. Experiments 1 and 3 were conducted with adult participants and address syntactic and semantic theories, although results may also have implications for cognitive-developmental theory. Experiments 2 and 4 were conducted with adults and preadolescents. These experiments address cognitive-developmental theory as well as the other two views; these experiments can also provide converging evidence for the data of Experiments 1 and 3.
Experiments 1 and 3 are illustrated by reference to Lea (1995). Lea presented stories containing MP premises in an Inference condition and a No-Inference condition in which the story's last sentence was slightly altered. The fourth and final sentence distinguished stories in the two conditions:

Special Agent Evans was trying to crack a Cuban cigar smuggling ring. As he entered the dark warehouse he whispered into his walkie-talkie, “If the dog starts to bark, then we know there’s tobacco around.”

“Ok, I’ll be listening,” replied his partner, Newstead, who was monitoring the situation from their unmarked patrol car.

INF: There was quiet for a few minutes, then the silence was broken by the sound of Evans’ dog barking.

NO INF: There was quiet for a few minutes, and Newstead wondered whether he was close enough to hear the sound of the dog barking. (Lea, 1995, p. 1470)

Participants made lexical decisions to letter-string probes presented immediately after each story. When the letter strings were actual words, the words were those that are most strongly associated with concepts in conclusions to the argument (Jenkins, 1970; Keppel & Strand, 1970); in the above example, the word would be SMOKE, the primary associate of tobacco.

Participants responded significantly faster to words presented after Inference versions than to the same words presented after No-Inference versions. This result presumably occurred because Inference versions supplied input conforming to the MP schema, but No-Inference versions did not. Participants inferred the MP conclusion (e.g., tobacco), and the conclusion primed words to which it was strongly related.

This result, though suggestive, does not favor one theory over another. Standard mental models theory predicts that participants can draw MP inferences based only on the initial representations of a conditional premise and the categorical premise. Lea’s priming effect seems compatible with mental models theory. A more informative test of
the two theories would be to include stories that supply the premises for an AC inference. Standard mental models theory proposes that this inference is easily drawn because the inference also requires only the initial representations of the conditional and categorical premise. Experiments 1 and 3 therefore replicated Lea (1995) but included stories with AC premises.

The priming task used in my experiments was speeded word recognition, rather than lexical decision. In speeded word recognition, participants decide whether a certain probe word was actually presented in the story and answer yes or no as quickly as possible; for each critical story, the probe word represented the focal noun in the inference. If a participant has just drawn an inference including a given concept, that concept should be more activated in memory. Assuming that the concept is represented as a proposition in a typical predicate-argument structure (e.g., Kintsch, 1998) the concept corresponding to the focal noun (argument) should be activated. Recognition of a word denoting that concept should thus be facilitated compared to a baseline condition that requires more effortful memory retrieval. On the other hand, if an inference involving that concept has not been drawn, recognition of that word should not be facilitated. Although speeded word recognition has not been used in the study of deductive inference, the task has a substantial history in other areas of discourse processing. It has figured in studies of anaphoric inference (McKoon & Ratcliff, 1980), predicting and bridging inferences (Fincher-Kiefer, 1995, 1996; McKoon & Ratcliff 1986, 1989, 1992), sentential connectives' influence on integration of text (MacDonald & Just, 1989; Millis & Just, 1994), and the basic processes of sentence comprehension (e.g., Gernsbacher, 1997; Gernsbacher, Hargreaves, & Beeman, 1989).
One caveat: as the example story suggests, the key word in the inferred conclusion is not the only concept that could prime a word probe. Tobacco is not the only possible source of priming for the probe smoke in the sample story; the word cigar in the first sentence could be another. However, the Inference and No-Inference versions share the first three sentences, so this source of priming is held constant. Additionally, in my experiments, the fourth sentence for in the No-Inference version of each story contained many, if not all, of the words in the Inference version that could presumably prime the lexical decision probe. This design factor is necessary to ensure that sources of priming unrelated to making an inference are controlled (Keenan & Jennings, 1995). Other sources of priming can never be ruled out with absolute certainty when the stimuli are as complex as short stories, but obtaining predicted results using the proper control (No-Inference) condition would give such results more credibility.

To test predictions of standard and modified mental models theories, it is necessary to use conditionals that vary in the strength of the relationship between their antecedents and consequents. Thus, the conditionals in Experiment 1 were constructed with strongly associated antecedents and consequents, which were identified in a preliminary study described in Chapter 2. In Experiment 3, these same conditionals were modified so that they instead involved weakly associated propositions (identified in the same preliminary study). All other aspects of design and procedure in the two experiments were identical.

Experiments 2 and 4 used a recognition memory task. In this task, the participant reads texts that allow, but do not state, certain inferences. Afterward, the participant is asked to verify whether certain information was presented in these texts, and some of this
information will be the content of the inferences. Presumably, if the participant draws the inference, it will be encoded in the participant's memory representation of the text and will tend to be remembered as actually appearing in the text.

The use of various recognition tasks to study inferences is common in both the developmental and the adult psycholinguistic literature. Developmentally, these memory tasks have been used to study spatial inferences (e.g., Brainerd & Reyna, 1993; Reyna & Kiernan, 1994), causal inferences (e.g., Reyna & Kiernan, 1994; Paris & Upton, 1976), instrument inferences (Marx, 1992; Paris & Lindauer, 1976), and knowledge-based elaborative inferences (e.g., Barnes, Dennis, & Haefele-Kalvaitis, 1996). A few experiments have also studied deductive inferences, including those based on conditionals, with evidence suggesting that MP inferences are drawn by children and university students (Franks, 1996). In the adult literature, many seminal studies of inference in discourse processing used memory measures (e.g., Bransford, Barclay, & Franks, 1972; see review in Singer, 1994). Other adult studies have used memory measures to assess the effect of sentential connectives on text integration (e.g., Caron, Micko, & Thuring, 1988; Murray, 1997). Still other studies have used such measures to assess the tendency to draw transitive inferences based on set relations and other relations (e.g., Griggs, 1976; Griggs & Warner, 1982; Nguyen & Revlin, 1993). Finally, one set of experiments used a memory measure to find that MP inferences were made during comprehension; these researchers did not examine invalid inferences (Lea, et al., 1990).

This task is appropriate for use with preadolescents and adults. However, the task is methodologically questionable with younger children. They are less likely to integrate information across sentences to make inferences, especially bridging inferences, because
they may forget premise information (e.g., Ackerman & Mannes, 1993; Barnes, et al., 1996). In some situations children may not notice pairs of contradictory sentences within a story (Markman, 1979). Also, many studies suggest that younger children overall are less likely to remember story information veridically and thus more likely to confuse inferences with what was actually presented (Ackerman & Mannes, 1993; Paris & Carter, 1973; Paris & Lindauer, 1976; but see Reyna & Kieman, 1994). Therefore, testing children younger than 10 years old in this type of task can yield findings that are difficult to interpret – on the one hand, children are less likely to integrate information, but on the other hand, they are more likely to respond based on gist. However, reviews of studies using older, preadolescent participants suggest that these memory limitations should be greatly reduced (Ackerman & Mannes, 1993; Barnes, et al., 1996).

The research design is the same as that for the priming study: participants read equal numbers of MP and AC stories, some in Inference versions and some in No-Inference versions (again, which story version each participant received was counterbalanced). An example of one of the stories illustrates the method:

Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was. He thought that if it was dark outside, then it was night. Still feeling sleepy, Frank arose to look out a window. INF: He saw that it was dark outside. NO INF: He wondered whether it was dark outside.

Note that the Inference version of this story contains sentences fitting the MP argument form. After reading each story and answering the comprehension question, participants were presented with three sentences, in random order. For each sentence, participants decided whether the information in the sentence had been explicitly
presented in the story or not. One Inference test sentence was a conclusion to an AC or MP argument. Referring to the example given above, this test sentence was The time of day was night. A second test sentence was a Paraphrase of a proposition actually presented in the story; an example was Frank glanced out of a window. Finally, a third test sentence was a Changed sentence that altered one noun phrase stated in a text sentence but otherwise used the same words involved in that proposition. An example of a changed sentence was Frank woke up on his bunk bed. Three test sentences of the same types were presented after each filler story, although inference sentences following these stories were not deductive consequences of these stories.

Predictions are analogous to those for the online priming task; the critical test sentences are the Inference test sentences. Mental logic theories predict a high rate of false alarms to MP-Inference stories; that is, participants should agree that the information in the conclusion was in fact stated in the story. Participants should decide that AC inferences were not stated since no such inference procedure is applied in comprehension. Standard mental models theories predict that false alarm rates should be relatively high for both types of inferences. The modified mental models theory predicts that the type of conditional premise used (i.e., containing strongly associated propositions versus weakly associated propositions) should impact results. With strongly associated antecedents and consequents, participants should false alarm on inference test sentences following MP and AC inference stories; in the case of weak associations, participants should false alarm only on inference test sentences following MP stories. Finally, the cognitive-developmental theory suggests that preadolescents should false alarm on both
AC and MP inferences because they cannot veridically represent conditional arguments and treat them as biconditionals. Adults, however, should false alarm only on MP inferences.

As with the online priming experiments, modified mental models theory was tested by altering the content of conditional premises used in the critical stories. Thus, the conditionals in Experiment 2 were constructed with strongly associated propositions, which were sampled from the same pool of conditionals used in Experiment 1. In Experiment 4, these same conditionals were altered so that they instead involved weakly associated propositions (these were sampled from the same pool of conditionals used in Experiment 3). Otherwise, all aspects of design and procedure between Experiments 2 and 4 were identical.

Conditionals in these experiments were written so that they would not be considered causal-temporal, or “lawlike” (Geis & Zwicky, 1971). This step was taken because it is generally agreed that such conditionals should not be normatively analyzed according to standards of propositional logic. They may instead be evaluated by assessing the necessity and sufficiency of antecedent for consequent, by computing the conditional probability of the consequent given the antecedent, or by some other analysis (Sanford, 1989). No reasoning theory predicts that propositional syntax or semantics would be the most important determinant of reasoning performance in such cases, hence this type of conditional was avoided.

The online priming experiments and the recognition memory experiments can provide converging evidence for one another. No one method of studying inferences during comprehension can give unambiguous results, since all methods are indirect (see
Potts, et al., 1988; Singer, 1994; Trabasso & Suh, 1993). For example, recognition memory judgments as to whether information representing inferences has actually been presented may be guided by general plausibility judgments rather than the generation of inferences. One must use multiple methods to determine whether the results yielded by each seem to converge on a common phenomenon. The reported experiments attempt to do exactly that, though of course converging results will be stronger for adults.
CHAPTER 2

PRELIMINARY EXPERIMENT

The purpose of this study was to identify a set of propositions for use in creating conditionals that consisted of strongly associated antecedents and consequents, as well as conditionals consisting of weakly associated antecedents and consequents.

This study generated a list of noun-attribute pairs that could be used to construct conditional premises. Participants in this experiment heard a list of nouns, one word every 20 seconds. For each noun, they listed the first attributes of the thing denoted by the noun that came to their mind in response to the word. This type of “free association” technique has been used to identify items that are strongly linked in semantic memory (e.g., Keppel & Strand, 1970). Strongly associated items such as this have produced predictable results in both psycholinguistic studies (e.g., Kintsch, 1998; Singer, 1994) and conditional inference studies (e.g., Cummins, 1995; Markovits, et al., 1998).

Participants listed up to three attributes for each word. From these listings, the most frequent attribute given in response to each noun could be found, as could one infrequently generated attribute. Thus, two noun-attribute pairs could be generated for each noun. The noun + most frequent attribute pair was used to generate a conditional premise that was considered strongly associated (in terms of the relationship between
antecedent and consequent), and the noun + infrequent attribute pair was used to create a conditional considered weakly associated with respect to its antecedent-consequent relationship.

Method

Participants. A total of 52 undergraduates from The Ohio State University participated voluntarily as part of an introductory psychology course. All participants were fluent English speakers.

Materials. The critical experimental materials consisted of a list of 66 common nouns, drawn from the nouns listed in Jenkins (1970), arranged in a random order. Additionally, each participant received a small booklet. The first page of the booklet contained instructions. The three subsequent pages contained blanks in which participants were to write their responses, with three blanks on each line.

Procedure. Participants were tested in groups of 6-10. The experimenter first read the instructions aloud. Instructions informed participants that the experimenter would read a list of nouns, one word at a time. After each noun, the participant had 20 seconds to list attributes that were associated with the thing that each noun represented. The experimenter informed participants before the study that an “attribute” could be any type of characteristic of that thing. Participants were to list the first attributes that came to mind, listing no more than 3 for each noun. The experimenter then read the words at a rate of one word every 20 seconds (a stopwatch aided in keeping time). The entire procedure took about 20 minutes.
Results

By tabulating participants’ responses, I could identify the most frequent attribute given in response to each noun. It should be noted that the numbers of participants included for the nouns night, sea, and priest were only 50, 49, and 51, respectively, because a few participants appeared to have misunderstood those words. To have a suitable number of stimuli for the experiment, I selected 40 noun-attribute pairs as the basis of stimuli construction. These pairs were those with the forty highest percentages for most frequent attributes. The most frequent attribute was given by 44-98% of participants in response to these nouns, with a mean of 64.1% (SD = 13.8%) participants listing the most frequent attribute for a given noun.

For the forty selected nouns, an infrequent attribute was also selected to construct weakly associated conditionals. Because most attributes in response to a given noun were infrequent, it was decided to randomly select an attribute that was listed by 5 to 8 participants (9.6 to 15.4% of participants). For these infrequent attributes, the mean percentage of participants listing them was 11.2% (SD = 2.0%). Table 2.1 lists the 40 nouns selected, along with the most frequent attribute and infrequent attribute chosen for each.

Conditional premises were constructed based on the noun-attribute pairs. As noted in the introduction, the conditional stimuli were written in a way that avoided causal/temporal conditionals; additionally, inspection of Table 2.1 reveals that very few of these conditionals could be considered analytically definitional. The conditionals were either of the circumstantial type in which the occurrence of one event leads to, but does
<table>
<thead>
<tr>
<th>Noun</th>
<th>Most Frequent Attribute (%)</th>
<th>Infrequent Attribute (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grass</td>
<td>Green (98.08)</td>
<td>Water it (9.62)</td>
</tr>
<tr>
<td>Lamp</td>
<td>Gives light (90.38)</td>
<td>Has a switch (11.54)</td>
</tr>
<tr>
<td>Scissors</td>
<td>Cuts things (88.46)</td>
<td>Made of metal (15.38)</td>
</tr>
<tr>
<td>Blossom</td>
<td>A flower (84.62)</td>
<td>Grows (9.62)</td>
</tr>
<tr>
<td>Snow</td>
<td>Cold (84.31)</td>
<td>Found in winter (15.38)</td>
</tr>
<tr>
<td>Night</td>
<td>Dark (82.00)</td>
<td>Cold (12.00)</td>
</tr>
<tr>
<td>Candy</td>
<td>Sweet (76.92)</td>
<td>Colorful (11.53)</td>
</tr>
<tr>
<td>Lemon</td>
<td>Sour (76.92)</td>
<td>Small (9.62)</td>
</tr>
<tr>
<td>Crowd</td>
<td>Has people (76.92)</td>
<td>Annoying (9.62)</td>
</tr>
<tr>
<td>Noise</td>
<td>Loud (75)</td>
<td>Scary (9.62)</td>
</tr>
<tr>
<td>Stove</td>
<td>Used to cook (73.08)</td>
<td>In kitchen (11.54)</td>
</tr>
<tr>
<td>Leaf</td>
<td>On Trees (71.15)</td>
<td>Has many colors (11.54)</td>
</tr>
<tr>
<td>Sheep</td>
<td>Has wool (71.15)</td>
<td>Lives on farms (13.46)</td>
</tr>
<tr>
<td>War</td>
<td>Killing (69.23)</td>
<td>Sad (9.62)</td>
</tr>
<tr>
<td>Sea</td>
<td>Water (67.35)</td>
<td>Big (10.20)</td>
</tr>
<tr>
<td>Bath</td>
<td>Water (67.31)</td>
<td>Uses soap (15.38)</td>
</tr>
<tr>
<td>Book</td>
<td>Is read (67.31)</td>
<td>Learn from it (15.38)</td>
</tr>
<tr>
<td>Bed</td>
<td>Sleep in it (63.46)</td>
<td>Has blanket (11.54)</td>
</tr>
<tr>
<td>Hand</td>
<td>Has fingers (63.46)</td>
<td>Useful (11.54)</td>
</tr>
<tr>
<td>Hammer</td>
<td>Pounds nails (61.54)</td>
<td>Heavy (11.54)</td>
</tr>
<tr>
<td>Truck</td>
<td>Big (59.62)</td>
<td>Hauls things (9.62)</td>
</tr>
<tr>
<td>Mountain</td>
<td>High (59.62)</td>
<td>Has trees (9.62)</td>
</tr>
<tr>
<td>Table</td>
<td>Wooden (59.62)</td>
<td>Place things on it (11.54)</td>
</tr>
<tr>
<td>Thief</td>
<td>Steals (59.62)</td>
<td>Wears black (9.62)</td>
</tr>
<tr>
<td>Child</td>
<td>Small person (55.77)</td>
<td>Cute (13.46)</td>
</tr>
<tr>
<td>Mother</td>
<td>Caring (55.77)</td>
<td>Cooks (9.62)</td>
</tr>
<tr>
<td>River</td>
<td>Water (53.85)</td>
<td>Has boats (9.62)</td>
</tr>
<tr>
<td>Heat</td>
<td>Warm (53.85)</td>
<td>Uncomfortable (9.62)</td>
</tr>
<tr>
<td>House</td>
<td>Lived in (53.85)</td>
<td>Has a roof (13.46)</td>
</tr>
<tr>
<td>Priest</td>
<td>Works in church (52.94)</td>
<td>Dressed in black (9.80)</td>
</tr>
<tr>
<td>Fruit</td>
<td>Sweet (51.92)</td>
<td>Has seeds (9.62)</td>
</tr>
<tr>
<td>Spider</td>
<td>Has eight legs (51.92)</td>
<td>Little (9.62)</td>
</tr>
<tr>
<td>Carpet</td>
<td>Soft (50)</td>
<td>Fuzzy (15.38)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Yellow (50)</td>
<td>White (9.62)</td>
</tr>
</tbody>
</table>

Table 2.1: The noun-attribute pairs used to create conditional premises in the four reported experiments, along with percentages of participants listing each of these attributes.
Table 2.1 (continued)

<table>
<thead>
<tr>
<th>Noun</th>
<th>Most Frequent Attribute (%)</th>
<th>Infrequent Attribute (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whiskey</td>
<td>Makes one drunk (44.23)</td>
<td>Smells (9.62)</td>
</tr>
<tr>
<td>Food</td>
<td>Eat it (44.23)</td>
<td>Hot (9.62)</td>
</tr>
<tr>
<td>Butterfly</td>
<td>Colorful (50)</td>
<td>An insect (9.62)</td>
</tr>
<tr>
<td>Window</td>
<td>Made of glass (50)</td>
<td>Square (9.62)</td>
</tr>
<tr>
<td>Head</td>
<td>Contains brain (50)</td>
<td>Round (13.46)</td>
</tr>
<tr>
<td>Socks</td>
<td>Keep feet warm (48.08)</td>
<td>Soft (9.62)</td>
</tr>
</tbody>
</table>

not cause, the occurrence of a second event (Scholnick & Wing, 1991), and other conditionals expressed a relationship between a thing and a characteristic of that thing. It should also be noted that, because the relationships between antecedents and consequents were always meaningful in some sense, all of the conditionals would be considered at least somewhat “relevant,” in terms of Ricco (1990). Thus, if it were found that participants seemed to draw invalid AC inferences with these conditionals, results could not be ascribed to the lack of a relevance relationship between antecedents and consequents.
CHAPTER 3

EXPERIMENT 1

In this experiment, adults read short stories in either the MP or AC inference forms (along with appropriate control stories), one sentence at a time. Immediately after the last sentence of each story, they decided whether a probe word had been presented in the story and pressed one of two buttons on a response box to answer yes or no. Some of the probe words represented focal concepts in the MP or AC inferences (depending on experimental condition). If participants are in fact drawing these inferences, they should respond to these words more quickly than participants who respond to the same words, but read control versions of the stories instead of the inference versions. Syntactic theories suggest that participants should show priming (faster responses to probe words) for MP inferences, whereas both versions of mental models theory suggest that participants should show priming for each type of inference.

Method

Participants. A total of 64 undergraduates from The Ohio State University (37 women and 27 men) with a mean age of 18.7 years (SD = 1.7 years) voluntarily participated as part of an introductory psychology course. All participants were fluent English speakers.
**Materials.** The critical stimuli consisted of 40 stories, each four sentences long. Each story contained in its second sentence a conditional premise formed from one noun-attribute pair. Each story was written in four versions that corresponded to the four experimental conditions. Examples of each version, using several noun-attribute pairs, are displayed in the Appendix. The MP-Inference version contained second and fourth sentences fitting the form *If Attribute then Noun; Attribute*. The MP-No Inference version was identical except that the fourth sentence mentioned *Attribute* without asserting it to be the case. The AC-Inference version contained second and fourth sentences of the form *If Noun then Attribute; Attribute*. The AC-No Inference version was identical to the AC-Inference version except that the fourth sentence mentioned *Attribute* without asserting it.

Each probe word, which was the noun of that noun-attribute pair, was focal in the MP and AC inferences suggested by the MP-Inference and AC-Inference versions. Also, the probe was the same in all four versions of each story. Importantly, within each argument form, the Inference and No-Inference versions were identical except for the fourth sentence, as described above. However, even the fourth sentences were written to be as similar as possible and to contain many of the same concepts, to control for simple interlexical priming effects that can inflate estimates of the priming effect due to generation of inferences (Keenan & Jennings, 1995).

The 40 critical stories were randomly split into four groups of ten stories each. Participants were also assigned randomly to one of four groups (*n* = 14 to 17 per group). Story groups and participant groups were combined in a Latin square design. There were
thus four groups of participants, with members of each group responding to only one
version of each critical story. Each participant responded to 10 stories in each of the four
experimental conditions, and no participant received the same critical story twice.

Along with the 40 critical stories (all of which required Yes answers to the word
probes), there were 72 four-sentence fillers. Forty of these fillers were similar to the
critical stories—they fit one of the four versions and were based on other noun-attribute
pairs. For each of these fillers, the probe word was not given in the story, so they
required No answers. The final 32 fillers were each four sentences long but lacked
conditional premises. Half of these fillers required Yes answers and half required No
answers. Participants thus received 112 stories in all.

**Design and procedure.** The experimental design was a 2 Argument Form (MP,
AC) X 2 Version (Inference, No Inference) design. Both factors were within-subjects.
Stimulus presentation and data collection were controlled by a PC that ran Superlab Pro
2.0 for Windows (Cedrus Corp., 1997).

Participants were tested individually. Each was told that she would read a series
of stories on the computer, and that after each story, a word would be presented. She was
to decide whether the word appeared in the story and respond as quickly as possible by
pressing one of two buttons (one for yes, another for no) on a four-button response box.
Also, she would be asked a yes/no comprehension question after each story (questions
were included simply to encourage attentive reading, and their answers did not pertain to
any conditional inference that the participant could draw). Half of the questions had yes
answers, and half had no answers. Participants were not told to reason, and reasoning
was never mentioned.
Each participant received 12 practice stories before the 112 experimental stories. The 12 practice stories were not labeled as "practice." They were simply a separate block of stories that were included to allow participants to develop some facility in using the response box (pilot testing indicated that participants tended to make many errors in the first several stories, possibly due to the requirement to respond quickly). The participant read each story in self-paced fashion. The first sentence appeared, and as soon as she finished it, she pressed a third key, marked NEXT, on the response box, to clear the screen and display the next sentence. This process continued until the fourth sentence. As soon as the participant pressed NEXT, the screen cleared and the word probe appeared in asterisks (** WORD **). Once she responded, the screen cleared and the comprehension question appeared. Once the question was answered, the screen cleared, and PRESS NEXT TO CONTINUE appeared. As soon as the participant did so, the procedure repeated for the next story. All stimuli were presented in the same centered position on-screen in 14-point type. Order of story presentation was randomized for each participant by the program. The overall duration of the experiment was approximately 45-50 minutes.

Results and discussion

In analyses of reaction times, wrong answers were discarded. Then, for each participant, reaction times that were at least 2.5 standard deviations above or below the participant's grand mean were discarded, a procedure that removed 2.5% of all responses. Participants' mean reaction times to the word probes in each condition, along with error
Table 3.1: Mean reaction times in milliseconds (SE’s in parentheses), and mean error rates in percentages, by experimental condition, Experiment 1.

<table>
<thead>
<tr>
<th>Argument Form</th>
<th>Inference</th>
<th>No-Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>1063 (29), 10.3</td>
<td>1143 (32), 13.6</td>
</tr>
<tr>
<td>AC</td>
<td>1083 (35), 8.4</td>
<td>1191 (35), 11.4</td>
</tr>
</tbody>
</table>

Mean reaction times were analyzed in two 2 X 2 repeated measures analyses of variance (ANOVAs), with Argument Form and Version as within-subject factors. The first analysis ($F_1$) used mean reaction times for each participant in each experimental condition and thus used error terms based on subject variability. The second analysis ($F_2$) used mean reaction times for each item in each experimental condition and thus error terms based on item variability. This data analytic technique is standard in psycholinguistic research, following Clark (1972), and has been used in virtually every study of discourse processing since then. Results of the first ANOVA found a main effect for Argument Form, $F_1(1, 63) = 4.24, p < .05$, with the MP condition faster than the AC condition. The first ANOVA also yielded a main effect for Version, $F_1(1, 63) =$
33.88, \( p < .001 \), with Inference versions faster than No-Inference versions. Results of the second ANOVA also gave main effects for Argument Form, \( F_2(1, 39) = 4.64, p < .05 \), and Version, \( F_2(1, 39) = 21.91, p < .001 \). The pattern of these results was identical to that in the \( F_1 \) analyses. In neither analysis was the Argument Form X Version interaction significant, \( F_1(1, 63) = 1.22, p > .25 \), and \( F_2(1, 39) = 3.19, p = .082 \).

A repeated measures 2 X 2 ANOVA, with the same factors as before, was also conducted on the numbers of errors committed by each participant in the four experimental conditions. This analysis was conducted to look for any speed-accuracy tradeoffs that could compromise interpretation of results (i.e., if Inference versions were faster but also led to more errors, one could not interpret the reaction time data in terms of inferencing). In this analysis, a significant effect was found only for Version, \( F(1, 63) = 5.00, p < .05 \), with Inference versions (\( M = 9.4\% \), \( SD = 9.3\% \)) associated with fewer errors than No-Inference versions (\( M = 12.5\% \), \( SD = 11.3\% \)). Note that a speed-accuracy tradeoff cannot account for the priming results because accuracy in the Inference conditions was slightly higher than accuracy in the No-Inference conditions.

Thus, participants’ mean response times to probes following both MP- and AC-Inference stories were faster than responses to those probes following the corresponding No-Inference stories. Because the probes were focal nouns in non-presented inferences, participants seem to draw MP and AC inferences while reading. The reported data are consistent with mental models theories but difficult to capture in syntactic theories.

These results are only suggestive. This experiment is not a complete test of modified mental models theory; that test awaits Experiments 3 and 4. Obviously, this experiment does not test cognitive developmental predictions. A third issue is the
significant and unforeseen effect for Inference Form found in this experiment. There is no ready explanation for this finding. One researcher has observed that people more readily draw “forward” conditional inferences from antecedent to consequent, such as MP, than “backward” conditional inferences from consequent to antecedent, such as AC (Evans, 1993). The “forward” order may reflect how conditionals function in natural language (Sanford, 1989), and if-then conditionals are processed more quickly when possibilities specified by the premises are presented in the order of antecedent-consequent (Oberauer & Wilhelm, 2000). The faster MP times could thus reflect the relative ease of forward processing, though this explanation can only remain conjecture.

From a psycholinguistic perspective, the online priming effects may signify transient inferences that are only “minimally encoded” (McKoon & Ratcliff, 1992) and not fully encoded into memory representations of the texts. Some data do suggest that some types of predictive inferences are drawn only momentarily but not encoded into memory if the context does not strongly support them (Keefe & McDaniel, 1993; McKoon & Ratcliff, 1986). Thus, participants in this experiment might not agree that the texts actually contained these inferences if they are tested with an explicit memory measure requiring more deliberate judgment. On the other hand, if participants agree that the inferences had actually been presented, one could have more confidence that the inferences were being made. To examine this possibility, Experiment 2 was conducted.
CHAPTER 4

EXPERIMENT 2

Experiment 2 addresses some of the concerns that could be raised about Experiment 1. It provides a converging test for the results of Experiment 1 by looking for evidence that participants encode MP and AC inferences, at least with strongly associated conditional premises, into their memory representations of the texts. Also, Experiment 2 tested a sample of preadolescents, which allows a test of cognitive-developmental theory along with the other two theoretical views.

In this experiment, participants read short stories that conformed to either the MP or AC premise forms, although conclusions were not given. The experimental design was the same as Experiment 1. The stories were a subset of those used in Experiment 1. After reading each story, participants made yes/no judgments as to whether the information in three test sentences had actually been presented in the stories. For the stories conforming to the MP and AC inference forms, some of these test sentences represented these inferences. If participants do draw these inferences during comprehension, they should tend to agree that the information in these inferences (i.e., the test sentences) had been presented; they should not do so after control stories.
Theoretical predictions here are relatively straightforward. The syntactic prediction is that participants in both age groups should be more likely to false alarm on sentences representing MP inferences (when those sentences follow MP-Inference stories) than on sentences representing AC inferences. Both mental models theories predict that all participants should instead false alarm on test sentences representing MP and AC inferences when those sentences follow Inference stories. Finally, cognitive developmental theory suggests that adults should be more likely to false alarm on MP inferences, whereas preadolescents should tend to false alarm on MP and AC inferences.

Method

Participants. A total of 70 undergraduates (40 women and 30 men) with a mean age of 19.3 years (SD = 2.0) from The Ohio State University participated as part of an introductory psychology course. All were fluent English speakers. The sample also included 24 sixth-graders (13 boys and 11 girls) with a mean age of 11.7 years (SD = .45). All children attended a middle school in a suburb of Columbus, Ohio, and participated only if their parents had signed and returned a consent form. These children were primarily Caucasians. Adults were tested in an on-campus lab, and children were tested in a quiet classroom at their school.

Materials. The critical stories were 20 stories, minus probe words, selected at random from those used in Experiment 1; recall that each story was written in four versions (MP-Inference, MP-No Inference, AC-Inference, and AC-No Inference). The most frequent attribute for these 20 nouns were listed by a mean of 65.9% of participants (SD = 16.4%) Again, refer to the Appendix for example stories. Participants first
answered a yes/no comprehension question after each story (used to encourage participants to read attentively rather than skimming). Then, they received three test sentences, one of each of the following types:

1) *Inference* test sentences represented either MP or AC inferences (depending on condition); the contents of these inferences were not actually presented.

2) *Paraphrase* test sentences presented the concepts of one of the sentences from the story with wording altered from the original.

3) *Changed* test sentences presented the concepts from one of the story sentences, but one noun phrase was changed.

The important test sentences were the Inference sentences; the Paraphrase and Changed sentences were added because they should have elicited *Yes* and *No* responses, respectively, and prevented the formation of a simple response bias.

The 20 critical stories were split randomly into four groups of 5. Each participant was also assigned randomly to one of four groups. For college students, from 16 to 19 participants appeared in each group; for sixth-graders, *n = 6* per group. Participant groups and story groups were combined in a Latin square design, so that each participant responded to one version of each story. Each participant received 5 stories in each of the four experimental conditions, and never received a given story more than once. Each participant also responded to 15 non-conditional filler stories that had been selected from the non-conditional fillers used in Experiment 1. Each participant thus received 35 stories total. Following each filler, participants also received three test sentences—a Paraphrase,
a Changed sentence, and an Inference sentence. These Inference sentences, however, represented non-logical elaborative and bridging inferences that readers might have made, rather than inferences consistent with propositional logic.

**Design and procedure.** The experimental design included the within-subject factors of Argument Form (MP, AC) and Version (Inference, No Inference). Stimulus presentation and data collection were controlled by a PC that ran Superlab Pro 2.0 for Windows (Cedrus Corp., 1997).

Participants were tested individually. Each was told that she would read a series of stories on the computer, and that after each story, she would answer a question and then decide, for each of three test sentences, whether the information in each test sentence had been presented in the story. An example story was presented for the participant to read, and the participant then responded to two test sentences about this story, one Paraphrase and one Changed. The participant then received feedback about these test sentences; the feedback demonstrated that she should answer based on how well the information in each test sentence matched the content of the original story, rather than how well the test sentences matched the original stories' exact wording. The participant was not told to reason, and reasoning was never mentioned.

Each story was presented in a sentence-by-sentence process identical to that of Experiment 1. This process repeated until the fourth sentence; once the space bar was pressed, the comprehension question appeared. The participant responded to the question by pressing the “Z” key for yes and the “/” key for no. After this, each test sentence was presented one at a time, in random order. The participant pressed “Z” for yes if she believed that the information in the sentence had been presented in the story, and “/” for
no if she believed that the information in the sentence had not been presented. After the participant responded to all test sentences, the screen read PRESS SPACE BAR TO CONTINUE; when the participant hit the space bar, the entire procedure repeated for the next story. The order of story presentation was randomized for each participant by the program. The entire experiment lasted approximately 20-25 minutes.

Results and discussion

Analyses are presented separately for sixth-graders and adults, for multiple reasons. Data for sixth-graders were collected at a different time period than those for adults, and the settings for the two age groups (a classroom versus a lab room for adults) were different. Additionally, sample sizes for the two age groups were quite different, leading to greater overall variability in sixth-graders’ data than in the adult data. The latter considerations pose great problems for analysis of variance of these data, if age groups are analyzed separately (Keppel, 1991). However, after both age groups are discussed, age differences are inspected indirectly via a comparison of standardized differences between mean acceptance rates, a common effect size comparison (Cohen, 1988). Also, critical theoretical predictions, which pertain to how different conditional forms are processed, involve comparisons between observed rates of MP and AC inferences within each age group. Thus, although age comparisons could also be informative (for example, by comparing overall inference rates), the central predictions to be tested can be addressed by analyzing each age group separately.

Adults. Initial analyses of the comprehension question indicate that participants read the stories attentively, in that the mean proportion of errors was quite low ($M = 7.9\%, SD = 5.2\%$). Also, participants generally correctly rejected the Changed test
<table>
<thead>
<tr>
<th>Argument Form</th>
<th>Inference</th>
<th>No-Inference</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>73.1 (3.8)</td>
<td>19.1 (2.6)</td>
<td>46.1 (2.7)</td>
</tr>
<tr>
<td>AC</td>
<td>74.6 (3.5)</td>
<td>20.9 (2.5)</td>
<td>47.7 (2.5)</td>
</tr>
<tr>
<td>Column Means</td>
<td>73.9 (3.3)</td>
<td>20.0 (2.3)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Mean acceptance rates in percentages (SE’s in parentheses) for Inference test sentences by experimental condition, Undergraduates, Experiment 2.

sentences ($M$ errors = 6.2%, $SD = 6.1\%$) and correctly accepted the Paraphrase test sentences ($M$ errors = 10.4%, $SD = 7.2\%$). These participants seem to have taken the task seriously.

Mean acceptance rates for Inference sentences in each experimental condition are presented in Table 4.1. A pair of 2 X 2 repeated measures ANOVAs, with Argument Form and Version as within-subject factors, found a main effect of Version, $F_{1}(1, 69) = 348, p < .0001$; and $F_{2}(1, 19) = 281, p < .0001$. In neither analysis was the Argument Form main effect or the Argument Form X Version interaction significant, $F_{1}(1, 69) < 1, ps > .4$; and $F_{2}(1, 19) < 1, ps > .44$. Additionally, mean acceptance rates for MP- and AC-Inference conditions were above chance, both $t_{s}(69) > 6, ps < .0001$, one-tailed. Mean acceptance rates for MP- and AC-No Inference conditions were below chance, $ts(69) < -11, ps < .0001$, one-tailed.
It should be noted that the dependent variable in the item \( F_2 \) analysis was the proportion of acceptances for each item in each condition; proportions were necessary because each item did not appear in each condition (i.e., in each participant group) an equal number of times. Because proportions as a dependent variable may violate the homogeneous variance assumption of general linear models (Cohen & Cohen, 1983), the item proportions were transformed via Cohen and Cohen’s (1983) recommended arcsine transformation (i.e., \( DV = 2 \times \text{arcsin}((\text{proportion})^{1/2}) \)) and the item analysis was repeated on the transformed dependent variable. However, this analysis repeated the original \( F_2 \) results: a main effect emerged for version, \( F_2(1, 19) = 185, p < .0001 \). The Argument Form and Argument Form X Version effects were both non-significant, both \( F_2s(1, 19) < 1 \).

**Sixth-graders.** These participants also performed well on the comprehension questions (\( M \) errors = 11.5\%, \( SD = 6.3\% \)), as well as the Changed test sentences (\( M \) errors = 7.6\%, \( SD = 4.7\% \)) and the Paraphrase test sentences (\( M \) errors = 12.7\%, \( SD = 6.5\% \)).

Mean acceptance rates for Inference sentences in each condition are presented in Table 4.2. Repeated measures ANOVAs with the same design as that used to analyze the adult data found only a main effect for Version, \( F_1(1, 23) = 61.4, p < .001, F_2(1, 19) = 65.6, p < .001 \). In neither analysis was the Argument Form effect significant, \( F_1(1, 23) < 1, F_2(1, 19) = 1.01, p > .3 \). The Argument Form X Version interaction was also not
Table 4.2: Mean acceptance rates in percentages (SE’s in parentheses) for Inference test sentences by experimental condition, Sixth-graders, Experiment 2.

<table>
<thead>
<tr>
<th>Argument Form</th>
<th>Inference</th>
<th>No-Inference</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>69.2 (6.8)</td>
<td>31.7 (5.9)</td>
<td>50.4 (5.0)</td>
</tr>
<tr>
<td>AC</td>
<td>75.0 (5.4)</td>
<td>33.0 (5.6)</td>
<td>54.2 (5.1)</td>
</tr>
<tr>
<td>Column Means</td>
<td>72.1 (5.4)</td>
<td>32.5 (4.9)</td>
<td></td>
</tr>
</tbody>
</table>

significant in either analysis, with both $F_1(1, 23)$ and $F_2(1, 19) < 1$. Also like the adult data, participants’ mean acceptance rates in the MP- and AC-Inference conditions were above chance, respectively, $t(23) = 2.82, p < .005$, one-tailed; and $t(23) = 4.62, p < .0001$, one-tailed. Participants’ mean acceptance rates in the MP- and AC-No Inference conditions were below chance, respectively, $t(23) = -3.1, p < .005$, one-tailed; and $t(23) = -2.97, p < .005$, one-tailed. (Because each item appeared an equal number of time in each conditions for the sixth-graders, the dependent variable in the item analysis was not a proportion, and no transformation was needed.)

Finally, to examine any age differences in the tendency to draw inferences, difference scores representing the difference between each participants’ number of acceptances in the Inference conditions (collapsed across argument form) and number of acceptances in the No-Inference conditions (again collapsed over form) were computed.
Then, the means of these difference scores for the adults and sixth-graders were compared by estimating the standardized difference between these means. The measure used to make this estimate was Cohen's $d$ (Cohen, 1988), which is the difference in group means divided by an estimate of the pooled groups' standard deviation. The resulting score is an effect-size measure in standard-deviation units.

To aid in understanding this comparison, mean acceptance rates in the Inference and No-Inference conditions, divided by age group, are presented in Figure 4.1. For the difference scores computed above, Cohen's $d$ for the difference between adults and sixth-graders is 0.59. This measure indicates that the mean difference in acceptance rates for Inference test sentences in the Inference and No-Inference conditions was indeed smaller for sixth-graders than adults. In particular, as Figure 4.1 shows, sixth-graders were somewhat more likely than adults to respond yes in the No-Inference conditions. However, because both age groups still accepted reliably more Inference test sentences following both MP- and AC-Inference stories, this age difference does not seem to be related to the tendency to draw either type of inference.

Overall, when participants of both age groups were presented with sentences representing inferences after reading an MP- or AC-Inference story, they reliably responded that the information in the inferences had appeared in the stories. When the same sentences were presented after No-Inference stories, participants reliably rejected them. However, this result appears to be stronger for adults than for sixth-graders. Nevertheless, these results, combined with those of Experiment 1, suggest that participants draw MP and AC inferences during comprehension.
Figure 4.1: Mean acceptance rates (+SEs) for Inference test sentences by story version (collapsed over argument form) and age group, Experiment 2.
Sixth-graders tended to accept more of the Inference test sentences in the No-Inference conditions than did adults. Preadolescents sometimes rely on verbatim memory more often than do adults in comprehension tasks of this sort, because they do not distinguish as well between verbatim and gist memory traces (Brainerd & Reyna, 1993). Because the inference test sentences literally appeared in the story (as part of the conditionals), sixth-graders may have tended to accept Inference test sentences at higher rates overall; this is just what would be expected if these participants tended to rely a bit more on verbatim memory. However, the data pattern for sixth-graders across experimental conditions is quite similar to the adult results, with Inference conditions leading to many more acceptances than the No-Inference conditions with both MP and AC stories. Thus, the results with sixth-graders cannot be explained as merely resulting from a greater reliance on verbatim memory.

These results converge with those of Experiment 1. The critical experiments, however, will involve conditional premises formed from weakly associated propositions. First, the use of this type of content can help determine which version of mental models theory is most consistent with the conditional inferences that people draw during comprehension. The standard theory suggests that the strength of association between premises will not matter, because people should still draw both MP and AC inferences. The modified mental models theory suggests that people should draw MP inferences with weakly associated propositions, but not AC inferences.

A second reason for doing these critical experiments is to address potential criticisms of the stimuli in Experiments 1 and 2. The conditional premises in these experiments involved strongly associated nouns and attributes of the things that the nouns
denoted. The priming results found in Experiment 1 could have reflected nothing more than the transient activation given to one concept by the other (e.g., night and dark are strongly linked, so the words may have primed one another regardless of the propositional form of the premises). Likewise, because each pair was strongly associated, the memory results of Experiment 2 could have been due to plausibility judgments that were due to nothing more than concept activation, regardless of logical syntax. These explanations do not seem very likely, given that the critical Inference and No-Inference sentences in each story contained the same key words that could have caused such transient activation. Nevertheless, replicating those experiments with stimuli that are not strongly linked in memory would help resolve this concern.

Experiment 3 is therefore a replication of the online priming study (Experiment 1), but with weakly associated antecedents and consequents used in place of the strongly associated ones. This experiment was conducted with adults only. Experiment 4 repeats the recognition task of Experiment 2, again with weakly associated stimuli used in the place of strongly associated ones.
CHAPTER 5

EXPERIMENT 3

This experiment is a replication of Experiment 1. The design, procedure, and materials are exactly the same. The only change is in the stimuli: instead of the strongly associated attributes used in Experiment 1, weakly associated attributes are used. To recap relevant theoretical predictions, syntactic theories suggest that MP inferences, largely irrespective of content, should occur during comprehension whenever discourse fits the MP schema (Braine & O’Brien, 1991; Lea, 1995). AC inferences would be difficult to predict according to the syntactic account. Standard mental models theory indicates that both MP and AC inferences should be drawn, at least if it is assumed that searching for counterexamples does not occur during comprehension (Johnson-Laird & Byrne, 1991). The modified mental models theory instead suggests that only MP inferences should be drawn. Participants should be less likely to draw AC inferences because conditional premises with weakly associated propositions lead participants to represent alternative antecedents, which block AC inferences (Markovits, 1993; Quinn & Markovits, 1998).
Method

Participants. A total of 56 introductory psychology undergraduates from The Ohio State University (Mean age = 19.6 years, SD = 2.0), including 33 women and 23 men, took part voluntarily in partial fulfillment of a course requirement. All participants were fluent English speakers.

Materials. The same 40 critical stories used in Experiment 1 were also used here. As in Experiment 1, each story was written in four versions (again, see Appendix for examples). However, the stories were modified so that the attribute used in each conditional premise was the infrequent attribute selected for that noun (as identified in the Preliminary Experiment). To accommodate the infrequent attributes, eight of the 40 critical stories had to be slightly modified to make them sensible, but the remaining 32 were left intact. As in Experiment 1, the 40 stories were randomly split into four groups of 10, and combined with four groups of participants (n = 14 per group) in a Latin square design to ensure that each participant only received one version of each story. The 72 fillers used in Experiment 1 (40 conditional fillers that required “no” answers and 32 non-conditional fillers) were also used in this experiment, as were the 12 practice trials that were not counted.

Design and procedure. These were identical to those of Experiment 1.

Results and discussion

In analyses of reaction times, wrong answers were discarded. Then, for each participant, reaction times that were at least 2.5 standard deviations above or below the participant’s grand mean were discarded, a procedure that removed 2.7% of all responses.
Analysis of participants' answers to the comprehension questions after each story suggested that participants were reading the stories attentively; the mean percentage of errors for these questions was 11.6%, $SD = 4.4%$.

Participants' mean reaction times in each condition, along with error rates, appear in Table 5.1. Mean reaction times were analyzed in repeated measures ANOVAs with Argument Form and Version as within-subject factors. In the subject analysis, there was a significant effect of Version, $F_1(1, 55) = 29.99, p < .001$, whereas the Argument Form effect and Argument Form X Version interaction were both nonsignificant, both $F_1s(1, 55) < 1$. This pattern was repeated in the item analysis, which found a significant effect for Version, $F_2(1, 39) = 13.71, p < .001$, but not for Argument Form or for the Argument Form X Version interaction, both $F_2s(1,39) < 1$.

Finally, to look for the presence of any speed-accuracy tradeoffs that would compromise interpretation of the data, a 2 X 2 repeated measures ANOVA, with the same factors as before, was conducted on the numbers of errors that participants committed in each condition. Neither the main effects nor the interactions were significant in this analysis, all $F$s(1, 55) < 0.5. This concern thus seems to be an unimportant issue with these data.

Thus, participants seem to have drawn both MP and AC inferences with these stimuli as well. Unlike Experiment 1, this study did not find faster overall times for MP stories, so the conjecture about a preference for "forward" conditional inferences must remain conjecture. Results of this experiment have several implications. First, the evidence for AC inferences along with MP inferences suggests that participants are not applying logically valid inference schemas to input, as the syntactic position suggests
Table 5.1: Mean reaction times in milliseconds (SE’s in parentheses), and mean error rates in percentages, by experimental condition, Experiment 3.

(Braine & O’Brien, 1991; Lea, 1995; Rips, 1994). This finding also suggests that the cognitive-developmental position, that adults represent conditionals in veridical fashion, needs to be specified more clearly. Such veridical representation does not seem to occur during comprehension, but cognitive-developmental theory does not distinguish between inferences during comprehension and deliberate reasoning.

Finally, these results are consistent with the standard mental models position that participants draw MP and AC inferences during comprehension without searching for counterexamples to conclusions. On the other hand, the modified mental models theory, which has been developed to explain many phenomena associated with conditional reasoning, was not supported. That theory predicts that conditionals formed from weakly associated antecedents and consequents should not promote AC inferences because alternative antecedents are retrieved automatically and thus block AC inferences (Markovits, 1993).
A larger implication of these findings is that conditional reasoning needs to be distinguished from conditional inference during comprehension. This speculation runs counter to the claim that logical reasoning is largely a subset of comprehension (Braine, 1978; Braine & O'Brien, 1991), and it is also inconsistent with the position that counterexamples to inferences may be available in comprehension (e.g., Quinn & Markovits, 1998). However, some researchers argue that true logical reasoning depends on the application of conscious strategies, whether search for counterexamples or something else (Evans, 2000; Johnson-Laird, et al., 1999). It remains to replicate the recognition memory task of Experiment 2 with the contents used in the present experiment, both to look for converging evidence and to obtain developmental data.
CHAPTER 6

EXPERIMENT 4

This experiment is a replication of the recognition memory procedure in Experiment 2. The conditional premises in this experiment were formed from nouns and infrequently generated attributes of them found in the preliminary study; thus, the same noun-attribute pairs used in Experiment 3 and twenty others were used in this experiment. Adult and preadolescent participants read stories conforming to MP and AC inference forms (without conclusions) and then decided whether the information in certain test sentences had been presented in each story. The critical test sentences were those that represented either MP or AC inferences. Syntactic theories suggest that participants in both age groups should be more likely to false alarm on sentences representing MP inferences (when those sentences follow MP-Inference stories) than on sentences representing AC inferences. Standard mental models theory predicts that all participants should instead false alarm on test sentences representing MP and AC inferences when those sentences follow Inference stories. Modified mental models theory instead predicts that participants should not false alarm on AC inferences, because participants do not
draw them, but MP inferences should be accepted. Finally, cognitive developmental theory implies that adults might false alarm on only MP inferences, whereas preadolescents might tend to false alarm on MP and AC inferences.

Method

Participants. The participants for this experiment included 48 introductory psychology undergraduates (19 females and 29 males) with a mean age of 20.5 years ($SD = 3.1$). These participants took part voluntarily in partial fulfillment of a course requirement. The sample also included 19 preadolescents (9 male and 10 female) with a mean age of 11.2 years ($SD = .37$). These participants were primarily Caucasians, and they took part only if their parents signed and returned a consent form. Undergraduates were tested in an on-campus lab, and preadolescents were tested in a quiet room at their school.

Materials. The same 20 critical stories and test sentences used in Experiment 2 were used. The infrequent attributes involved were listed by a mean of 11.4% of participants ($SD = 2.1$). As in Experiment 2, each story was written in four versions (refer to the Appendix for examples). The stories were changed so that the attribute used in each conditional premise was the infrequent attribute selected for that noun (as identified in the Preliminary Experiment). To accommodate the infrequent attributes and keep the stories sensible, two of the 20 critical stories had to be slightly modified, but the remaining eighteen were left intact. The 20 stories were randomly split into four groups of 5, and combined with four groups of participants in a Latin square design to ensure that each participant only received one version of each story. For college students,
$n = 12$ for each participant group; for sixth-graders, $n = 4$ or $5$ per participant group. The 15 non-conditional fillers and their test sentences from Experiment 2 were also used in this study.

**Design and procedure.** The design and procedure were exactly the same as in Experiment 2.

**Results and discussion**

The plan of data analysis is the same as that in Experiment 2: age groups are analyzed separately, and age effects are examined indirectly via effect size estimates.

**Undergraduates.** Inspection of the answers to the comprehension questions indicated that these participants appeared to be reading the stories attentively ($M$ errors = 11.2%, $SD = 7.5$%). Participants also tended correctly to reject the changed test sentences ($M$ errors = 6.2%, $SD = 6.5$%) and to accept the paraphrase test sentences ($M$ errors = 10%, $SD = 6.3$%).

Participants' numbers of acceptances of Inference test sentences in each experimental condition were analyzed in 2 X 2 repeated measures ANOVAs with Argument Form and Version as within-subject factors (see Table 6.1). These analyses revealed a Version effect that was significant by subjects, $F_1(1, 47) = 67.9, p < .001$, and by items, $F_2(1, 19) = 199.9, p < .0001$. The main effect of Argument Form and the Argument Form X Version interaction were not significant in either analysis, all $Fs < 1$, $ps > .39$. (No transformation was needed in the item analysis because each item appeared an equal number of times in each condition.) The mean acceptance rates of Inference test sentences were also compared to chance rates. These analyses revealed that the mean acceptance rate of Inference sentences in the MP-Inference condition was above chance,
<table>
<thead>
<tr>
<th>Argument Form</th>
<th>Inference</th>
<th>No Inference</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>60.8 (9.1)</td>
<td>23.8 (4.2)</td>
<td>42.3 (3.9)</td>
</tr>
<tr>
<td>AC</td>
<td>59.2 (4.9)</td>
<td>20.8 (3.7)</td>
<td>40.0 (3.5)</td>
</tr>
<tr>
<td>Column Means</td>
<td>60.0 (4.6)</td>
<td>22.3 (3.6)</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: Mean acceptance rates in percentages (SE’s in parentheses) for Inference test sentences by experimental condition, Undergraduates, Experiment 4.

$t(47) = 2.13, p < .05$, one-tailed; whereas the mean acceptance rate of Inference sentences in the MP-No Inference condition was below chance, $t(47) = -6.19, p < .0001$, one-tailed.

The mean acceptance rate for Inference sentences following AC-Inference stories was above chance, $t(47) = 1.88, p < .05$, one-tailed; and the mean acceptance rate following AC-No Inference stories was below chance, $t(47) = -7.95, p < .0001$, one-tailed.

**Sixth-graders.** Inspection of answers to the comprehension questions revealed that these participants made generally made few errors ($M$ errors = 13.8%, $SD$ = 6.7%), so they seemed to have read the stories attentively. Error rates with Paraphrase test sentences ($M = 15.5\%, SD = 7.9\%$) and Changed test sentences ($M$ errors = 8.7%, $SD = 6.4\%$) were also low overall.
Participants' mean acceptance rates for Inference sentences across experimental conditions are presented in Table 6.2. As the table suggests, 2 X 2 repeated measures ANOVAs (with Argument Form and Version as factors) found a significant main effect for Version, $F_1(1, 18) = 32.06, p < .001; F_2(1, 19) = 64.20, p < .001$. The Argument Form effect missed significance in both analyses, $F_1(1, 18) = 1.99, p = .18; F_2(1, 19) = 1.47, p = .24$. The Argument Form X Story Version interaction was also not significant, both $F_1(1, 18)$ and $F_2(1, 19) < 1$. When the proportions used in the item analysis were transformed via the arcsine transformation (Cohen & Cohen, 1983), the result pattern persisted. The main effect for Version was still significant, $F_2(1, 19) = 49.78, p < .0001$, but Argument Form effect was not significant, $F_2(1, 19) = 1.33, p = .26$; nor was the Argument Form X Version interaction significant, $F_2(1, 19) < 1$. 

Table 6.2: Mean acceptance rates in percentages (SE’s in parentheses) for Inference test sentences by experimental condition, Sixth-graders, Experiment 4.

<table>
<thead>
<tr>
<th>Version</th>
<th>Argument Form</th>
<th>Inference</th>
<th>No Inference</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td>58.9 (6.9)</td>
<td>25.3 (6.3)</td>
<td>42.1 (5.4)</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>67.4 (6.7)</td>
<td>29.5 (5.6)</td>
<td>48.4 (5.0)</td>
<td></td>
</tr>
<tr>
<td>Column Means</td>
<td>63.2 (6.2)</td>
<td>27.4 (5.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acceptance rates for the Inference test sentences in each condition were also compared to chance. As with adults, both MP- and AC-No Inference conditions led to acceptance rates that were significantly below chance; respectively, $t(18) = -3.94, p < .001$, one-tailed, and $t(18) = -3.67, p < .001$, one-tailed. However, whereas acceptance rates in the AC-Inference condition were above chance, $t(18) = 2.59, p < .001$, one-tailed; acceptance rates in the MP-Inference condition could not be distinguished from chance, $t(18) = 1.29, p = .11$, one-tailed. This last result is still in the predicted direction, however, and the acceptance rate for Inference test sentences is very similar to that of the adults in the MP-Inference condition. A lack of statistical power may explain the inability to distinguish this result from chance.

As with Experiment 2, a difference score was computed for each participant to measure the difference between number of acceptances of Inference test sentences in the Inference conditions and No-Inference conditions, collapsing across argument form. Then Cohen’s $d$ was used to estimate differences, if any, between adults’ and sixth-graders’ mean difference scores. To aid in understanding this comparison, mean acceptance rates in the Inference and No-Inference conditions, as a function of age group, are presented in Figure 6.1. Cohen’s $d$ for the difference in adults’ and sixth-graders’ mean difference scores was 0.064. This result indicates that sixth-graders may have had a slight tendency to answer yes more so than adults, with the difference more pronounced in the No-Inference conditions. However, unlike Experiment 2, the estimated difference between age groups was quite small (less than 1/10 of a standard deviation). It is also
Figure 6.1: Mean acceptance rates (+SEs) for Inference test sentences by story version (collapsed over argument form) and age group, Experiment 4
important to point out that sixth-graders still accepted Inference test sentences in the Inference conditions at reliably higher rates than in the No-Inference conditions, just as the adults did.

Results with undergraduates and sixth-graders thus largely replicate the pattern found in Experiment 2. Whenever an Inference test sentence followed either an MP or AC argument, of which the Inference sentence was the conclusion, participants tended to agree that they had read the information in the Inference sentence. Whenever an Inference sentence was presented after a story that did not fit either form, participants tended to say that they had not read the information in it.

The acceptance rates of Inference test sentences in MP- and AC-Inference conditions were lower than those of Experiment 2. In that experiment, participants accepted Inference test sentences in the MP- and AC-Inference conditions roughly 70-76% of the time, on average. The corresponding acceptance rates in this experiment ranged from approximately 59% to 67%. It is possible that this decrease is simply a plausibility effect: strongly associated attributes like those of Experiment 2 may seem more likely in the story context, and participants are therefore more likely to agree that the stories actually contained inferences mentioning those attributes. Whether or not this is the case, the overall pattern in acceptance rates found in this experiment is very similar to that found in Experiment 2.

These results are again consistent with the standard mental models theory predictions that MP and AC inferences occur in comprehension. The results also make sense given what is known about human discourse processing in general (this idea will be elaborated in the General Discussion). Experiment 4’s results are not compatible with the
predictions of syntactic or cognitive-developmental theories, and they are also hard to reconcile with modified mental models theory. Across all 4 experiments, the results suggest that a factor known to impact deliberate conditional reasoning — the semantic memory association between antecedent and consequent — is not important in determining the conditional inferences drawn during comprehension. In turn, reasoning and comprehension seem to require clear conceptual separation.
CHAPTER 7

GENERAL DISCUSSION

This chapter will focus on several issues. First, the results from the four experiments will be overviewed with respect to critical theoretical predictions. Next, the implications of these data for theories of conditional processing, both in terms of deliberate reasoning tasks and otherwise, will be discussed. Finally, I will discuss the larger relationship between comprehension and reasoning.

Overview of findings

In Experiment 1, adult participants read short stories that conformed to the logically valid MP argument form (If $P$ then $Q; P; therefore Q$) or the logically invalid AC argument form (If $P$ then $Q; Q; therefore P$). Argument conclusions were never presented. The antecedent and consequent in each conditional seemed to be strongly linked in semantic memory, based on results of the Preliminary Survey. After each story, participants responded to a probe word denoting the main concept in the argument’s conclusion. Participants’ mean response times to probes following both MP- and AC-Inference stories were significantly faster than responses to those same probes when they followed the corresponding No-Inference control stories. This decrease in reaction time
compared to controls suggests that both MP and AC stories prime their conclusions. The MP condition was faster overall than the AC condition, although reasons for this finding are unclear. The lattermost finding did not replicate in Experiment 3.

In Experiment 2, adult and sixth-grade participants read the same sorts of stories and had to decide whether the information in test sentences representing the arguments’ conclusions had been presented in the stories. When the stories conformed to a conditional argument form (MP or AC), participants reliably agreed that the information in those conclusions had been presented. However, when the same test sentences were presented after No-Inference control stories, participants reliably rejected them. This data pattern occurred for both sixth-graders and undergraduates, although sixth-graders were somewhat more liberal in their “yes” responses than undergraduates. The two experiments together thus suggest that participants will draw MP and AC inferences during text comprehension, with argument validity apparently an unimportant factor.

Experiments 3 and 4 replicated the first two experiments, but with a different type of content. Each conditional’s antecedent and consequent were weakly associated. This variation is known to decrease the frequency with which participants draw AC inferences in deliberate reasoning tasks, while leaving the tendency to draw MP inferences unaffected (Cummins, 1995; Janveau-Brennan & Markovits, 1999; Quinn & Markovits, 1998; Schroyens, et al., 2000). As in Experiment 1, however, both MP- and AC-Inference stories of Experiment 3 seemed to prime their conclusions, as compared to MP- and AC-No Inference stories. Reaction times in Experiment 3 were overall slightly slower than those of Experiment 1, but this is likely to reflect a plausibility/context effect rather than an influence of the stories’ propositional forms. Experiment 4 suggested that
adult and sixth-grade participants encode MP and AC inferences into their memory representations of the stories even when each conditional’s constituent propositions were not strongly associated. This finding was more tenuous in Experiment 4 than in Experiment 2. The difference may again reflect a plausibility effect; i.e., the conclusions to arguments based on weakly associated conditionals were not perceived to be quite as plausible as those based on stories containing strongly associated conditionals.

In the recognition tasks of Experiments 2 and 4, which involved sixth-graders and adults, the response patterns were essentially the same across age groups. Participants seemed to encode both MP and AC inferences. The only difference between sixth-graders and adults was that sixth-graders were slightly more likely to false alarm on inference test sentences when the sentences were presented following No-Inference control stories (this tendency was more pronounced in Experiment 2 than in Experiment 4). Possibly, task instructions (i.e., respond to test sentences based on whether the information in them had appeared in the story) caused sixth-graders to rely slightly more on verbatim memory for the stories. Such a reliance would lead these participants to respond yes to sentences representing inferences because those inference sentences were literally presented in the stories (i.e., the words in each inference constituted part of the conditional that appeared in the story).

Indeed, children and preadolescents can rely on verbatim memory traces more so than adults in some recognition tasks (Brainerd & Reyna, 1993). Participants in this age group also may not exploit information available in memory to make inferences with quite the same frequency as adults (Barnes, et al., 1996). Therefore, reliance on verbatim memory may be a fallback for preadolescents when they are not certain whether or not
some piece of information was explicitly stated in a text. However, even if this is true, the difference between Inference and No-Inference conditions was quite pronounced even for the sixth-graders. Those participants also seemed to draw MP and AC inferences given the appropriate input.

Importantly, the priming experiments (1 and 3) and the recognition experiments (2 and 4) yielded results that in general converged on the same conclusion. Convergence is important because no single method is regarded as an unambiguous indicator of inferences in comprehension (e.g., Kintsch, 1998; Trabasso & Suh, 1993). Recognition judgments about sentences presented after a story may reflect only a general plausibility assessment, for example (Alba & Hasher, 1983). The general strategy in the reported work has therefore been to look for converging evidence from multiple methods. The existence of this evidence lends much support to the conclusions offered here. Of course, conclusions can never be certain; further converging evidence could be obtained by, for example, measuring latencies to pronounce the probe words used in the priming tasks of Experiments 1 and 3 (see Potts, et al., 1988).

As readers have no doubt realized, the conditionals used in the reported studies were embedded in story contexts that were much richer than the typical conditional reasoning problem. However, because theories of conditional processing are often intended to apply to such contexts (Braine & O’Brien, 1991; Johnson-Laird, 1983; Lea, 1995), the use of this type of content seems justified. Additionally, extending such theories to discourse contexts raises the issue of the relationship between comprehension and reasoning, a subject of great importance to the present work.
Implications for theories of conditional processing

For syntactic theories, the repeated finding of both MP and AC inferences during comprehension is problematic. Syntactic theories propose that people automatically extract the propositional form of conditionals and apply the MP inference schema to these forms, whenever premises are available in working memory (Braine, 1990; Braine & O'Brien, 1991; Lea, 1995; Lea, et al., 1990). Because the MP and all other inference schemas postulated by syntactic theories are valid, the finding of AC inferences in comprehension suggests theoretical incompleteness. Furthermore, this result suggests that people may not automatically parse discourse into logically valid syntax; people may attempt (with variable success) such parsing only when asked to reason. Indeed, some researchers argue that logical reasoning performance depends upon the deliberate application of one or more strategies, which may vary across individuals, rather than automatic comprehension processes (Evans, 2000; Johnson-Laird, Savary, & Bucciarelli, 2000; Stanovich, 1999).

One syntactic response to the existence of AC inferences could be that conditionals may be mis-represented as biconditionals (e.g., Braine, 1978; Rumain, et al., 1983). The biconditional-interpretation hypothesis suggests that whenever presented with an If $P$ then $Q$ conditional, participants maintain a representation of both the conditional and its converse If $P$ then $Q$ in working memory. Thus, someone hearing (a) If $P$ then $Q$ assumes that (b) If $Q$ then $P$ is also the case. In this case $Q$, while a consequent in (a), is an antecedent in (b). Therefore, according to this interpretation, people make an MP inference in (b), which appears as if it were an AC inference in (a). However, no data exist that support this speculation. While the proportions of false
alarms evidencing AC inferences in Experiments 2 and 4 were above 70% and 59%, respectively, other research suggests that participants rarely represent a conditional’s converse along with the conditional itself (Rader & Sloutsky, 2000b). Furthermore, syntactic theories do not include biconditional inference schemas, suggesting that this argument form should not be processed directly in comprehension even if participants do represent conditionals as biconditionals. Therefore, the biconditional interpretation cannot easily account for the occurrence of AC inferences.

A second syntactic explanation for AC inferences is that people draw invited inferences; that is, someone hearing (a) If \( P \) then \( Q \) is led to believe that (b) If not-\( P \) then not-\( Q \) is also true (Braine & O’Brien, 1991; Geis & Zwicky, 1971). As has been noted, no evidence yet indicates that such inferences occur in comprehension. However, assume that they do. AC inferences could occur in the following manner. When presented with \( Q \), which is the negation of the consequent in (b), one concludes that \( P \) is also the case, which is the negation of the antecedent in (b). Such an inference would conform to a valid modus tollens (MT) inference, specifically of the form If not-\( P \) then not-\( Q \); \( Q \); therefore, \( P \). [This presentation of the MT rule may seem unusual, but the basic MT syntax is If \( P \) then \( Q \); not-\( Q \); therefore not-\( P \). In other words, the consequent of the conditional is negated, and the antecedent is negated as a conclusion. The example here follows this syntax – the negation of not-\( Q \) is a substitution instance of \( Q \), and the negation of not-\( P \) is a substitution instance of \( P \).]

However, this putative explanation suffers two drawbacks. First, no syntactic theory posits an MT schema; in fact they typically argue that MT inferences must be drawn through conscious application of a reductio ad absurdum strategy (Braine, et al.,
1984; Rips, 1994). Secondly, numerous reviews concur that reasoners typically draw MT inferences much less frequently than MP inferences (Evans, et al., 1993; Johnson-Laird & Byrne, 1991; Rips, 1995). Given that the evidence for MP and AC inferences appeared to be equally strong in my experiments, the invited inference hypothesis seems unable to account for AC inferences during comprehension. In sum, it appears unsatisfactory to explain away AC inferences in comprehension as merely the result of invited inferences or biconditional representations. Syntactic theories thus seem to require substantial modification in light of the present results.

Stein (1996), in discussing the rationality debate that has occupied many cognitive scientists, philosophers, and psychologists over the past three decades, attacks an implicit assumption seemingly held by many syntactic proponents (such as Braine, 1978, 1990; Cohen, 1981; Macnamara, 1986; Rips, 1994). This assumption is that a meaningful analogy exists between linguistic competence and performance, on the one hand, and reasoning competence and performance, on the other hand. Linguists have long held that linguistic competence is not necessarily impugned by errors in performance. For example, someone who utters a tongue-twister is still capable of forming and uttering the correct statement; the error reflects perhaps trouble in sequencing the movements needed to form the words, phonological similarity between elements that are misplaced, etc. In general, we do not decide that such performance errors signify a lack of some element of linguistic competence. The ready analogy to reasoning is that errors in performance do not necessarily reflect errors in competence — someone who, for example, draws the AC inference may still possess the correct understanding of the conditional but commit the error due to a slip of attention, a pragmatically-motivated inference, etc. Indeed, all of
the authors cited above make just such appeals, and the more radical in the group argue that human reasoning competence must therefore be normatively correct as a matter of principle (Cohen, 1981; Macnamara, 1986).

However, Stein (1996) argues that linguistic competence and performance are linked in an important way that reasoning competence and performance are not. Describing the linguistic competence of a person (i.e., the set of rules/knowledge used to form and understand language) also describes the norms of grammaticality for that person; if his/her competence were to change, then what normatively could be a grammatical statement would also be different. On the other hand, suppose someone possesses a normatively correct reasoning competence. Something happens to alter that person’s competence, so that she no longer understands or applies the principle modus ponens. Would we then argue that that person’s new reasoning competence was normatively appropriate? Probably we would not; we would not be so ready to assume that failure to apply modus ponens was merely a performance error. Instead we might assume that, by failing to apply a principle that underlies many types of inferences, the principle was in fact lacking from that person’s competence, and that the person was making mistakes.

Furthermore, Stein (p. 230) suggests that even persons lacking a normative principle of reasoning might be able to learn it, perhaps by bootstrapping the principle onto an already-existing rule or heuristic that serves to guide correct inferences in a variety of situations. In this vein, it is noteworthy that students in majors that emphasize the abstraction of necessary and specific relations among phenomena, as well as a search for contradictions (e.g., law, psychology), show large improvement in the ability to
reason with conditionals (Nisbett, Fong, Lehman, & Cheng, 1987). On the other hand, developmental, neurological, and even some genetic evidence suggest that the competence underlying the ability to learn language is largely innate, suggesting that linguistic competence is not nearly so flexible. Stein notes that comparable evidence for the innateness of reasoning is lacking (although some evidence does exist concerning regions of the brain involved in different types of reasoning; e.g., Waltz, Knowlton, Holyoak, Boone, Mishkin, Santoz, Thomas, & Miller, 1999).

The issue will not be further belabored, but the point is that simply falling back on a competence-performance distinction to "explain away" erroneous inference patterns, as many syntactic theorists have, does not seem to be a defensible proposition. At the very least, compelling evidence that could support this proposition is currently lacking.

Semantic theories, in particular the standard mental models theory (Johnson-Laird & Byrne, 1991), predict that the propositional representation of a conditional formed during comprehension consists mainly of a conjunction of the antecedent and consequent. If further input supplies the premise needed for either an MP or AC inference, this inference is likely to be drawn during comprehension because both inferences can be based on the representations formed during comprehension. If participants do not engage in deliberate search for counterexamples to their inferences during comprehension, then MP and AC inferences seems especially likely. All four experiments yielded data supporting this prediction. The finding of AC inferences during comprehension is novel; only one author to my knowledge has previously reported such data (Franks, 1997) and
that was in a recognition memory task possessing methodological characteristics that made the task more similar to deliberate reasoning. No study has heretofore demonstrated the existence of AC inferences in an online priming situation.

The reported findings suggest that the modified mental models theory requires revision (Markovits, 1993). This account, which focuses on conditional reasoning, assumes that participants will retrieve alternative antecedents to conditionals with likelihoods inversely proportional to how strongly each conditional’s antecedent and consequent are linked in memory. Strongly associated propositions prevent the retrieval of alternative antecedents, but weakly associated propositions make such retrieval much more likely. Strongly associated propositions thus lead participants to endorse both MP and AC inferences at high rates, but weakly associated propositions lead participants to endorse only MP inferences, with AC inferences much less frequent (Cummins, 1995; Janveau-Brennan & Markovits, 1999; Klaczynski & Narasimham, 1998; Markovits, 1985; Markovits, et al., 1998; Quinn & Markovits, 1998; Schroyens, et al., 2000; Thompson, 1994). These enhancement/suppression effects associated with AC have been found with participants across a wide age range. The natural extrapolation from these findings is that parallel effects should be found in comprehension when content is varied in this manner. The prediction seems reasonable if the retrieval of alternatives is truly an online process (Quinn & Markovits, 1998).

However, no such differences were found in the present experiments. It is therefore reasonable to speculate that these alternatives are retrieved only when participants engage in deliberate reasoning (an idea discussed in more detail below when reasoning and comprehension are addressed). Directly presenting counterexamples may
lead to their use in deliberate conditional reasoning tasks (Byrne, 1989; Byrne, et al., 1999; Rumain, et al., 1983). Also, some of these studies have used conditional premises whose semantic and pragmatic content strongly suggests counterexamples to inferences (e.g., Markovits, et al., 1998). However, all of these studies involved explicit reasoning. Currently no evidence, including the present experiments, suggests that search for counterexamples or any other deliberate reasoning strategy comes into play when participants are not explicitly requested to reason.

Because MP and AC inferences in comprehension appear to be drawn based only on initial representations of premises, a tentative conclusion is that such inferences represent only an initial phase of reasoning and not the entire process. Indeed, in most conditional reasoning studies using semantically impoverished content, MP inferences are more numerous than AC inferences (Byrne & Tasso, 1999; Johnson-Laird, 1999; Rips, 1994; but see Evans, et al., 1995). At least some participants seem to apply a more deliberate strategy to these problems. The AC suppression and enhancement effects described earlier are robust phenomena as well. However, because MP and AC inferences in comprehension seem to be equally robust based on the reported work, it seems that these deliberate strategies, whatever they may be, are not brought into play. This conclusion is also consistent with the finding that, across a variety of other reasoning problems, most participants tend to draw inferences that are most easily available, even when they are normatively inappropriate (Evans, et al., 1999; Newstead, et al., 1999; Stanovich, 1999; see also Sloman, 1996).
With respect to developmental concerns, no developmental theory of conditional processing seems easily reconciled with the reported results. Cognitive-developmental theories argue for a shift from preadolescence to adulthood in how conditionals are represented (Byrnes & Overton, 1986; Overton, 1990; Ward & Overton, 1990; see also Franks, 1996). Preadolescents such as the sixth-graders in this study are said to treat conditionals as biconditionals, hence they should draw AC as well as MP inferences (Given the problems already discussed with the notion of a biconditional representation, this is questionable). Adults, in contrast, should at least be capable of representing conditionals veridically, representing all the possibilities with which a conditional is compatible. They should thus be less likely to commit AC inferences than younger participants.

Results do not support these developmental hypotheses because adults and preadolescents did not differ in their MP and AC inference rates. Also, the cognitive-developmental assumptions may be criticized on two grounds. First, the assumption that participants represent all possibilities compatible with a conditional, including cases with a false antecedent, seems to make unreasonable demands on working memory. A variety of reasoning studies, including not only studies of propositional reasoning but also studies of syllogistic reasoning and other reasoning with quantifiers (e.g., Johnson-Laird & Bara, 1984; Johnson-Laird, Byrne, & Tabossi, 1989; Klauer & Oberauer, 1995), demonstrate that reasoning performance deteriorates drastically as the number of possibilities to be considered in memory decreases; even two possibilities often lead to a large decrement in performance. If people do routinely represent all possibilities, such
error patterns should not occur. Working memory limitations may play a role in this phenomenon, and participants may represent only one possibility, as a default, for a given proposition (e.g., Rader & Sloutsky, 2000a; Sloutsky & Goldvarg, 1999).

Secondly, cognitive-developmental theories seem exceedingly vague in that they make no attempt to distinguish different phases of the inference process, and no attempt to detail the cognitive processes underlying reasoning competence. Indeed, Overton (1990) argues that these theories need not specify such processes in too much detail. However, this position seems untenable if it is accepted that representations formed during comprehension serve as inputs into deliberate reasoning processes. More pragmatically, some specification of processing phases and representations could help set boundaries for the theory. Given the theory’s rational emphasis, it is not clear whether cognitive-developmental theory intends to apply to inferences during comprehension. Only one author has suggested as much (Franks, 1996); other authors have not addressed this issue (e.g., Moshman, 1990; Overton, 1990). If these issues were addressed, one could devise a more informative test of the cognitive-developmental point of view.

Perhaps critically, if the theory intends to describe deliberate, higher-level reasoning, findings regarding inferences in comprehension could feasibly be integrated with the developmental account, which does have some success in predicting age differences in deliberate hypothesis-testing and modal reasoning tasks (Byrnes & Overton, 1986; Pieraut-Le Bonniec, 1980; Scholnick & Wing, 1988; Ward, et al., 1990). However, cognitive-developmentalists have tended to speak of conditional representations and structures in a global manner that does not seem to distinguish among tasks or task goals.
What about other theories of conditional processing? A wide range of theories exist that purport to explain conditional reasoning in specific contexts. For example, one theory assumes that conditionals dealing with permission automatically trigger innate, evolved "cheater-detection" algorithms (Cosmides, 1989). Another proposal (Manktelow & Over, 1991) is that conditional reasoning essentially involves a computation of utilities associated with various outcomes and social roles implied in the conditionals. Still other proposals argue that conditional reasoning relies upon the application of abstract but domain-specific schemas (Cheng & Holyoak, 1985), while another account relies upon linguistic pragmatics (Light, et al., 1989).

However, these accounts suffer from two difficulties. First, all of those mentioned above, as well as many others, were developed and tested almost exclusively in the context of the Wason selection task. This task, described earlier, may invoke hypothesis-testing processes and linguistic heuristics distinct from processes involved in deductive reasoning. Secondly, most of these accounts do not seem to generalize beyond the specific problem contents within which they were developed. Indeed, syntactic, semantic and cognitive-developmental theories seem to offer the best attempts at domain-general accounts of reasoning.

An approach worth mention is the rational theory of Oaksford, Chater, and colleagues (Oaksford & Chater, 1996; Oaksford, Chater, Grainger, & Larkin, 1997; Oaksford, Chater, & Larkin, 2000). These authors have proposed that conditional reasoning cannot be based on a mental logic because such processes are computationally intractable in humans. Instead, they propose that participants process conditional arguments by computing, at some level, conditional probabilities, rather than applying
inference schemas or forming mental models (Chan & Chua, 1994; and Stevenson & Over, 1995, give other probabilistic accounts of conditional reasoning). In the case of MP, the factor determining whether one draws the inference is the probability of an exception to the conclusion. People assume such exceptions to be rare (Oaksford & Chater, 1996), hence MP is almost always endorsed in most contexts. With the other conditional inferences, participants draw conclusions to an argument in direct proportion to how probable they regard the categorical premise of the argument. Thus, an AC inference (If P then Q; Q; therefore P) is more likely to be drawn as the participant’s assessment of the probability of Q increases.

Oaksford, Chater, and colleagues have produced results supporting the predictions of their rational probabilistic theory for a variety of phenomena (Oaksford & Stenning, 1992; see also Schroyens, et al., 2000). Importantly, one study demonstrated that the predictions held when probability information was NOT supplied to participants; instead, conditional premises were used in which the antecedents and consequents varied in probability as assessed by other participants (Oaksford, et al., 2000, Experiment 3). This type of assessment is similar in spirit to that carried out by other researchers who have examined the enhancement/suppression effects associated with AC and other inferences (e.g., Cummins, 1995). This raises the speculation that these effects might be due to participants’ assessments of the probability of the conclusion for each argument, rather than the retrieval (or lack thereof) of alternative antecedents.

However, this speculation has difficulties when applied to the current experiments. Because content manipulations in my studies had no appreciable effect on the tendency to draw inferences in comprehension, probabilistic effects that Oaksford and
Chater would predict could not have occurred. It seems unlikely that participants are computing any type of probabilistic inference during comprehension. Probabilistic effects may only come into play when participants are given directions to reason.

One other conditional reasoning theorist who should be mentioned is Evans (1984, 1993; Evans & Over, 1996). He has advocated the heuristic-analytic theory of reasoning. According to this view, participants’ representations of conditionals are guided by preconscious heuristics that focus attention on possibilities consistent with the conditional; in particular, those propositions that are explicitly stated are typically represented, but other possibilities (e.g., possibilities involving a false antecedent) are not (Evans, 1984). Analytic reasoning processes then operate upon these representations to produce conclusions. Note that the assumption that participants tend to represent what is stated accords with the standard mental models assumption that people tend to represent explicit true possibilities. Furthermore, although Evans (1984) did not speculate on the nature of the analytic processes, he has since aligned himself with standard mental models by adopting its assumptions about representation, premise combination, and search for counterexamples (Evans, 1993, 2000; Evans & Over, 1996). Thus, although the findings and arguments in this work could be said to support Evans’ theory as well, the fact that Evans has adopted the standard mental models account makes that point relatively inconsequential.

The relationship between comprehension and reasoning

A basic tenet of syntactic theories is that conditional reasoning is largely a component of language comprehension and procedural in nature (Braine, 1990; Braine & O’Brien, 1991; Lea, 1995; Macnamara, 1986). Current findings do not support this
argument. In particular, a factor known to influence the relative rates at which participants deliberately commit MP and AC inferences (the availability of alternative antecedents, as measured by strength of association between antecedent and consequent) appears to have no effect on the tendency to draw inferences during comprehension. This suggests that deductive reasoning is not exhausted by comprehension.

Other studies seeking to find deductive inferences in comprehension also hint at this conclusion. For example, transitive inferences involving unquantified relational terms (e.g., \(A\) is longer than \(B\); \(B\) is longer than \(C\); therefore, \(A\) is longer than \(C\)) and involving quantifiers (e.g., All the \(A\)'s are \(B\)'s; All the \(B\)'s are \(C\)'s; therefore, All the \(A\)'s are \(C\)'s) have been examined in several text comprehension studies. Valid transitive inferences with unquantified terms seem to be drawn even when conclusions are not presented, and invalid inferences are usually rejected (Evans, et al., 1993; Favrel & Barrouillet, 2000). However, with quantified terms, people draw fewer and fewer valid inferences as the amount of intervening material between premises increases (e.g., Griggs, 1976), and they draw many invalid inferences (e.g., participants given All the \(A\)'s are \(B\)'s may agree that they read All the \(B\)'s are \(A\)'s). These effects occur even though participants are much better at recognizing valid inferences and logical relationships between such premises when they are given explicit reasoning tasks with similar materials (Griggs & Warner, 1982; Nguyen & Revlin, 1993). Apparently, people may initially represent All the \(A\)'s are \(B\)'s as being synonymous with All the \(B\)'s are \(A\)'s, leading to the high proportion of false inferences when premise terms are adjacent in the text. Also, intervening material places demands on working memory, so that valid inferences that involve premises separated in the text become less likely (Favrel &
Barrouillet, 2000). This pattern of data reinforces the general conclusion that inferences during comprehension may not reflect the actual application of whatever deductive reasoning ability people actually possess. Additionally, these data suggest that many deductive inferences in comprehension are based on initial representations of discourse information, and that considerations of validity, necessity, and so forth are generally unutilized at this phase.

Comprehension is a rapid automatic process that seeks to establish a coherent representation of a text (Gernsbacher, 1997; Kintsch, 1998; Zwaan & Radvansky, 1998), and one could argue that both MP and AC inferences accomplish this goal. These inferences maintain coherence by allowing a person to minimize the number of possibilities to be considered, a factor to which people seem to be sensitive (Johnson-Laird, 1999; Sloutsky & Goldvarg, 1999). It can also be argued that these inferences, by connecting different propositions in a story, maintain a general causal flow between events and their consequences, a flow which comprehension processes seem to seek automatically (Van den Broek, 1994).

It is intriguing that one general finding from the discourse comprehension literature is consistent with a tenet of the standard mental models theory: people find it easiest to create and maintain only one discourse model. Recall that Johnson-Laird and colleagues (Johnson-Laird, 1999; Johnson-Laird, et al., 1992) argue that people prefer to represent as few possibilities as possible. This hypothesis is similar in spirit to the idea that comprehenders prefer to work with only representation of a discourse. Many studies demonstrate, for example, that comprehenders find it easiest to process discourse that contains one definite spatial arrangement and time course; indefinite arrangements slow
down comprehension and are not integrated as well in memory (Zwaan & Radvansky, 1998). As another example, comprehenders try to integrate events in a text into a single coherent causal chain; processing difficulty results, and memory is poorer, if this integration cannot take place (Van den Broek, 1994). A discourse model representing a single situation (possibility) seems to be the default.

Thus, considerations from discourse processing suggest that comprehenders try to maintain only one possibility when representing a text; drawing MP and AC inferences may allow the maintenance of a single-possibility, coherent representation. Additionally, it was argued in the introduction that many inferences and problem representations may be bounded and logically incomplete, on grounds of reducing cognitive complexity and being applicable in many real-world situations. The postulations of standard mental models theory, in terms of inferences drawn in comprehension, are consistent with both views. It thus seems reasonable to assume that inferences in comprehension will show such "boundedness," and that deliberate reasoning involves the application of further strategies. These strategies may involve of search for counterexamples or even be suppositional in nature, and different individuals may use different strategies (Evans, 2000; Johnson-Laird, et al., 2000; Stanovich, 1999). However, reasoning cannot be reduced to comprehension processes to the extent that syntactic theories have suggested.

A consequence of this position is that, just because some inferences drawn in comprehension appear to mimic those drawn during deliberate reasoning, one can assume neither that both inferences were drawn using the same processes nor that they represent the same abilities. Thus, an inference conforming to the MP argument form may not
reflect the application of deductive competence, and an AC inference in comprehension may not reflect the lack of such competence. Instead, the nature of the inferences drawn may vary based on the task.

Subsequent studies could further elucidate the relationship between comprehension processes and reasoning, to examine the nature of representations that underlie reasoning, comprehension, and the overlap between the two. For example, some studies of the enhancement/suppression of conditional inferences have examined causal conditionals such as *If the brake is depressed, then the car slows down* (e.g., Cummins, 1995; Quinn & Markovits, 1998; Thompson, 1994). Additionally, causal inferences are widely studied in research on discourse processing, and it is known that other sentential connectives (e.g., *because*) can facilitate causal inferences and lead to the integration of propositions into memory (Caron, Micko, & Thüring, 1988; Millis & Just, 1994; Murray, 1997). and even young children are sensitive to causal constraints in a story (Casteel, 1993; Van den Broek, 1989). The reported experiments could be replicated with causal conditionals to see whether causal conditionals also facilitate inferences. Such experiments could also determine whether the memory association effects found in conditional reasoning tasks have parallels in discourse comprehension in this domain (i.e., systematic suppression and enhancement of AC and MP inferences based on variations in the causal connection between antecedent and consequent). I did not find such a parallel, but this does not eliminate the possibility that it could exist in the causal domain. However, if the parallel were not found and instead the present findings were
replicated, this would lend credence to the hypothesis that conditional inferences in comprehension are only initial, non-deductive inferences needed to maintain coherence, with deliberate reasoning depending on a further stage or stages of strategy usage.

Finally, these experiments could be repeated with younger children to give a more complete developmental picture, although some methodological difficulties would need to be overcome. Nevertheless, it would be interesting to determine whether the present findings hold for younger age groups. Also, studies demonstrate that young children experience difficulty even in drawing valid conditional inferences when the premises are contrary to fact, but adults have practically no problem (Markovits & Vachon, 1989, 1990). If the hypothesis in the present work is correct, then adults and children as well should draw such inferences in comprehension when such conditionals are presented in plausible stories. Such a finding would confirm the hypothesis that inferences in comprehension serve to maintain a single possibility and a coherent mental model of the discourse, rather than drawing on ideas about logical necessity and validity.

Another possible topic for future studies would be an examination of individual differences, if they exist, in the tendency to draw different conditional inferences in comprehension. Some studies with adults (Markovits, 1984, 1985) and with children (Janveau-Brennan & Markovits, 1999) suggest that participants who are more likely to think of alternative antecedents to conditionals are also less likely to commit the fallacies in reasoning tasks. One intriguing finding of another study is that participants who were more likely to draw the AC inference were also less-skilled readers; no individual differences existed in the tendency to draw MP inferences because they were drawn at ceiling rates by all readers regardless of reading level (Franks, 1997). (It is commonplace
to find that poor readers draw fewer inferences of many types than do more able readers
(Oakhill, 1996). This result suggests important implications for theories of conditional
processing, inasmuch as none of them directly address the issue of individual differences.

In sum, both adult and sixth-grade participants appear to draw both valid MP and
invalid AC inferences in comprehension. Furthermore, a factor known to influence the
tendency to draw invalid AC inferences, for participants of both age groups, in deliberate
reasoning tasks has no appreciable effect on the tendency to draw the same sorts of
inferences in comprehension tasks. These findings argue against the idea that
conditionals in comprehension are processed through the application of logically valid
inference schemas. Instead, participants seem to form economical but logically
incomplete mental models based on the conditionals, as considerations of discourse
processing and inference in other domains suggest. Furthermore, the findings imply that
comprehension processes and reasoning processes need to be conceptually separated.
| Table A.1. | Stimulus story for the noun *night*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics. |
| Table A.2. | Stimulus story for the noun *lamp*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics. |
| Table A.3. | Stimulus story for the noun *snow*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics. |
| Table A.4. | Stimulus story for the noun *candy*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics. |
| Table A.5. | Stimulus story for the noun *book*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics. |
Table A.6. Stimulus story for the noun *thief*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
<table>
<thead>
<tr>
<th>Sentence</th>
<th>MP Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was.</td>
</tr>
<tr>
<td>2</td>
<td>He thought that if it was dark/cold outside, then it was night.</td>
</tr>
<tr>
<td>3</td>
<td>Still feeling sleepy, Frank arose to look out a window.</td>
</tr>
<tr>
<td>Inference</td>
<td>He saw that it was dark/cold outside.</td>
</tr>
<tr>
<td>No-Inference</td>
<td>He wondered whether it was dark/cold outside.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sentence</th>
<th>AC Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Frank woke up on his couch after taking a long nap, and realized that he didn’t know what time it was.</td>
</tr>
<tr>
<td>2</td>
<td>He thought that if it was night, then it was dark/cold outside.</td>
</tr>
<tr>
<td>3</td>
<td>Still feeling sleepy, Frank arose to look out a window.</td>
</tr>
<tr>
<td>Inference</td>
<td>He saw that it was dark/cold outside.</td>
</tr>
<tr>
<td>No-Inference</td>
<td>He wondered whether it was dark/cold outside.</td>
</tr>
</tbody>
</table>

Probe presented after the story in Experiments 1 and 3: ** NIGHT **

Inference sentence presented for recognition after the story in Experiments 2 and 4

The time of day was night.

Table A.1. Stimulus story for the noun *night*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
<table>
<thead>
<tr>
<th>Sentence</th>
<th>MP Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jill looked at the strange object in her attic, and wondered what it could be.</td>
</tr>
<tr>
<td>2</td>
<td>She thought to herself, “If it gives light/has a switch, then it is a lamp.”</td>
</tr>
<tr>
<td>3</td>
<td>Jill walked through the crowded attic over to the object.</td>
</tr>
</tbody>
</table>

| Inference | After studying the object for a few minutes, Jill discovered that it did give light/have a switch. |
| No-Inference | Jill studied the object and tried to discover whether it gave light/had a switch. |

<table>
<thead>
<tr>
<th>Sentence</th>
<th>AC Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jill looked at the strange object in her attic, and wondered what it could be.</td>
</tr>
<tr>
<td>2</td>
<td>She thought to herself, “If it is a lamp, then it gives light/has a switch.”</td>
</tr>
<tr>
<td>3</td>
<td>Jill walked through the crowded attic over to the object.</td>
</tr>
</tbody>
</table>

| Inference | After studying the object for a few minutes, Jill discovered that it did give light/have a switch. |
| No-Inference | Jill studied the object and tried to discover whether it gave light/had a switch. |

**Probe presented after the story in Experiments 1 and 3: **LAMP**  
Inference sentence presented for recognition after the story in Experiments 2 and 4**

The object that Jill found was a lamp.

Table A.2. Stimulus story for the noun *lamp*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
<table>
<thead>
<tr>
<th>Sentence</th>
<th>MP Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Newt was traveling the next day and needed to know what the weather would be like at his destination.</td>
</tr>
<tr>
<td>2</td>
<td>He believed that if his destination was <em>cold/had weather found in winter</em>, then there would be snow.</td>
</tr>
<tr>
<td>3</td>
<td>Newt decided to watch a weather forecast before he packed his suitcase.</td>
</tr>
<tr>
<td>Inference</td>
<td>Eventually, a weather report on TV informed him that his destination would be <em>cold/have weather found in winter</em>.</td>
</tr>
<tr>
<td>No-Inference</td>
<td>He found a TV weather forecast and watched in case the forecaster predicted <em>cold/winter weather</em>.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sentence</th>
<th>AC Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Newt was traveling the next day and needed to know what the weather would be like at his destination.</td>
</tr>
<tr>
<td>2</td>
<td>He believed that if there was snow, then his destination would be <em>cold/have weather found in winter</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Newt decided to watch a weather forecast before he packed his suitcase.</td>
</tr>
<tr>
<td>Inference</td>
<td>Eventually, a weather report on TV informed him that his destination would be <em>cold/have weather found in winter</em>.</td>
</tr>
<tr>
<td>No-Inference</td>
<td>He found a TV weather forecast and watched in case the forecaster predicted <em>cold/winter weather</em>.</td>
</tr>
</tbody>
</table>

**Probe presented after the story in Experiments 1 and 3: **SNOW**

**Inference sentence presented for recognition after the story in Experiments 2 and 4**

At Newt's destination, there was snow.

Table A.3. Stimulus story for the noun *snow*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
Alicia was pleased when her father told her that he had gotten her a surprise.

She guessed that if it was *sweet/colorful*, then it was candy.

She asked him excitedly, “What did you buy me?”

Her father smiled and told her that he had bought something *sweet/colorful*.

Alicia would be pleased to receive something *sweet/colorful*.

Table A.4. Stimulus story for the noun *candy*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
Table A.5. Stimulus story for the noun *book*. In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
<table>
<thead>
<tr>
<th>Sentence</th>
<th>MP Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hilary was a member of a jury at the courthouse, and she wondered why the defendant in the first case had been arrested.</td>
</tr>
<tr>
<td>2</td>
<td>If he had been caught stealing/wearing black, then he was a thief.</td>
</tr>
<tr>
<td>3</td>
<td>The prosecutor began to tell the jury why the defendant was being put on trial.</td>
</tr>
</tbody>
</table>

**Inference**
Before long, Hilary heard that the defendant had been caught stealing/wearing black.

**No-Inference**
Hilary supposed that the defendant might have been caught stealing/wearing black.

<table>
<thead>
<tr>
<th>Sentence</th>
<th>AC Versions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hilary was a member of a jury at the courthouse, and she wondered why the defendant in the first case had been arrested.</td>
</tr>
<tr>
<td>2</td>
<td>If he was a thief, then he had been caught stealing/wearing black.</td>
</tr>
<tr>
<td>3</td>
<td>The prosecutor began to tell the jury why the defendant was being put on trial.</td>
</tr>
</tbody>
</table>

**Inference**
Before long, Hilary heard that the defendant had been caught stealing/wearing black.

**No-Inference**
Hilary supposed that the defendant might have been caught stealing/wearing black.

**Probe presented after the story in Experiments 1 and 3:**
**THIEF**

**Inference sentence presented for recognition after the story in Experiments 2 and 4**
The defendant was a thief.

---

Table A.6. Stimulus story for the noun *thief*: In sentence 2, the Inference sentence, and the No-Inference sentence, the strong associate precedes the weak associate in italics.
BIBLIOGRAPHY


McCawley, J. D. (1981). *Everything that linguists have always wanted to know about logic but were ashamed to ask.* Chicago: University of Chicago Press.


123


