Since the inception of the current deduction paradigm (Henle, 1962), researchers have been challenged to construct problems that accurately measure an individual’s reasoning ability and to illuminate the mental processes associated with reasoning. Much of the research conducted within the deduction paradigm has utilized abstract material; however, some researchers question the ecological validity of using abstract material, stating that real world reasoning problems always incorporate some kind of concrete or thematic content (Hertwig, Ortmann, Gigerenzer, 1997). As a result, several lines of research have arisen to investigate the various effects of thematic material. However, few studies have directly compared individuals’ performances using abstract and thematic material. Moreover, research regarding this topic has not produced a coherent theory or data set of the effects of thematic material. Thus, the purpose of the present study is to directly compare abstract and pragmatic performances on a immediate inference reasoning task within the framework of the mental model theory.

In the field of human reasoning, one of the most influential and studied theories is that people create mental models when solving problems. First proposed by Johnson-Laird (1983), the mental model theory attempts to reveal the cognitive processes associated with deduction, a fundamental aspect of human intelligence. Deduction is the application of reasoning specific mental processes to produce logically valid inferences from a series of information (Polk & Newall, 1995). The model theory explains deduction as a process of three components: (a) model formation, encoding a problem into a mental representation in the form of a model; (b) conclusion formation, searching for a putative conclusion consistent with, but not the same as the premise; and (c) conclusion validation, searching for counterexamples, or alternative models, to validate or falsify the conclusion (Evans, Handley, Harper & Johnson-Laird, 1999). Conclusion validation is the most critical component of the model theory, because it is the only stage where deductive specific mechanisms are utilized – the model formation and conclusion formation stages are simply processes of comprehension. The model theory states that conclusions are deemed valid if there are no counterexamples that contradict it.

The Effects of Pragmatic Content on Syllogistic Reasoning

Performances on reasoning tasks have been known to be affected by the type of material presented to the participants. In this study, we examined performance differences between pragmatic and abstract material when solving immediate inference problems using necessity instructions. Participants were given 16 immediate inferences with abstract material and 16 immediate inferences will pragmatic material. An overall significant facilitation effect was found for pragmatic material (p<.01). However, the pragmatic facilitation effects were found for specific problems, not for the entire problem set. Thus, the effects of the study’s specific pragmatic material cannot be determined at this time.

*Faculty supervisor
The origins of the mental model theory and the current deduction paradigm have been largely shaped by Peter Wason, specifically by the Wason selection task (Wason, 1966). This task uses logic as the normative standard to measure human reasoning. The task employs the logical structure of an “if, then” conditional inference problem (Wason & Shapiro, 1971). As a result of this benchmark, the majority of the research on human reasoning uses logic as the means to assess reasoning competency. Specifically, there has been an extensive amount of work conducted using the four conditional inferences (De Neys, Schaeken & D'Ydewalle, 2003):

1. The logically valid, modus ponens (MP), e.g., “if $p$ then $q$, $p$ therefore $q$.”
2. The logically invalid, affirmation of the consequent (AC), e.g., “if $p$ then $q$, $q$ therefore $p$.”
3. The logically invalid, denial of the antecedent (DA), e.g., “if $p$ then $q$, not $p$, therefore not $q$.”
4. The logically valid, modus tollens (MT), e.g., “if $p$ then $q$, not $q$, therefore not $p$.”

The rules of logic are based upon mathematical principles that were created to demonstrate a method of fabricating axioms that are absent of contradictions (Evans, 2002). The premise (if $p$ then $q$) of the four arguments is assumed to be true; thus, the arguments’ logicality is based solely on its syntactic form (Thompson, 1994). As a result, when solving a logic problem, a correct evaluation is one in which the form of the conclusion logically follows the premise (i.e., MP is logically valid because the form of “$p$ therefore $q$” is consistent with the premise). Abstract material (i.e., arbitrary letters of the alphabet, numbers, and shapes) has provided the best method of measuring an individual’s ability to apply logic, or to think logically. The purpose of using abstract material was to determine if reasoning logically was an innate, domain-general cognitive process (Evans, 2002). Introducing real-world material was believed to confound the logicality of participants’ responses by eliciting biases due to differences of knowledge bases and differences in contextual understanding. Thus, because logic has been the normative standard within the modern reasoning paradigm, the majority of the reasoning research has used abstract material to assess human reasoning (Newstead, Ellis, Evans & Dennis, 1997).

However, many researchers question the ecological validity of laboratory reasoning tasks that use abstract material (George, 1999; Henle 1962; Hertwig et al., 1997; Mankalow & Over, 1991; Newstead et al., 1997; Oaksford & Charter, 1998). Specifically, Hertwig and colleagues (1997) argued that, because real world problems always incorporated some kind of content and context, abstract conditional inference problems did not map well onto everyday reasoning processes. Moreover, they stated that problems like the Wason’s selection task failed to measure any kind of social rationality because standard logic was not the appropriate tool to use when measuring the social structure of an environment.

In response to the argument for greater ecological validity, a host of researchers have incorporated, and extrapolated numerous effects from content and contextually-rich materials. Within the framework of the mental model theory of reasoning, Cummins, Lubart, Alksnis, and Rist (1991) found that conclusion validation is affected by the number of alternative causes and disabling conditions that are readily retrievable from the semantic memory system. Within a conditional inference, an alternative cause is a cause other than the one stated in the premise that can lead to the putative conclusion. For example, given the logically invalid AC form conditional:

When it is cloudy you, cannot see the sun.
You can not see the sun, therefore it is cloudy.

For this problem, an alternative cause for not being able to see the sun would be that it is night time. Likewise, a disabling condition is a conclusion other than the one stated that can result from the given premise. Cummins and colleagues (1991) found that the more alternative causes and disabling conditions that were available in an individual’s semantic memory system, the less likely they were to endorse the logically invalid AC and DA form conditionals. This led to the development of the semantic memory framework of conditional reasoning, which states that possible alternative causes and disabling conditions are stored in the semantic memory system and retrieved when assessing a conditional inference problem (De Neys, Schaeken, & D’Ydewalle, 2002, 2003).

Markovits and Quinn (2002) also cited memory retrieval as a key process for reasoning and generating alternative models, stating that the speed of retrieval correlated with a greater number of accurate responses to conditional inference problems. They conducted a study that was designed to measure the ability of participants to retrieve pertinent information from their semantic memory. Participants were asked to generate two causes that would produce a given conclusion as quickly as they could. For example, participants were asked to list two causes for the conclusion: “a dog scratches a lot.” For each conclusion given, there was an alternative cause with a strong association (in this case “the dog has fleas”) and several alternative causes with weak associations (i.e., the dog’s skin is dirty, or the dog has a skin disease). They found that gener-
ally people quickly generated the strong association alternative and that the speed of retrieval of a weak association alternative was clearly correlated with the ability to correctly evaluate AC and DA form conditional inferences.

Despite the abundant evidence that content-based material, in various ways, affects the ability of participants to generate logically appropriate responses to conditional inference problems, few studies have been conducted under the framework of the mental model theory that directly analyze the differences between abstract and thematic reasoning tasks. Research that has examined the differences between abstract and thematic content has produced discordant results (Evans, 2002; Griggs & Cox, 1983; Mankelow & Evans, 1979; Wason & Shapiro 1971, experiment 2). Several studies have found that thematic material, compared to abstract material, significantly improved performance on the Wason selection task (Bracewell & Hidi, 1974; Johnson-Laird, Legrenzi, & Sonino Legrenzi, 1972; Wason & Shapiro, 1971, experiment 2). These findings, however, have been challenged by numerous studies that have not found significant facilitation effects for thematic material (Brown, Keats, Keats & Seggie, 1980; Griggs & Cox, 1983; Mankelow & Evans, 1979; Yachnin & Teney, 1982).

The discrepancies between the two findings may be due, at least in part, to differences in the familiarity with the thematic material used, and in the types of thematic material used. For example, Johnson-Laird and colleagues (1972), who found a significant thematic facilitation effect, instructed British participants, who were familiar with a specific rule in the British postal system that was pertinent to the study, to pretend that they were postal workers and verify a rule such as “If a letter is sealed, then it has a 50 lire stamp on it.” In their attempt to replicate Johnson-Laird et al.’s (1972) findings, Griggs & Cox (1983), who used American participants who were unfamiliar with the British postal rule, completely failed to find a significant facilitation effect of thematic content. Taking the participants’ unfamiliarity of the postal rule into consideration, Griggs and Cox suggest that thematic material can facilitate performance on the Wason selection task by a process of memory-cueing that allows for the retrieval of pertinent alternative or falsifying causes (i.e., alternative models).

Overall, research on this topic has not produced a coherent theory or data set about the effects of thematic material. Moreover, the majority of this research has used the Wason selection task as the means to compare abstract and thematic content. The present study differs in that it utilizes immediate inference problems (similar to syllogisms) as a means to compare abstract and pragmatic content. The present study also differs from previous abstract and thematic comparison research in that the material utilized for the present study was pragmatic; it was designed to incorporate situations that are representative of the real world. Often the thematic material used in previous research was fabricated from arbitrary content that was not necessarily representative of real world contexts. The present study’s goal was to present pragmatic material that would be familiar to participants. Thus, because conditional inference problems require the participant to assume that the premise is true, each pragmatic premise was constructed to actually represent a situation that could occur in the real world.

The purpose of designing pragmatic premises consistent with the real world was to facilitate any potential memory-cueing that could lead to the retrieval of pertinent alternative models (thus the goal was to avoid arbitrary pragmatic content that would not involve any memory retrieval processes).

Immediate Inference Problems

Syllogisms are conditional inference problems that combine two premises and a conclusion, for example, All A are B (premise 1); All B are C (premise 2); Therefore, All A are C (conclusion). A syllogism’s conclusion and premises must contain four quantified forms, which are described as “moods,” and classified by the abbreviations A, E, I, and O (Evans et al., 1999). These are the following: A, universal affirmative (All P are Q); E, universal negative (No P are Q); I, particular affirmative (Some P are Q); O, particular negative (Some P are not Q).

Immediate inferences are problems that use one mood as a premise, and another mood as a conclusion, for example, All P are Q: Therefore, some Q are P. The order of the terms (P and Q) can also be exchanged; thus there are seven possible conclusions that follow each of the four premise forms (excluding the repetition of the premise). Immediate inferences, like syllogisms, have conclusions that can be classified into three types: Necessary (the conclusion must be true given that the premise is true); Possible (the conclusion could be true, but it is not necessarily true given that the premise is true); Impossible (the conclusion cannot be true given that the premise is true).

Recent attention has been devoted to peoples’ ability to differentiate necessary, possible, and impossible conclusions. To determine if a conclusion to a syllogism or immediate inference problem is necessary, all possible models must be exhaustively searched (to rule out any models that could possibly falsify the putative conclusion). But to determine if a conclu-
sion is possible, an individual only has to search for one model that leads to a correct conclusion.

Evans and colleagues (1999, experiment 1) examined the effects of Necessity instructions and Possibility instructions for an immediate inference task. Consistent with their predictions, they found that people were more likely to accept conclusions as possible than as necessary.

The present study is an adaption of Evans et al.’s (1999) experiment 1. The present study uses only Necessity instructions, because the focus of the research was to examine performance differences between abstract and pragmatic content. For each immediate inference problem using abstract content, there was a corresponding immediate inference problem of the same form that used pragmatic content. Based on the burgeoning research that has shown that memory retrieval and model generation are substantial components to the reasoning process (Cummins, 1995; Cummins et al., 1991; De Neys et. al., 2002, 2003; Markovits & Quinn, 2002) we predict that pragmatic content would facilitate better performance than abstract content on an immediate inference reasoning task.

Method

Participants

Eighty-seven undergraduate students (67 women, 20 men; M = 22.4 years) with a major or minor in psychology at a Midwestern comprehensive university participated in this study. Participants were mostly white students from middle to upper class backgrounds. Participants were selected from four upper level psychology courses and were asked to voluntarily complete the study as part of a classroom activity.

Materials

Each participant received a packet with the consent form, instructions, and the immediate inference problems, all printed on white computer paper. Participants completed the immediate inference problems using a paper and pencil format. The instructions were printed on an individual sheet and read as follows:

The following packet contains 32 syllogisms. Answer each syllogism to the best of your knowledge. Each syllogism contains a premise (GIVEN THAT) and a conclusion (IT IS NECESSARY THAT). You are asked to mark “Yes” or “No” based on if the conclusion is absolutely necessary. Answer “Yes” to a conclusion only if it is necessarily true given that the premise is true. There is no time limit.

Sixteen immediate inference problems were selected: four with necessary conclusions, eight with possible conclusions, and four with impossible conclusions. For each immediate inference problem there was an abstract form and a pragmatic form, such as:

Abstract:

GIVEN THAT
All M are X
IT IS NECESSARY THAT
Some X are M
☐ Yes ☐ No

Pragmatic:

GIVEN THAT
All squirrels are mammals
IT IS NECESSARY THAT
Some mammals are squirrels
☐ Yes ☐ No

Refer to Appendix for a complete list of the pragmatic problems. Each sheet of paper contained four immediate inference problems. Two classes (45 participants) received 16 consecutive, arbitrarily-ordered pragmatic problems, followed by 16 consecutive, arbitrarily-ordered abstract problems. The other two classes (42 participants) received the consecutive, arbitrarily-ordered abstract problems first and the consecutive, arbitrarily-ordered pragmatic problems second. Letters were randomly selected for the content of the abstract problems. The pragmatic problems were constructed such that the content of the problems were designed to represent a situation, or instance, that could occur in real life. The content selected was designed to be familiar to the participants. The content included topics of basic scientific knowledge, such as the temperature when water freezes, as well as many day-to-day activities, such as recycling paper. The content of the pragmatic problems also had to adhere to the logical structure of the immediate inference. Thus, even if a participant was not familiar with the content of the pragmatic problem, he or she could still correctly solve the problem by means of assessing its logical structure. The goal, however, was to utilize content that was familiar to participants in a way that involved potential model generation and information retrieval from the participants’ memory systems.

Design and Procedure

The study utilized a 2 X 2 within-subjects repeated measures design. The independent variables were the immediate inference type (abstract and pragmatic) and the order in which the participants received the immediate inference problems (abstract first or prag-
matic first). Each participant completed both the abstract problems and the pragmatic problems.

Participants were divided into four groups and completed the immediate inference problems in a classroom setting with individual desks. Each participant worked independently at his or her own pace. Participants were briefly instructed about the content and purpose of the study, and instructed that they could discontinue the study at anytime without penalty. Participants were asked to mark their age, gender, and whether they had had a formal course in logic. Two participants were excluded for having a formal course in logic because they would have had prior exposure to the rules of conditional inferences. Participants were instructed to read each question carefully. The problems were distributed approximately 25 min before the class period officially ended, so although there was no artificial time constraint, realistically participants had about 25 min to complete the problems. Upon completion, participants handed in the survey and were free to leave the classroom.

Results

A significant facilitation effect was found for the problem type. Participants had significantly fewer incorrect responses for the pragmatic problems (M = 5.48, SD = 2.02) than for the abstract problems (M = 6.22, SD = 1.98), F(1,85) = 14.3, p < .01; figure 1). Although the results were significant, η² = .143 indicated a relatively small effect size. There was no order effect for the problem type, F(1,85) = 2.19, p > .05 indicating that participants were not significantly affected by the order (abstract first or pragmatic first) of the problems.

There was a large range (2-92%) of percentage endorsements of incorrect responses for problem forms is consistent with results of Evans and colleague’s (1999) previous research. These data support the hypothesis that some forms of immediate inference problems lead to an initial model that supports the conclusion (Evans et al., 1999). These data support this hypothesis even when pragmatic material is introduced (see Table 1).

Given that there was not a consistent facilitation effect of pragmatic material, the reason for the specific facilitations cannot be determined. It cannot be determined if the specific pragmatic material of each question was the cause of the specific facilitation effects, or if any type of pragmatic material would have produced similar facilitating effects. Table 1 indicates that some problem forms resulted in a pragmatic facilitation effect, while other problem forms resulted in an abstract facilitation effect. The reasons for the specific facilitation effects cannot be determined at this time due to the preliminary nature of the content utilized for this study. Further research is needed to validate the effects of our specific pragmatic material and its interaction with the problem forms.

Although the pragmatic problems were designed to represent general knowledge-type situations that could occur in the real world, one reason for the inconsistent facilitation effects could have been that participants may have had different levels of knowledge.

Discussion

Overall, our results indicate that there was a significant facilitation effect of pragmatic material; however, the facilitation only occurred with certain items and not others (see Table 1). The large range (2-92%) of percentage endorsements of incorrect responses for problem forms is consistent with results of Evans and colleague’s (1999) previous research. These data support the hypothesis that some forms of immediate inference problems lead to an initial model that supports the conclusion (Evans et al., 1999). These data support this hypothesis even when pragmatic material is introduced (see Table 1).

Given that there was not a consistent facilitation effect of pragmatic material, the reason for the specific facilitations cannot be determined. It cannot be determined if the specific pragmatic material of each question was the cause of the specific facilitation effects, or if any type of pragmatic material would have produced similar facilitating effects. Table 1 indicates that some problem forms resulted in a pragmatic facilitation effect, while other problem forms resulted in an abstract facilitation effect. The reasons for the specific facilitation effects cannot be determined at this time due to the preliminary nature of the content utilized for this study. Further research is needed to validate the effects of our specific pragmatic material and its interaction with the problem forms.

Although the pragmatic problems were designed to represent general knowledge-type situations that could occur in the real world, one reason for the inconsistent facilitation effects could have been that participants may have had different levels of knowledge.

![Figure 1](image-url)
about the specific material used for the pragmatic problems. Also, differences in interpretation of the pragmatic material may have contributed to the inconsistent facilitation effect. Thus, our results may indicate a form or content specific, but not a generalized facilitating effect of pragmatic material when solving immediate inference problems.

In addition, the number of alternative causes and disabling conditions was unknown for the pragmatic problems. Cummins and colleagues (1991), demonstrated that the number of possible alternative causes and disabling conditions can affect an individual’s ability to correctly evaluate logic problems. Thus, because of the lack of information about the number of models for each problem, the present study cannot determine if the specific pragmatic material used facilitated any memory-cueing or retrieval of pertinent models. Future studies could be conducted to validate the number of alternative models and disabling conditions present for each pragmatic problem.

Our results, however, support the hypothesis that thematic facilitation effects are highly dependent on the specific content and context. In a manipulation of the Wason selection task where participants were given different instructions and perspectives, Manktelow and Over (1991) demonstrated that the effects of thematic content varied depending on the instructions and the participant’s perspective. Our data also support Grigg’s (1983) hypothesis that thematic content does not simply improve the logicality of people’s responses.

The theoretical implications of our study indicate that pragmatic material is capable of inducing facilitation effects on immediate inference problems. However, pragmatic content does not produce a generalized facilitating effect on immediate inference problems. Our data indicate that the specific pragmatic material used, as well as the problem form may interact to create a localized facilitation effect. Further research is needed to untangle the many sources of

### TABLE 1

Percentage of Incorrect Responses for the Individual Immediate Inference Forms for Abstract and Pragmatic Content.

<table>
<thead>
<tr>
<th>Premise</th>
<th>Conclusion</th>
<th>Logic</th>
<th>Abstract</th>
<th>Pragmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>All A are B</td>
<td>All B are A</td>
<td>P</td>
<td>39</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Some B are A</td>
<td>N</td>
<td>22*</td>
<td>7*</td>
</tr>
<tr>
<td></td>
<td>No B are A</td>
<td>I</td>
<td>16*</td>
<td>2*</td>
</tr>
<tr>
<td></td>
<td>Some B are not A</td>
<td>P</td>
<td>33</td>
<td>47</td>
</tr>
<tr>
<td>Some A are B</td>
<td>No A are B</td>
<td>I</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Some A are not B</td>
<td>P</td>
<td>87</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Some B are A</td>
<td>N</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Some B are not A</td>
<td>P</td>
<td>78*</td>
<td>32*</td>
</tr>
<tr>
<td>No A are B</td>
<td>Some A are B</td>
<td>I</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Some A are not B</td>
<td>N</td>
<td>36</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Some B are A</td>
<td>I</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>No B are A</td>
<td>N</td>
<td>37*</td>
<td>15*</td>
</tr>
<tr>
<td>Some A are not B</td>
<td>Some A are B</td>
<td>P</td>
<td>74*</td>
<td>87*</td>
</tr>
<tr>
<td></td>
<td>No A are B</td>
<td>P</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Some B are A</td>
<td>P</td>
<td>72</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Some B are not A</td>
<td>P</td>
<td>72</td>
<td>69</td>
</tr>
</tbody>
</table>

Notes. N = Necessary, P = Possible, I = Impossible. Alpha was adjusted to p<.003 using the Bonferroni correction.

* Indicates a significant effect (p<.003) of the immediate inference type
Future research should continue to strive for greater ecological validity by incorporating reasoning problems and tasks that are designed to represent real world reasoning situations. Models of reasoning processes should account for an individual’s knowledge of material about which he or she is reasoning, as well as the specific demands the reasoning task makes. Exploring interactions among content, context, and participant’s knowledge will contribute to an understanding of reasoning processes.

References


APPENDIX

Pragmatic Inference Problems

1. GIVEN THAT
   All squirrels are mammals
   IT IS NECESSARY THAT
   Some mammals are squirrels
   □ Yes □ No

2. GIVEN THAT
   Some 9th graders are freshmen
   IT IS NECESSARY THAT
   Some freshmen are 9th graders
   □ Yes □ No

3. GIVEN THAT
   No pure water will remain liquid at 32˚F
   IT IS NECESSARY THAT
   Some pure water does not remain liquid at 32˚F
   □ Yes □ No

4. GIVEN THAT
   At a company, no managers are college graduates
   IT IS NECESSARY THAT
   At a company, no college graduates are managers
   □ Yes □ No

5. GIVEN THAT
   Some students who study hard do not perform well on the test
   IT IS NECESSARY THAT
   Some students who study hard do perform well on the test
   □ Yes □ No

6. GIVEN THAT
   When a road is covered with ice, some cars do not slow down when the break is depressed
   IT IS NECESSARY THAT
   When a road is covered with ice, no cars slow down when the break is depressed
   □ Yes □ No

7. GIVEN THAT
   In an apartment building, some things that make people feel cool are not air conditioners
   IT IS NECESSARY THAT
   In an apartment building, some air conditioners are things that make people feel cool
   □ Yes □ No

8. GIVEN THAT
   In a high school, some basketball players are not tall people
   IT IS NECESSARY THAT
   In a high school, some tall people are not basketball players
   □ Yes □ No

9. GIVEN THAT
   In a box of matches, some matches that were struck are matches that were lit
   IT IS NECESSARY THAT
   In a box of matches, some matches that were lit are not matches that were struck
   □ Yes □ No

10. GIVEN THAT
    Some cars in a parking ramp are Hondas
    IT IS NECESSARY THAT
    Some cars in a parking ramp are not Hondas
    □ Yes □ No

11. GIVEN THAT
    All 12oz drinks are soda cans
    IT IS NECESSARY THAT
    Some soda cans are not 12oz drinks
    □ Yes □ No

12. GIVEN THAT
    All glasses of pure water will freeze in a room with a temperature of 32˚F
    IT IS NECESSARY THAT
    All rooms with a temperature of 32˚F will freeze a glass of pure water.
    □ Yes □ No

13. GIVEN THAT
    All ocean water is salt water
    IT IS NECESSARY THAT
    No salt water is ocean water
    □ Yes □ No

14. GIVEN THAT
    At a police station, some dogs are cop dogs
    IT IS NECESSARY THAT
    At a police station, no dogs are cop dogs
    □ Yes □ No

15. GIVEN THAT
    No engineers are employed at the company
    IT IS NECESSARY THAT
    Some engineers are employed at the company
    □ Yes □ No

16. GIVEN THAT
    At a household, no paper is recycled
    IT IS NECESSARY THAT
    At the household, some recycled material is paper
    □ Yes □ No